Personalized Dynamic Recommendation System Using Hybrid Approach in Web Mining

Priyanka B. Tiwari[1], Krunal J. Panchal[2]

[1]CE, LJIET, GTU
Ahmadabad, Gujarat, India
priyankabtiwari26@gmail.com

[2]CE, LJIET, GTU
Ahmadabad, Gujarat, India
krunaljpanchal@gmail.com

Abstract - *Today's rapid exponential growth in the use of internethas given boom to all areas having access to it. This has helped creating great opportunities and options available for the businesses as well as customers. But this rapid growth has resulted complex situation for the customers to find the product and services they really want due to variation in prices. Also, providing accurate, valuable and personalized information for the customers, has become crucial for the businesses. This paper contains a survey on different recommendation methods used for E-Services by different websites. The paper also provides comparison of different techniques used for dynamic recommendation with strength and weakness and proposed future work. Our result shows that the personalized recommendation for E-services in a dynamic manner not only has direct impact on customer's interest and gross-sale, but also increase the loyalty to/for the customer.*

Keywords – Web Data Mining, Data Acquisition, Data Pre-processing, Data Cleansing, Data Mart Development, Collaborative Filtering, Item-Based Filtering, Knowledge Based Filtering, Demographic Filtering

1. INTRODUCTION

The trend of E-services hasgrown exponentially these days in a rapid manner due to the indulgence of internet in all areas that have access to it. These areas include various industries like E-commerce, Tourism, Medical, Education, Music, Film industry and many more. This trend of E-Services has given boom to marketplaces, leading vast and boundless commercial information for the scenario. This has helped creating great opportunities for businesses as well as customers. The customers have ease in finding theservices and products of their choice whereason the other hand the businesses have the opportunity to offer their services directly tothe customers in dynamic manner. However, the customer have to process massive information from the commercial gateway before selection of the actual thing that they need or want.

Recommender System, as kind of web-based support system, actively suggests a set of limited and ranked items from all available set of items without taking direct input from the user. For this, in real situation it uses some form of data mining.
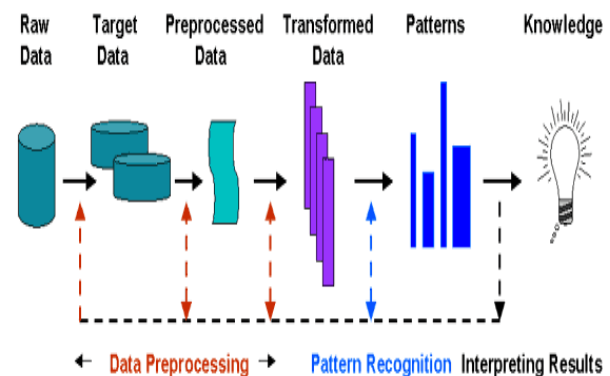

Figure1. Basic Data Mining Process [7]

2. BASIC RECOMMENDER SYSTEM

Before getting into the methods used in our survey, the general recommendation system and its purpose must be clear. The purpose of any recommendation is to provide the users, the list of selected items which either they have viewed or of those items which are highly ranked by others. The recommender system is the system that uses some method or algorithm for making recommendations dynamically from a class of data sets. The recommender system includes the following basic processes:

I. Web Data Mining

Web data mining is the application of data mining techniques to discover the patterns in web content, structure and usage in various applications.

A. Data Acquisition

Data acquisition refers to the collection of data for mining purpose. This data is collected from three main sources: (i) web server, (ii) proxy server and (iii) web client

B. Data Pre-processing

Data processing is cleansing, formatting and grouping of web log files into the meaningful session for the sole aim of utilizing it for web mining purpose.

C. Data Cleansing

Data Cleansing is the stage in which irrelevant/noisy entries are eliminated from log files, removal of entries with "Error" or "Failure" status.

D. Data Mart Development

Data Mart is a logical subset of data warehouse. During this phase, the pre-processed data is stored in the warehouse.

After the data is stored in the warehouse, different recommendation techniques can be implied.

II. Recommendation Techniques:

A. Content Based Filtering

The Content Based Filtering basically considers the attributes of the products. The algorithm recommends only best-matched products with similar attributes and the attributes get collected based on users' rating or the details.

B. Collaborative Filtering

The Collaborative Filtering approach basically considers the user-product interaction. This is based on collection and analysis of large amount of information available from user's behaviour, activity or preferences and also the prediction of what user would select based on similarity of other users.

C. Knowledge Based Filtering

The Knowledge Based technique is based on the explicit knowledge about interest of the user, item classification and recommendation standards. This is an alternative approach for Collaborative and Content based Filtering approach.

D. Demographic Filtering

The demographic technique is purely based on the users' information like gender, age, abilities, nationality, knowledge of languages, employment status, home ownership, location etc. The system does the recommendation based on the similarities in the demographic information.

E. Hybrid Approach

This approach implements the combination of two or more techniques of recommendation.

Issues in Recommendation System

1. The data in the web keeps changing, so there are issues in maintaining and updating the data warehouse.

2. The user's choice is different for each single user. So providing the recommendation in a personalized way it crucial.
3. Providing faster and more accurate recommendation to the client with desired qualities is complex.
4. Large number of in-frequent item set increase the space complexity and require too many data scans.
5. Uses complex algorithms in practical.

Due to these drawbacks, there is a necessity of making modifications in the recommendation techniques which are discussed in these papers. Let us go through the various techniques.

3. RELATED LITERATURE REVIEW

Numerous researches are available for web recommendation system. The term Web usage Mining was defined for the first time by Cooley and et.al which aims at predicting user's preferences and behaviour [10]. In [1], Prajyoti Lopes, Bidisha Roy, developed a dynamic recommendation system for all registered/unregistered visitors of the website. It was based on Action Based Rational Recommendation Technique which used formation of Lexical Patterns for products in each session and then its frequency and timestamp was counted. Based on the maximum frequency/timestamp, recommendation was made and it mainly focused on content based filtering. Another paper, presented recommendation to the user in efficient manner and schedule all activities of the users, using three different algorithms has namely HKB(Historical Knowledge Base), Cookies and Time Scheduling [2]. The data in the HKB was processed and ontology was made from where the matrix was used to find patterns for mining and ranking. For cookies and time scheduling algorithms also an ontological data was formed. And from that ontological data, pre-scheduling of tour was determined based on users search. Only disadvantage was that it resulted more accurately for predetermined schedules. In [3], Xuesong Zhao, KaifanJi, implemented the vitality of recommender system of present e-commerce for tourism e-commerce. For this Collaborative Filtering with implication of association rules had been used. In User based collaborative filtering, matrix was constructed for evaluations on different products, then nearest neighbour set was found using KNN and then recommendation result was generated for target users. For final recommendation results, Association Rules was applied to discover relation between large data sets by finding minimum support and minimum confidence thresholds.

In [4], Xiaosheng Yu, Shan Sun, showed two personalised recommendation technologies: 1) Content Based and 2)Collaborative Filtering and a framework had been designed based on web mining. In this the web log files and web contents were pre-processed to obtain customer and business document. Then the recommendation engine did recommendation based on the customer's browsing record. Customers' business document used collaborative filtering, where as web content document used content-

based filtering. Also Case Study of Amazon.com online bookstore was shown which used Item-to-Item Collaborative Filtering, which didn't match between customers but between the items purchased. It resulted in increasing loyalty to customers. Liu Hai- Ling , Li Jun-Husi, Peng Jun, in [5] with objective to provide online recommendation system for optimized travel routes with least costs by integrated application of GIS, data mining, presentation recommendation technology and Web and GIS Technology based on both user and attraction data. For this whole recommendation system consisted of three layers namely data source, business and presentation layer and the services are divided into three parts that is ordinary users system administrationand system function. And the system consisted of personal module, travel information, geographic information, and system module. And thus provides tailored travel plans. The whole system used matrix and then mining for recommendation. Aymen El Kalifi ,Firas Ben Kharrat, Rim Faiz, presented a method for recommendation using live user interaction for providing best list of hotels[6]. For this, recommender algorithm used multidimensional user behaviour. In the proposed system, Web Services API traced all the user action, which got stored in database layer with additional information like date, time,user, kind of action, items.And then in recommendation layer, actions were grouped by type and then prediction was calculated. Each prediction had specific weight. Finally one RS was returned to Web Services API, for results.Use of java API (open service) has been done to extract hotel data from website TripAdvisor. In paper [7], Jen Hsiang Chen, Kuo-ming Chao, Nazaraf Shah have focused inpredicting active users' interest by getting a set of highly rated tourism places collecting information about preferences and taste of other tourists. For this, Item-Based Collaborative Filtering has been used. Also, an approach to minimize the travelling cost by scheduling travelling path from a set of selected places using genetic algorithm has been done. In first step, process adopted Recommender Lab mechanism to search high rating of top places. In second procedure, it selected top 10 highly rated places, which applied genetic algorithm to minimize the tourists budget and then search for minimum cost path for covering maximum places. Another paper [8], focus on tourism e-commerce. For this, Collaborative Filtering and Association Rules of mining were implied for representing web mining based tourism e-commerce recommender system. Using CF, recommendation of top N list of products and services was done by finding nearest neighbour set and filtering implied item-based clustering or user-based clustering. Then, association rules were applied with an aim to increase the cross sales of tourism e-commerce. The whole recommendation composed of two parts: Offline Module and Online Module. Offline Module included data collection from sources, then pre-processing and data mining. The online part consisted of the whole recommendation engine that provided real-time personalized recommendation service. Honvia Zhang, Yuan Yang, in [9], focused on developing a personalized recommendation system model based on customer feedback, where results could be dynamically adjusted based on customer feedback information. The whole process consisted of two parts: Offline Mining and Online Recommendation. In Offline Mining, product data as well as customers' feedback was collected and processed. In online mining, the recommendation engine matched the mode according to customers' present conversation and produced personalised recommendation page set and got enclosed in the latest page of customer inquiry. The feedback included both Implicit and explicit feedbacks. The overall browsing behaviour of the user was stored in tabular form and degree of customers' interest was calculated from there based on time spent and number of hits.

A lot number of other methods and techniques have been implied by different researches and scholar.

4. RESEARCH GAP

The problem exists that therecommendation techniques that have been applied are having either user based, item based or action based in particular creating limitation for the users. And the model providing recommendation using both techniques together are based on predetermined scheduling which makes the system rigid.

5. CONCLUSION

The study shows that Action Based Rational Recommendation Technique is the best among all other recommendation techniques giving around 65-70% efficiency to unregistered and 80-85% to registered users. The emergence of E-Services recommendation makes a convenient, simple and easy to use system, solving predicament of customer and businesses which is resulted due to information overload. In future, we are working on development of a system applying the hybrid approach is possible to increase personalization to the unregistered user as well as registered ones in more dynamic manner.

6. REFERENCES

1) Lopes Prajyoti, and Bidisha Roy. "Dynamic Recommendation System Using Web Usage Mining for E-commerce Users." *Procedia Computer Science* 45 (2015): 60-69. doi:10.1016/j.procs.2015.03.086.
2) AgarwalJuhi, Nishkarsh Sharma, Pratik Kumar, VisheshParshav, AnubhavSrivastava, and R.h. Goudar. "Intelligent Search in E-Tourism Services Using Recommendation System: Perfect Guide for Tourist." *2013 7th International Conference on Intelligent Systems and Control (ISCO)* (2013): 410-415. doi: 10.1109/ISCO.2013.6481190.
3) Zhao, Xuesong, and KaifanJi. "Tourism E-commerce Recommender System Based on Web Data Mining." *2013*

8th International Conference on Computer Science & Education (2013): 1485-1488. doi: 10.1109/ICCSE.2013.6554161.

4) Yu, Xiaosheng, and Shan Sun. "Research on Personalized Recommendation System Based on Web Mining." *2010 International Conference on E-Business and E-Government (2010):* 346-349. doi: 10.1109/ICEE.2010.95.

5) Hai-Ling Liu, Jun-Huai Li, and Jun Peng. "A Novel Recommendation System for the Personalized Smart Tourism Route: Design and Implementation." *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)* (2015): 291-96. doi: 10.1109/ICCI-CC.2015.7259400.

6) Elkhelifi, Aymen, Firas Ben Kharrat, and Rim Faiz. "Recommendation Systems Based on Online User's Action." *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (2015): 485-490.* doi: 10.1109/CIT/IUCC/DASC/PICOM.2015.69.

7) Chen, Jen-Hsiang, Kuo-Ming Chao, and Nazaraf Shah. "Hybrid Recommendation System for Tourism." *2013 IEEE 10th International Conference on E-Business Engineering (2013): 156-161.* doi: 10.1109/ICEBE.2013.24.

8) Zhao, Xuesong, and KaifanJi. "Tourism E-commerce Recommender System Based on Web Data Mining." *2013 8th International Conference on Computer Science & Education (2013): 1485-488.* doi: 10.1109/ICCSE.2013.6554161.

9) Zhang, Hongxia, and Yuan Yang. "An E-Commerce Personalized Recommendation System Based on Customer Feedback." *2011 International Conference on Management and Service Science* (2011): 1-3. doi:10.1109/ICMSS.2011.5998970.

10) Cooley, Robert, BamshadMobasher, and Jaideep Srivastava. "Data Preparation for Mining World Wide Web Browsing Patterns." *Knowledge and Information Systems* 1.1 (1999): 5-32. Web.

11) Wang, Ya-Min, Xue-Ling Han, and Xiao-Wei Liu. "E-commerce Recommendation System Based on CBR and Web Log Mining." *2011 IEEE 18th International Conference on Industrial Engineering and Engineering Management* (2011): 311-315. doi:10.1109/ICIEEM.2011.6035165.

Priyanka Tiwari received the bachelor degree in information technology from Gujarat Technological University, in 2014. Perusing master degree in computer engineering from Gujarat Technological University.

Krunal J. Panchal. Presently, he is assistance professor in PG Department of LJIET, Ahmedabad.