Application of Data Mining Technology in E-Commerce

Cheng Yu, Xiong Ying
(Computer School, Hubei University of Technology, Wuhan 430068, China)

ABSTRACT: ***The rapid development of the Internet has brought an opportunity for the development for enterprises. Web mining in the status of e-commerce websites have became more and more important. In the e-commerce, the data mining is helpful of the discovery to trade development tendency, of the correct decision-making made by the enterprise. This article mainly summarizes the present electronic commerce's Web data mining method, and classifying to its objects, analyzing the network data mining's function,described the use of electronic commerce website in a number of web mining application, and offering the reference for the future electronic commerce practice, especially its software's development and the application.***

KEYWORDS:E-Commerce, Web data Mining, Web Log, Data Preprocessing

I. INTRODUCTION

E-commerce refers to individuals and businesses through the internet network, using digital electronic exchange in business data and business operations. At present, there are online business advertising and electronic clearing house, online ordering, banking, payment and settlement, such as the form of various types of e-commerce. For its low-cost, convenient, safe, reliable, free from constraints of time and space advantages, E-commerce give prominence to the global trade gradually.

Data Mining is accompanied by the development of data warehouse technology. Data mining is to help business users to deal with a lot of data and subsequently to discovery its regularity implied. It requires a lot of incomplete noisy, fuzzy and random data to be extracted from potentially useful information and knowledge. Sometimes the process of data mining, also known as that of discovering knowledge, which is exiting in e-commerce data mining, is a data mining technology, extracted from some interesting and useful web file or web service. It is a comprehensive technology related to Internet technology, artificial intelligence, computer languages, information science, statistics and other fields.

II. CLASSIFICATION OF WEB DATA MINING OBJECTS

Web Data Mining is applied in web environment by data mining technology, which is potential and useful model or information discovered from lots of data of web files and websites. Web Data Mining carries on web deposit-fetching model, web structure and rule, and dynamic seeking. Generally, Web Data Mining should be divided into three categories: web content mining, web structure mining and web usage mining. Web content mining is process of fetching knowledge from file or its description. Web structure mining is to deduce knowledge from web structure and super link relation. Web usage mining mainly withdraw useful model from web-visiting. Web mining categories are shown as figure 1.

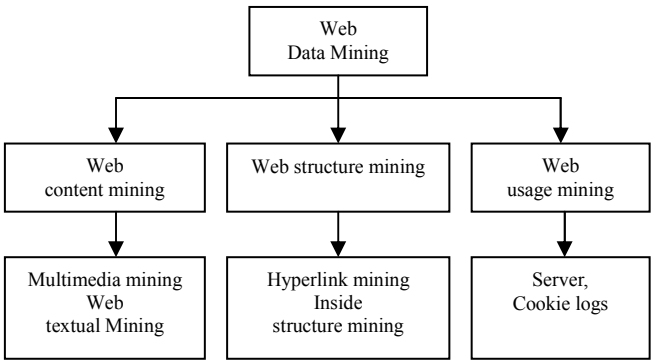

Figure 1. web mining categories

III. E-COMMERCE METHODS IN DATA MINING

Different mining goals of e-commerce can be used with different methods of data mining. Data mining methods are a lot of the major categories, including three categories: statistical analyses, knowledge discovery and the mining methods based on forecasting model.

A. Statistical Analysis

Statistical analysis of data is mainly used to check the rule of mathematics, and then use statistical models and mathematical models to explain these laws. Methods commonly used include linear analysis, nonlinear analysis, regression analysis, logistic regression analysis, and multivariate analysis, as well as time series analysis. Statistical analysis methods help to find the relationship between large amounts of data, for example, when asked to identify the model of sequence data, abnormal data to help select the appropriate data for the statistical model including the multi-dimensional table, partition, sort, at the same time appropriate to generate a chart provided to the analysts, statistical function is the appropriate statistical tools through to complete the regression analysis, multivariate analysis, data management for finding detailed information, visiting the subset, and deleting the redundant.

B. Knowledge Discovery

Knowledge Discovery originated in artificial intelligence and machine learning uses a data-search process to collect data information, which indicates that the relationship between the data elements and models, so as to find business rules and business fact. The use of data visualization tools and views is contributed to the development of analysis tools for past data mining in order to further enhance the capacity of data mining. Other data

978-0-7695-3930-0/09 $26.00 © 2009 IEEE
DOI 10.1109/IFCSTA.2009.76

mining methods such as visualization systems can be given with a multi-variable graphical analysis of data to help business analysts in carrying out knowledge discovery.

C. Mining Methods Based on Forecasting Model

Mining methods based on forecasting model is to applied machine learning and artificial intelligence into data mining system. Prediction model is based on the assumption: the consumption behavior of consumers has a certain repetitive and regularity. This enables analysis of business can be stored in a database to collect the transaction information to predict consumer behavior. According to specific consumer behavior, businesses marketing will be able to focus on the part of consumer.

Ⅳ. DATA MINING PROCEDURES IN ELECTRONIC COMMERCE ENVIRONMENT

Generally, the process of web data mining in E-commerce is consist of data source, data preprocessing, model construction and analysis. The detail procedure is shown as figure 2.

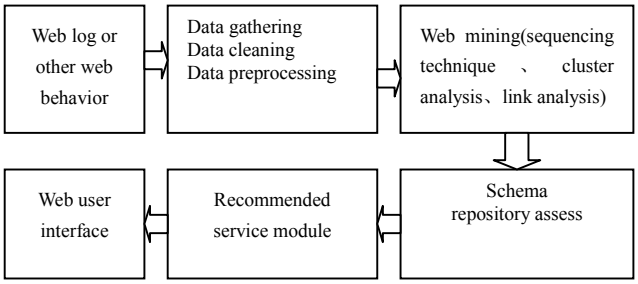

Figure 2 . Data Mining Procedures of E-Commerce

A.Select the Data Source

This step is responsibility for exacting analyzing data from available source, setting into the same analysis as a data-analysis source. The required data in data mining of e-commerce mainly comes from two aspects: the customer's registration information, including customer background information and customer transaction data in the past, and the click-stream of visitors stored in the server, which is used to study the behavior of customers.

B. Data Preprocessing

Data pre-processing clean Data up through the words filter, derivative processing, data's scatter, data sample record, summary, additional and sort, so as to settle up missing of data, redundancy, inconsistency and so on. A major task of data pre-processing is to organize the web site of the original log into a affairs' database for the use of data mining. This is the most critical data mining stage of e-commerce because of its complexity of the data.

C. Mining model Construction and Data Mining

The key to success is transferring data into a real algorithm data mining model for analysis. Construction of the model is related to researched e-commerce business. For example, the goal of model will reflect the various relevant factors to which all ages' customer respond, if the research objective is to analyze the customer's interesting in a commodity's response. Model finished also requires the accuracy, performance and understandable for comprehensive study. During the actual operation, the mining algorithms should be select to identify a suitable mining so that the data from the mass is certified as a valid, novel, potential, useful and ultimately understandable knowledge. Commonly EC-used data mining is some technology such as association rules, sequential patterns, classification and clustering techniques.

D. Analysis and use of the results

It is necessary to explain the results of the mining and evaluation. The results of EC-oriented data mining be suitability or not, depending on the problem to be solved. Just considering the accuracy of a model is no use. The most important is the actual applicability of the model. It is worth noting that it is only a estimate, detection can only be true in the actual applications of e-business, even if the choice of a variety may be made, such as mathematics or any other method of the objectivity of the evaluation, because the data mining model with a certain period of time may be more short-term rule. Users should assess the results of the mining, if satisfied, ending the mining process. Otherwise, the request of re-mining will go on. Model of effective Confirmed by mining tested can be applied in the knowledge managerial decision to enhance the competitiveness of enterprises.

Ⅴ. WEB MINING FUNCTION

Through the collection, processing and handling with a large amount of information related to consumer behavior, specific consumer groups or individuals habits, trends and demand could be identified to infer the corresponding consumer groups or the individual's future consumer behavior, and then to make decision of specific marketing, cost savings, improving efficiency, so as to bring more profit for enterprises.

1). Website optimization. Web designers no longer totally rely on the guidance of experts to design a qualitative website, but according to the visitor's information to modify the design, site structure and appearance. Website page content and links on items like supermarket display, together with for sales. Webmasters can also structure the page according to model the majority of visitors' browser to and cut their users web information space, so as to offer convenience for the majority of visitors.

2). Personalized website design. Emphasis on personalized identification information can bring customers in their own ways to access the site. To user frequently visiting place, personalized banner ads can be offered for personalized marketing services.

3). Retain old customers. E-commerce operator can be informed of the personal-loving visitors to understand more fully the needs of customers. Providing with products according to the unique needs of each category (or even a) customer can be helpful of improving customer satisfaction to achieve the purpose of retaining customers.

*4). Mining potential customers. T*hrough the analysis and explore the regulation of web log, the visitors may be classified to determine the classification of the key attributes and relationships, and then to identify potential customers of e-commerce to improve the quality of customer service.

5). Extend the residence-time of customers. In e-commerce, to make customers stay longer on the site should be aware of customer interest and needs of the dynamic demand, to which in accordance recommendation and adjust of website pages is necessary to provide unique some commodity information and advertising for customer satisfaction.

6). Reduce operating costs. Through Web mining, the company can analyze the future behavior of customers for targeted marketing, and decide the location of advertising based on interest in a product the visitor's browser so as to increase return rate. A reliable market feedback could be applied to reduce the company's operating costs.

7). Enhance the security of e-commerce. Web-content mining includes background information registering in the transaction database. It is a very important role for Clients to register their information in e-commerce, especially in the security, or on the restrictions for customers to access.

8). Improving the competition of enterprises. According to the customer not only the historical data, analysis of potential target markets and optimizing the business model of e-commerce site, can predict the demand trend ,or assess the demand for change, so as to help improving the competitiveness of enterprises.

Ⅵ. SUMMARIZATION

In short, valuable knowledge can be found because of the application of data mining technology in e-commerce. Based on this knowledge, Business can grasp the customer dynamics, track changes in the market, and target to make a correct decision-making, such as improving the site, personalizing users' pages, or retaining the preferential policies to the quality customer and so on. But in e-commerce for Web data mining there is still plenty of problems to be solved. For example, how to solve the storage of different countries in different regions of the Semantic Web data inconsistency, how to provide with more secure, and efficient service, there are a lot of work to do.

REFERENCES

[1] Hao Xincheng, Zhang Degang, Zhao Hai: *E-commerce data mining research.* Small Micro-computer Science[J], 2007(7) pp786~787

[2] Zhao Huanping: *Web data mining and its application in e-commerce*. Fujian Computer [J].2008 (1) pp167

[3] Shi Yan: *Applications of web mining techniques in e-commerce*. Scientific and Technological Information and Economic Development [J] 2006 (7) pp255 ~ 256

[4] Ling Chuan-Fan: Web mining techniques applications in e-commerce. Intelligence Journal [J] 2008 (1) pp93 ~ 94

[5] Agrawal R, Srikant R. Fast algorithms for mining association rules in large database [A]. Proceedings of the 1994 International Conference on VLDB [C]. San Francisco : Morgan Kaufmann Publishers , 1994.pp487- 499

[6] Witten I H, Frank E. Data Mining : Practical Machine Learning Tools and Techniques[M]. 2005.

[7] Margarent H. Dunham . Data Mining: A Tutorial-Based Primer 2005.

[8] J.B. Schafer, J.A. Konstan, and J.Rie1, Recommender Systems in E-Commerce. In ACM Conference on Electronic Commerce (EC99),1999.

[9] G. Salton and M.E. Lesk. Computer Evaluation of Indexing and Text Processing. Journal of the Association for Computing Machinery,1996,67-79