IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 9, November 2014.
www.ijiset.com
ISSN 2348 – 7968

Product Recommender System Using Hybrid Search Technique By Integrating Collaborative Filtering (CF) And Round Robin Fusion Algorithm

Ashwini B. Waghchoure , Priya W.Tonde , Kishori P. Sonawane ,Priyanka T. Yewale
Computer Engineering, Pimpri Chinchwad College of engineering
Pune, Maharashtra, India

***Abstract*— Recommender systems (RS) have been widely applied in many commercial e-commerce sites to help users deal with the information overload problem. Recommender systems provide personalized recommendations to users and, thus, help in making good decisions about which product to buy from the vast amount of product choices. Many of the current recommender systems are developed for simple and frequently purchased products like books and videos, by using collaborative-filtering and content-based approaches. These approaches are not directly applicable for recommending infrequently purchased products such as cars and houses as it is difficult to collect a large number of ratings data from users for such products. Search-based recommender's cannot provide personalized recommendations. For different users, the recommendations will be the same if they provide the same query regardless of any difference in their interest. A simple user profiling approach is proposed to generate users preferences to product attributes based on user product click stream data. The user profiles can be used to find similar-minded users accurately.**
Two recommendations approaches namely Round-Robin fusion algorithm (CFRRobin) and Collaborative Filtering-based Aggregated Query algorithm (CFAgQuery), is used to generate personalized recommendations based on the user profiles.

Keywords—Collaborative filtering, Userprofiling,Recommender System, Product Recommendation,Personalization

I. INTRODUCTION

Websites and applications that offer their users or customers an item or a social element, have been trying to recommend them relevant content in order to their items/elements which they are interested in. Increasing the time that user spends on the website and increasing the interest of the user to the items in the website are the main reasons why the recommender systems are being used. Accuracy and time-efficiency are the most common problems of recommender systems.

People often observed that irrelevant recommendations of the products are provided. User sometimes even says that the websites like Amazon (Site 1) or Netflix (Site 2) even though they are considered as having the best recommender systems, also provides irrelevant recommendation. Accuracy and time-efficiency are the most common problems of recommender systems.

In our paper, we are using two combined approaches that are CF technique with the search based approach to solve above problem. One approach is to generate a query, which is called Collaborative Filtering based Aggregated Query (CFAgQuery), by aggregating neighbour users profiles. The query is then used to retrieve products. For the second approach, the attributes of each product preferred by the users neighbours are used to form a new query for use by the search-based approach to retrieve similar products. The new query captures neighbour users preferences and provides more detailed content to the query which may be interested but may have been missed by the target user when she/he submits her/his query as well as a data fusion technique to merge the response must be developed for merging the result retrieved by each query.

In our paper we propose the Round-Robin fusion algorithm to merge and rank the products retrieved from the multiple queries. Therefore, the product recommendations will be generated based on the new query.

The paper is organized as follows. First, the related work will be briefly reviewed in Section II. Then, in Section III the proposed user profiling approach and the two recommendation approaches will be discussed in detail. Finally, the system design will be given in section IV.

II. LITERATURE SURVEY

A. Collaborative filtering technique:

Collecting Behavior, activity or preferences for analyzing large amount of information given by user is done by collaborative filtering. Collaborative filtering [3], [4] is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating).

Collaborative filtering algorithms often require (1) users' active participation, (2) an easy way to represent users' interests to the system, and (3) algorithms that are able to match people with similar interests.

B. Content based approach:

Content-based filtering [5], methods are based on a description of the item and a profile of the user's preference. In a content-based recommender system, keywords are used to describe the items; beside, a user profile is built to indicate the

ISSN 2348 – 7968

type of item this user likes. In other words, these algorithms try to recommend items that are similar to those that a user liked in the past (or is examining in the present).

C. Hybrid approach:

Hybrid approach, combination of collaborative filtering and content-based filtering could be more effective in some cases. Hybrid approaches can be implemented in several ways: by making content-based and collaborative-based predictions separately and then combining them; by adding content-based capabilities to a collaborative-based approach (and vice versa); or by unifying the approaches into one model.

D. User profiling:

User profiling [2] is defined as to generate a user profile that represents the user's preference or interest in products or product attributes. User profiles can be exploited by the recommendation generating process to recommend new potentially relevant items to users.User profiling uses a user's data that can be gathered either explicitly or implicitly from the user.

III. MATHEMATICAL MODEL

A. User profiling and Query aggregation module:

In this module we are going to compare attribute and the feature from user query with history database, user profile data and dataset taken from various web sites.

For such a comparison we have to find similarity amongst them. For this we are using following steps:

1) Cosine similarity:

Similarity between items i & j is computed by isolating the users who have rated them and then applying a similarity computation technique. Cosine-based Similarity – items are vectors in the m dimensional user space (difference in rating scale between users is not taken into account).

$$\text{Sim}(i,j) = \cos(\bar{\imath}, \bar{\jmath}) = \frac{\bar{\imath}, \bar{\jmath}}{||\bar{\imath}||^2 * ||\bar{\jmath}||^2}$$

2) Correlation-based Similarity:

Using the Pearson-r correlation (used only in cases where the uses rated both item I & item j).

$$\text{Sim}(i,j) = \text{corr } i,j = \frac{\sum u \epsilon U (Ru,i - \bar{R}i)(Ru,j - \bar{R}j)}{\sqrt{\sum u \epsilon U (Ru,i - \bar{R}i)^2} \sqrt{\sum u \epsilon U (Ru,j - \bar{R}j)^2}}$$

R(u,i) = rating of user u on item i.
R(i) = average rating of the i-th item.

3) Adjusted Cosine Similarity:

Each pair in the co-rated set corresponds to a different user. (Takes care of difference in rating scale).

$$\text{Sim}(i,j) = \frac{\sum u \epsilon U (Ru,i - \bar{R}i)(Ru,j - \bar{R}j)}{\sqrt{\sum u \epsilon U (Ru,i - \bar{R}i)^2} \sqrt{\sum u \epsilon U (Ru,j - \bar{R}j)^2}}$$

R(u,i) = rating of user u on item i.
R(u) = average of the u-th user.

Generating the prediction – look into the target users ratings and use techniques to obtain predictions.

4) Weighted Sum:

How the active user rates the similar items.

$$P_{u,i} = \frac{\sum all\ similar\ items, N(si, N * Ru, N)}{\sum all\ similar\ items, N(|si, N|)}$$

5) Regression:

An approximation of the ratings based on a regression model instead of using directly the ratings of similar items. (Euclidean distance between rating vectors).

$$\bar{R}'N = \alpha \bar{R}i + \beta + \epsilon$$

R'(N) = ratings based on regression.

IV. SYSTEM DESIGN

The following architecture diagram shows the use of approaches

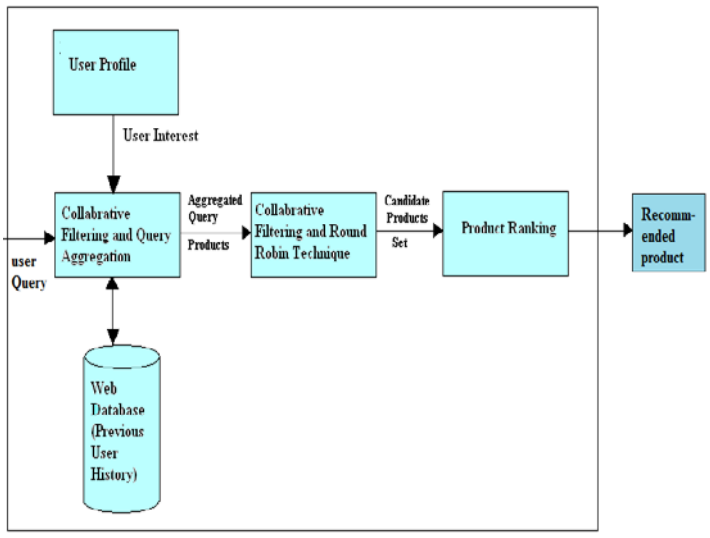

Figure . System Architecture

WEBSITES LIST

Site 1: Amazon
https://www.amazon.com
Site 2: Netflix
https://www.netflix.com

REFERENCES

ISSN 2348 – 7968

[1] N. Abdullah, Y. Xu, S. Geva, and M. Looi, Enhancement of infrequent purchased product recommendation using data mining techniques, in Proceedings of International Federation for Information Processing, World Computer Congress, Brisbane, Australia, 2010, pp. 57-66.

[2] M. Claypool, P. Le, M. Wased, and D. Brown, Implicit interest indicators, in Proceedings of the 6th International Conference on Intelligent User Interfaces, Santa Fe, New Mexico, USA, 2001, pp. 33-40.

[3] Y. Ge, H. Xiong, A. Tuzhilin, and Q. Liu, Collaborative filtering with collective training, in Proceedings of the 5th ACM Conference on Recommender Systems, New York, USA, 2011, pp. 281-284.

[4] Y. Hu, Y. Koren, and C. Volinsky, Collaborative filtering for implicit feedback datasets, in Proceeding of the 8th IEEE International Conference on Data Mining, Washington, USA, 2008, pp. 263-272.

[5] Ipek TATLI ,1395557,June,2009, **"**Food Recommendation System", Project Report.