A Novel E-Commerce Recommendation System Model based on the Pattern Recognition and User Behavior Preference Analysis

Ning Wang¹,²,³, Qiaoling Zhang⁴, Liejun Yang⁵, Mingming Chen⁶,⁷,⁸

¹ Dept. of Information and Mechatronics Engineering, Xiamen Huaxia University, Xiamen, Fujian
² Dept. of Automation, Xiamen University, Xiamen, Fujian
³ Fujian Province Engineering Research Center on New Generation of Information and Communication Technology and Wisdom Education
⁴ Dept. of Commerce, Xiamen Huaxia University, Xiamen, Fujian
⁵ Dept. of Computer Science, Ningde Teachers’ College
⁶ Dept. Of Information and Mechatronics Engineering, Xiamen Huaxia University, Xiamen, Fujian
⁷ Fujian Province Engineering Research Center on New Generation of Information and Communication Technology and Wisdom Education
⁸ University of Illinois at Springfield, United States

Abstract. This paper proposes the novel E-Commerce recommendation system model based on the pattern recognition and user behavior preference analysis. With the development of Internet and the generation of the huge amounts of data, information overload problem is increasingly serious, user drown in an ocean of data, it is difficult to effectively find themselves interested in the general information. Recommendation system technology was put forward and is widely used in this case, the recommendation system analysis of the user's past behavior records, while using the recommended algorithm automatically recommend users might be interested in the information to the general user. Standard utility of the log file format stored physical information about the client connection, if it can be some of the files stored in the mining analysis as can see the customer's behavior. From this starting point, we propose the new recommendation system architecture with the integration of pattern recognition algorithm that is innovative.

Keywords: E-Commerce, Recommendation System, User Behavior, Pattern Recognition, Preference Analysis.

1 Introduction

With the development of Internet and the generation of the huge amounts of data, information overload problem is increasingly serious, user drown in an ocean of data, it is difficult to effectively find themselves interested in the general information. Recommendation system technology was put forward and is widely used in this case,
the recommendation system analysis of the user's past behavior records, using the recommended algorithm automatically recommend users might be interested in the information to the general user. According to the recommendations adopted by the electronic commerce recommendation system technology, the electronic commerce recommendation system is mainly the following categories [1,2,3].

- Recommendation system based on collaborative filtering technology. Its advantage is able to discover new interesting and do not needs to consider characteristics of the goods, any form of commodity can be recommended. The disadvantage is that the user evaluation matrix for goods is very sparse, and with the increasing of the users of the system and the commodity, the performance of the system will be lower.
- Recommendation system based on content filter. Its advantage is simple and effective, and that recommends the fast response time. Defect is difficult to distinguish the quality of the commodity information and style, and can't discover new interesting for the user as can only be found and the user has been interested in similar products.
- Recommendation system based on the general data mining technology. Recommendation system based on the data mining technology through the browser user ratings with the method of data collection and recommendation, according to the results without the need for users to provide subjective rating information, so that the user is more convenient to use, at the same time with the higher credibility [4,5,6].

Data mining is the electronic commerce management activities that indispensable important tool, provide strong support for the right business decisions and reliable guarantee, but most of the traditional data mining methods can only analyze the homogeneous, isomorphic data, this for a large number of the heterogeneous text information on the Internet, log information, hyperlinks, etc. not applicable. Under this condition, in this paper, we propose the novel E-Commerce recommendation system model based on the pattern recognition and user behavior preference analysis.

2 User Behavior Preference Analysis

In the Internet in e-commerce, client access server or proxy server as will generate the corresponding server data, including log files and query data, and stored in a log file, the specific data source has the following several forms.

- The customer registration information. To better understand customers, must combine customer registration information and access log analysis, to draw more accurate judgment, and to provide more perfect service.
- The proxy server data. Its cache function reduces the Web server of network traffic, speed up the speed of Web page, at the same time will be a large number of users access to information through a proxy stored in the form of a log. On the analysis of such basic information can also help to customers browsing habits and target of induction and speculation.
- On the server side web data and log files. The purpose of the Web structure mining is found that the structure of the page, on the basis of classification
and clustering of the page in order to find authoritative page. Each provides
information resources on the WWW server has Web access log, as used to
record and accumulation of data about user interaction [7,8].

Standard utility of the log file format stored physical information about the client
connection, if it can be some of the files stored in the mining analysis as can see the
customer's behavior. For that example, by buying a product to the customer's domain
analysis, know from that country or region the number of buyers is more, accordingly
adjusted to this information in the electronic commerce online marketing strategy, to
increase business activities in the region or country. Standard combination log file
format is mainly refers to the meta information in log files, such as the beginning and
the general end of a session monitoring date, file size, etc.

E-commerce enterprise important degree of different, can according to customer's
goals through product innovation to improve target important product attributes and
weaken the appropriate weight small property, in order to enhance customer value. If
the design is a new and existing product with different properties of product, it not
only can better meet the customer's existing shopping the basic target of improving
the customer value can also create new type value for customers with new products to
attract the purchase, even with its new attributes to create a new market demand.
Improve can significantly increase the value of customers online market customer
satisfaction and loyalty, and positive influence on word of mouth and repeat purchase
online shoppers users. The figure one shows the principles.

Fig. 1. The Components of the User Behavior Preference Analysis

3 The Pattern Recognition and E-Business

The Web data mining in electronic commerce is that mainly about the customer
information and customer browsing and searching information on the website of the
digging, its purpose is to get to the user access patterns and browsing the behavior
analysis, so as to provide useful information for market decisions. The commonly
used data mining techniques in the e-commerce with path analysis, association rule
mining, clustering analysis and classification analysis, etc.
The PR techniques used for the e-business can be summarized as follows. (1) Correlation analysis technology. The purpose of correlation analysis is to dig out the hidden in the relationship between data, so as to find customers on the website access between the various files are interrelated, which found that the association rules. (2) Path analysis technology. Path analysis is a method of finding frequent access paths, it through to the Web server log file analysis customer access to the site visits, mined frequent access paths. Thus can understand customers love, of which pages to better design improvements, service to the customers. (3) Clustering analysis technology. In Web use mining, there are two types of clustering: user clustering and clustering of Web pages. User clustering mainly, all users are divided into several groups with the similar characteristics (or view) of user in a group, this kind of basic knowledge of electronic commerce and especially useful to provide users with personalized service. (4) Classification analysis technology. Classification analysis is based on the analysis of the sample data in the database, and make an accurate description for each category analysis model is set up or mining classification rules, and then use the classification rules to classify other data in the database [9].

For the mathematical description of the steps, we can summarize as follows.

\[
\text{Support}(X \Rightarrow Y) = \frac{|T \cup Y \subseteq T, T \in D|}{|D|}
\]

\[
\text{Confidence}(X \Rightarrow Y) = \frac{|T \cup Y \subseteq T, T \in D|}{|T, T \subseteq D|}
\]

Pattern discovery is to use various methods to find hidden patterns and rules. To deal with after pretreatment of data to get the corresponding transaction database, data mining on the basis of this, to two aspects: one is the transaction database sorting into corresponding to a certain mining technology of data storage form as the second is to use data mining algorithms mining out valid, novel, potential, useful, and ultimately understandable information and knowledge. Users of the service satisfaction is related to the performance of the system, users stay on the page for a long time, may not be to have a great interest in the page, the server status and due to display the page content transmission may cause the user to open the web page of waiting for a long time. For this reason, you need according to the proxy server access log to predict the user's request in time and space distribution.
4 Conclusion and Summary

This paper proposes the novel E-Commerce recommendation system model based on the pattern recognition and user behavior preference analysis. In the increasingly competition in the e-commerce market, and any information relating to the consumer behavior is invaluable for operators. Companies learn the user's access pattern is very important. The definition of the Web use mining is presented and a complete model framework, then the main steps of the latest research progress in the Web use mining conditions do the detailed elaboration and analysis, including: data acquisition, data preprocessing, patterns discovering, pattern analysis. Under this circumstance, this paper integrates the pattern recognition algorithm to propose the new perspective of the recommendation method that is innovative.

References


