ISSN (e): 2250 – 3005 || Volume, 07 || Issue, 3|| March – 2017 ||
International Journal of Computational Engineering Research (IJCER)

Recommender Systems Survey UsingData Partitioning Clustering

PratyushRanjan Mohapatra[1],Mohini Prasad Mishra[2] ,Kamalkanta Shaw[3]
[1,3]*Associate Professor, Department of Computer Science Engineering, Gandhi Institute For Technology (GIFT), Bhubaneswar*
[2] *Assistant Professor, Department of Computer Science Engineering, Gandhi Engineering College, Bhubaneswar*

Abstract
The Recommender systems have developed in parallel with the web for distinguish applications. They were initially based on demographic, content based and collaborative filtering .currently these systems are incorporating social information. In the future they will use implicit, local and personal information from the internet of things. This article provides an over view of recommender system as well as collaborative filtering methods and algorithms that how we can collect information for our day to day life using the clustering concept of data mining. It also explains their evolution, provides an original classification for these systems, identifies area of future implementation like collecting vital information about the user's use of things. And also develop certain areas of future implementation and develops certain area selected for past, present or future importance. Here in this paper the overall recommender system are enlighten and how they are effectively used for the day to day life and commercial life are used.
Keywords:*Recommender System, Collaborative Filtering, Clustering, Hybrid Filtering*

I. INTRODUCTION

Recommender Systems (RS) [4] collect information on the preferences of its users for a set of items (e.g., movies, songs, books, jokes, gadgets, applications, websites, travel destinations and e-learning material).The information can be acquired explicitly (typically by collecting users' ratings) or implicitly (typically by monitoring users' behavior, such as songs heard, applications downloaded, web sites visited and books read).RS make use of different sources of information for providing users with predictions and recommendations of items. They try to balance factors like accuracy, novelty, disparity and stability in the recommendations. The Recommender systems are basically subdivided into collaborative filtering, content based filtering, hybrid filtering.

Content-based filtering makes recommendations based on user choices made in the past (e.g. in a web-based e-commerce RS, if the user purchased some fiction films in the past, the RS will probably recommend a recent fiction film that he has not yet purchased on this website).

Demographic filtering is justified on the principle that individuals with certain common personal attributes (sex, age, country, etc.) will also have common preferences.

Collaborative Filtering allows users to give ratings about a set of elements (e.g. videos, songs, films, etc. in a CF based website) in such a way that when enough information is stored on the systems.

Hybrid filtering commonly uses a combination of CF with demographic filtering or CF with content-based filtering to exploit merits of each one of these techniques. Hybrid filtering is usually based on probabilistic methods such as genetic algorithms, fuzzy genetic neural networks, Bayesian networks, clustering and latent features.

With the purpose of being useful to the new readers of RS field, we have included in this survey some traditional topics: RS foundations,k-NearestNeighbors algorithm, cold start issues, similarity measure and evaluation of recommender system.

II. CLUSTERING

Clustering is a data mining tool has its roots in many application areas such as biology, security, business intelligence, and web search. The basic clustering technique includes partitioning methods, hierarchical methods, density based methods and grid based methods. Cluster analysis or simply clustering is the process of partitioning a set of data objects (or observations) into subsets. Each subset is a cluster, such that objects in a cluster are similar to one another, yet dissimilar to objects in clusters.

Clustering is also called data segmentation in some applications because clustering partitions large data sets into groups according to their similarity. Clustering can also be used for outlier detection
.Basically data clustering is under vigorous development.Abranch of statistics cluster analysis has been extensively studied, with the main focus of distance based clustering. Clustering is known as unsupervised learning also.

III. HOW TO USE CLUSTERING IN RECOMMENDATION?

As we know Recommendation systems (RS)[5] help to match users with items
– Ease information overload
– Sales assistance (guidance, advisory, persuasion…).Or we can say Recommender system recommends the perfect essentials for the user for his demand Or RS are software agents that elicit the interests and preferences of individual consumers […] and make recommendations accordingly. They have the potential to support and improve the quality of the decisions consumers make while searching for and selecting products online. Basically we know clustering is the data mining tool. The main purpose of clustering is to choose and collect the best similar goods for the user. We are wishing to use the concept of clustering in recommendations. Recommender system Recommend widely unknown items that users might actually like!.so by using the clustering concept we think we can definitely give a perfect streamline guidance to a novice user for his demand.
The most common and widely used recommendation system is
Collaborative filtering recommendation system.

IV. Collaborative Filtering Recommendation Collaborative filtering method allows users to give ratings about a set of elements (e.g. videos, songs, films, etc. in a CF based website)
insuchawaythatwhenenoughinformationisstoredonthesystems. The most prominent approach to generate recommendations that recommendation is used by large, commercial e-commerce sites and all the recommendations well-understood. Basically the collaborative filtering method used the "wisdom of the crowd" to recommend items which involves basic assumption and idea that Users give ratings to catalog items (implicitly or explicitly) and Customers who had similar tastes in the past, will have similartastes in the future will be the main achievement of this filteringmethod.

Collaborative Filtering methods play an important role in the recommendation, although they are often used along with other filtering techniques like content-based, knowledge-based or social ones.CF is based on the way in which humans have made decisions throughout history: besides on our own experiences, we also base our decisions on the experiences and knowledge that reach each of us from a relatively large group of acquaintances.Arecurrent theme in CF research is generating metrics to calculate with accuracy and precision the existing similarity for the users (or items). The most widely used algorithm for collaborative filtering is the k Nearest Neighbors (kNN).In the user to user version kNN executes the following three tasks to generate recommendations for an active user like (1) Determine k users neighbors (neighborhood) for the active user .(2)Implement an aggregation approach with the ratings for the neighborhood in items not rated by the user. (3)Extract the predictions from in step 2 then select the top N recommendations.

A. Similarity Measure in Collaborative Filtering
A metric or a Similarity Measure (SM) determines the similarity between pairs of users (user to user in CF) or the similarity between pairs of items (item to item in CF). For this purpose, we compare the ratings of all the items rated by two users (user to user) or the ratings of all users who have rated two items (item to item). Using the selected similarity measure, we produce the set of k neighbors for the active user. The k neighbors for user are the nearest k (similar) users to user.
Once the set of k users (neighbors) similar to active user has been calculated, in order to obtain the prediction of item i on user, one of the following aggregation approaches is often used: the average, the weighted sum and the adjusted weighted aggregation (deviation-from-mean).To obtain the top-n recommendations, we choose the n items, which provide most satisfaction to the active user according to our predictions.

V. BASIC PROBLEM IN RECOMMENDATION OR COLDSTART

Problem

The cold-start problem [2] occurs when it is not possible to make reliable recommendations due to an initial lack of ratings (when a new processing starts).The cold-start problem is often faced using hybrid approaches (usually CF-content based RS, CF- demographic based RS, CF-social based RS) We can

distinguish three kinds of cold-start problems:(1) new community. (2) New item (3) new user. The last kind is the most important in RS that are already in operation.

The new item problem arises because the new items entered in RS do not usually have initial ratings, and therefore, they are not likely to be recommended. In turn, an item that is not recommended goes unnoticed by a large part of the community of users.

The new user problem represents one of the great difficulties faced by the RS in operation. Since new users in the RS have not yet provided any rating in the RS, they cannot receive any personalized recommendations based on memory-based CF.

The new community problem refers to the difficulty, when starting up a RS, in obtaining, a sufficient amount of data (ratings) for making reliable recommendations.Two common ways are used for tackling this problem: to encourage users to make ratings through different means; to take CF-based recommendations when there are enough users and ratings.

Fig. 1: Recommender Systems

VI. EVALUATION OF RECOMMENDER SYSTEMS

The main contributions of this paper are: (1) A presentation of the explicit problems we find when we attempt to apply the commonly accepted framework for recommender systems, and try to measure, in a common way, the performance of any current type of these systems. (2) As a solution to the latter, we propose a new logical framework, general enough to incorporate in it every recommender system to date, while keeping it specific enough to obtain important and solid results. (3)Anew metric that measures the performance of any recommender system defined into the last framework limits. (4) A comparison between this new metric and the traditional ones.

Since RS research [5] began, evaluation of predictions and recommendations has become important. Research in the RS field requires quality measures and evaluation metrics to know the quality of the techniques, methods and algorithms for predictions and recommendations.

Evaluationmetricsandevaluationframeworkfacilitatecomparisons of several solutions for the same problem and selection from different promising lines of research that generate better results.

VII. ORIGINAL THEME OF THIS PAPER

Basically using the clustering method for recommendations is a new approach. We know that clustering having its vital concept by partitioning method, hierarchical method and density based method and grid based method.

In this paper we are trying to basically use the partitioning method of clustering for recommendation of the goods for the new users by taking the past experience of the neighbor user for the perfect recommendations. In recommender system Suppose a set of nobjects are present ,partitioning method constructs k partitions of the data ,where each partition represents a cluster and k<=n

.That is it divides the data into k groups such that each group must contain at least one object .On the other way partitioning methods conduct one level partitioning on data sets. The basic partitioning methods typically adopt exclusive cluster separation for the recommendation which is totally depend on users prediction which may be depends on item or user. Each object may belong to exactly one group. Most partitioning methods are distance based.

VIII. FUTURE WORK

We wish to use the hierarchical methods of clustering method incollaborative filtering approach in future.

IX. CONCLUSION

Recommender systems are providing to be a useful tool for addressing a partition of the information overhead phenomenon from the internet. When we are adding the most effective way of partition clustering in the recommendation concept that will be more successfulrequiring an exhaustive enumeration of the possible partitions.

REFERENCES

[1]. G.Adomavicious, J.Zhang,"On the stability of recommen- dations algorithms", In: ACM Conference on recommender systems, 2010
[2]. O.Arazy, N.Kumar, B.Shapira,"Improving social Recomm- ender systems", Journal IT Professional 11(4), (2009).
[3]. S.Abbar, M.Bouzeghoub, S.Lopez,"Context-aware recomm- ender systems:Aservice oriented approach", In: processing's of the third international workshop on personalized Access, profile management and Content Awareness in Databases,2009.
[4]. J.Bobadilla, F.Ortega, A.Harnando,"Recommender system survey".
[5]. F.Hernandez, E.Gaudioso,"Evaluation of recommender systems: A new approach, Expert system with applications".