Online based Content Recommender System based on Consumer Behavior Modeling

K. Thenmozhi[1], D. Sridhar[2]

[1]Assistant Professor
Department of Computer Science
Dr. NGP. Arts and Science College, Coimbatore, Tamilnadu, India
thennkrishnan@gmail.com

[2]Assistant Professor
Department of Computer Science
Dr. SNS Rajalakshmi College of Arts and Science, Coimbatore, Tamilnadu, India
ursridhardeva@gmail.com

Abstract: *Online usages are growing in popularly. Nowadays Most of the peoples are purchasing the products in online shopping. There are various online websites are available in the Internet. This paper presents the study of Online Based Content Recommender System Based on Consumer Behavior Modeling. Web Surfing has become a popular activity for many consumers who not only make purchases online, but also seek relevant information on products and services before they commit to buy. The proposed system used a web recommender that models user habits and behaviors by constructing a knowledge base using temporal web access patterns as input. Fuzzy logic is applied to represent real-life temporal concepts and requested resources of periodic pattern-based web access activities. The fuzzy representation is used to construct a knowledge base of the user's web access habits and behaviors, which is used to provide timely personalized recommendations to the user.*

Keywords: Web recommender, Association Rule Mining, Fuzzy logic, Patterns, Personalized Recommendations

1. Introduction

In www, there are various Online Shopping websites are available in the market. Each online shopping website uses different design concepts to attract the customers and make them to buy the products. Every product is categorized based on the keyword like laptop, mobile phone .Normally they are clustering the products based on Product type, Price and Quality. In the proposed research based on the consumer behavior that every successful consumers product selections are take it as input and provide the recommender system for new user.

Fast and robust clustering algorithms play an important role in extracting useful information in large databases. The aim of cluster analysis is to partition a set of N object into C clusters such that objects within cluster should be similar to each other and objects in different clusters are should be dissimilar with each other. Association analysis is the discovery of what are commonly called association rules. It studies the frequency of items occurring together in transactional databases, and based on a threshold called support, identifies the frequent item sets. Another threshold, confidence, which is the conditional probability than an item appears in a transaction when another item appears, is used to pinpoint association rules. Association analysis is commonly used for market basket analysis.

2. Association Rule Mining and Product Taxonomy

Given a set of transactions where each transaction is a set of items (item set), an association rule implies the form X) Y, where X and Y are item sets; X and Y are called the body

and the head, respectively. The support for the association rule X) Y is the percentage of transactions that contain both item set X and Y among all transactions. The confidence for the rule X) Y is the percentage of transactions that contain item set Y among transaction that contain item set X. The support represents the usefulness of the discovered rule and the confidence represents certainty of the rule.

Association rule mining is the discovery of all association rules that are above a user-specified minimum support and minimum confidence. Apriori algorithm is one of the prevalent techniques used to find association rules (Agrawal, Imielinski, & Swami, 1993; Agrawal & Srikant, 1994).Apriori operates in two phases. In the first phase, all item sets with minimum support (frequent item sets) are generated. This phase utilizes the downward closure property of support. In other words, if an item set of size k is a frequent item set, then all the item sets below (k 2 1) size must also be frequent item sets. Using this property, candidate item sets of size k are generated from the set of frequent item sets of size (k 2 1) by imposing the constraint that all subsets of size (k 2 1) of any candidate item set must be present in the set of frequent item sets of size (k 2 1). The second phase of the algorithm generates rules from the set of all frequent item sets.

Association rule mining has been widely used from traditional business applications such as cross-marketing attached mailing, catalog design, loss-leader analysis, store.

3. Review of Keyword based Systems

Several keyword-based recommender systems have been developed in a relatively short time, and it is possible to find them in various fields of applications, such as news, music,

e-commerce, movies, etc. Each domain presents different problems that require different solutions. In the area of Web recommenders, famous systems in literature are Leticia [49], Personal Web Watcher [62, 63], Skysail & Weber [70, 68], if Web [4], Am althea [66], and Web Mate [23]. Leticia is implemented as a web-browser extension that tracks the user's browsing behavior and builds a personalized model consisting of keywords related to the user's interests. It relies on implicit feedback to infer the user's preferences. For example, bookmarking a page is interpreted as strong evidence for the user's interests in that page. In a similar way, Personal Web Watcher learns individual interests of users from the Web pages they visit, and from documents lying one link away from the visited pages. It processes visited documents as positive examples of the user's interests, and non-visited documents as negative examples. Amalthaea uses specific filtering agents to assist users in finding interesting information as well. User can specify filtering agents by providing pages (represented as weighted vectors) closely related to their interests.

4. Customer Preference Analysis

The methodology applies the results of analyzing preference inclination of each customer to make recommendations. For this purpose, we propose a customer preference model represented by a matrix. The customer preference model is constructed based on the following three general shopping steps in online stores modified from works of Leets al. (2001):

1. **Click through:** The click on the hyperlink and the view of the web page of the product,
2. **Basket placement:** The placement of the product in the Shopping basket
3. **Purchase:** The purchase of the product, completion of a transaction.

A basic idea of measuring the customer's preference is Simple and straightforward. The customer's preference is measured by only counting the number of occurrence of {purchased products}. Hence, it makes sense to assign the higher weight to occurrences of purchased products than those of products only placed in the basket. Based on the customer preference we can show the product selling rate to the next consumer who want to buy a new product. So every actions of the consumers are monitored and every successful purchase information are stored for future analyze.
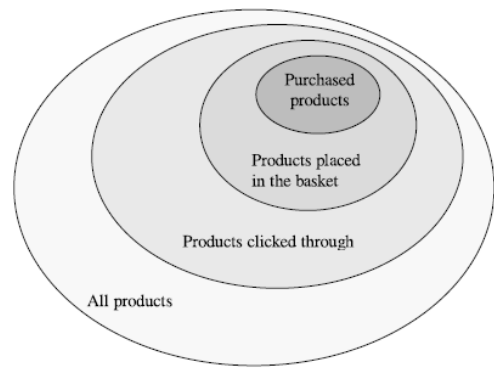
5. User Registration & Weblog Creation

The registration form is provided for each user for his identification. Each user is provided with a keyword to login. Web log is a data ware used to store the URL clicked by the user. The weblog enables easy identification of individual. It minimizes the time to analyze entire database. In this research is analyze of past consumer behavior. So, every consumer's information like name, age, gender, address, Email and contact number are gathered and stored in the database. Based on this field, we can easily analyze the requirements of the consumer. Every consumer must register their details then only allow them to buy the products. After the successful registration, a user account details are send to the consumer mail id.

In the research, we choose to use the fuzzy logic to find useful patterns. In constructing web usage Network we identify user's web access activities and construct a personal web usage Network from the personal access sessions of the user. Fuzzy theory is one of the appropriate techniques to describe such vagueness in information and incorporated into Formal Concept Analysis for representing both periodic attributes and resource attributes. To extract the correct search result records singular heading even if you have many acknowledgments from different component search engines can be merged into a single ranked list. This module is sometimes called an extraction wrapper. Since different search engines often format their results differently, a separate result extractor is usually needed for each tagged products

6. Personalized Recommender System

Recommender systems have the effect of guiding users in a personalized way to interesting objects in a large space of possible options. Content-based recommendation systems try to recommend items similar to those a given user has liked in the past. Indeed, the basic process performed by a content-based recommender consists in matching up the attributes of a user profile in which preferences and interests are stored, with the attributes of a content object (item), in order to recommend to the user new interesting items.

When a new customer wants to buy a product, after selecting the product the Recommender systems will show every product success rate in the market. The product sales percentage is nothing but the percentage of products purchasing rate of a particular product in the site. It was

automatically calculated based on the previous consumer behavior. The following formula show the way of calculating the recommendation of every product in the site that related to the particular keyword.

$$((n-r)/n) \times 100$$

Where,

r - Number of products not sold
s - Number of product sold
n - Total of number of products

Suppose the consumer wants to purchase laptops. In the shopping website, laptop1 and laptop2 are the two products available for selling in online. We are having 15nos (LAPTOP1) out of 50 laptop1 (n) of laptop1 and 20nos (x) out of 50(n) of Laptop2. If Consumer selected the laptop1 then the Recommender system shows the products sales percentage of Laptop1 and Laptop2 by using common keyword in laptop.

For Laptop1, product sales percentage rate 70% and Laptop2, product sales percentage rate 60%
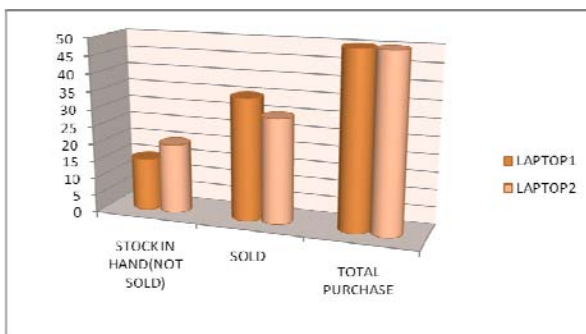
7. Conclusion

Thus the research paper depicts about the Web content Recommendation systems have become serious business tools and are re-shaping the world of e-commerce. Effective recommendations are a valuable service to the customers and a profitable service to the retailer. We hope that the work presented in this paper can be used to extend the knowledge in this area and open up new perspectives for future research.

References

[1] M.S.Yang," A Survey of fuzzy clustering" Mathl. Comput. Modelling Vol. 18, No. 11, pp. 1-16, 1993.
[2] Yoon Ho Choa, Jae Kyeong Kimb, Soung Hie Kima,"A personalized recommender system based on web usage mining and decision tree induction" Expert Systems with Applications 23 (2002) 329–342
[3] V. Bharanipriya , V.Kamakshi Prasad "Web Content Mining Tools: a comparative study" January-June 2011, Volume 4, No. 1, pp. 211-215
[4] Bamshad Mobasher, Robert Cooley, and Jaideep Srivastava. Creating adaptive web sites through usage-based clustering of urls. In Knowledge and Data Engineering Workshop, 1999.
[5] Bernardo Huberman, Peter Pirolli, James Pitkow, andRajan Kukose. Strong regularities in world wide web surfing. Technical report, Xerox PARC, 1998.
[6] C. Nuzman, I. Saniee, W. Sweldens, and A. Weiss, "A compound model for TCP connection arrivals, with applications to LAN and WAN," Computer Networks, Special Issue on Long-Range Dependent Traffic, vol. 40, no. 3, pp. 319–337, Oct. 2002.
[7] D. J. Hand, H. Mannila, and P. Smyth, Principles of Data Mining. Cambridge, MA: MIT Press, 2001.

Author Profile

Ms. K. Thenmozhi is Assistant Professor, Department of Computer Science at Dr.NGP Arts and Science College, Coimbatore, India. She received the B. Sc CS and M. Sc CS degrees from Dr.NGP Arts and Science College, Coimbatore, under Bharathiar University in the year 2004 and 2006. She received the M.Phil degree from Manonmaniam Sundaranar University, Madurai in 2007. Her area of interest is Data Mining.

Mr. D. Sridhar is Assistant Professor, Department of Computer Science at Dr.SNS Rajalakshmi College of Arts and Science, Coimbatore, India. He received the B. Sc CS from Sri Ramalinga Sowdambigai College of Science and Commerce, Coimbatore and MCA from Dr.SNS Rajalakshmi College of Arts and Science under Bharathiar University in the year 2006 and 2009. His area of interest is Data Mining.