

# WEB data mining applications in e-commerce

Yonghua Zhao

Department of Computer Science and Technology,  
Wuhan University of Technology,  
Wuhan, China  
abcyonghua@163.com

Hong Lin

Department of Computer Science and Technology,  
Wuhan University of Technology,  
Wuhan, China  
linhong6561@126.com

**Abstract** — Web data mining is a new technology that can find valuable model and information in mass data. With the development of the Internet, e-commerce has a strong trends, application of web-based data mining in e-commerce has become a hot topic. In this paper, the Web data mining and e-commerce is combined, it introduces the method and process of Web data mining in e-commerce, and describes the Web data mining applications in e-commerce.

**Index Terms** —Web, data mining, e-commerce.

## I. INTRODUCTION

With the rapid development of Internet in China, the industry's business model has changed. At present, great process has been made in Web e-commerce platform for its convenience and transaction fast. Competition for users is the key factor for e-commerce business in the increasingly fierce competition. If you can grasp customer needs, develop targeted business activities, not only can provide convenient trading mode and a wide choice for customers, but also make the e-commerce business to retain customers better. One of the solutions is Web data mining technology. We can get the user behaviour from the browsing behaviour of customers on Web and further analysis, then to find a solution. This will allow sellers know more about their customers' needs, and provide personalized according to customer preferences, then obtains the competitive advantage

## II. THE DATA RESOURCES IN THE E-COMMERCE PLATFORM

In e-commerce activities, it will leave the data on Web when customers browse information or do other acts, these data will be e-commerce Web data mining data source.

The first is the customer information, customer information includes not only the customer's registered personal data, but also includes customer's order information, customer's individual requirements and problems, as well as the customers to browse merchandise records and visited pages and so on behaviour information. Such information can help to analyse customer preferences and background, then forecast the future purchase behaviour of customers; The second is the commodity information on e-commerce website, each commodity has its own information, such as price and parameters, it is to be presented to customers through Web forms, and therefore also the product information is the Web mining object, such information can be found in goods between pages hidden links, and optimize the structure of the page;

Finally is the server information, server data will be generated on the server when client access the server, the data can be divided into server logs, error logs, cookie logs, through the analysis of these log files, you can optimize server system better, improve the operation efficiency, enabling customers to obtain a better user experience .

## III. THE CLASSIFICATION OF WEB DATA MINING

The object Web mining process including static Webpage, Web database, Web structure, user records and other information. In general, Web mining can be divided into three categories:

### A. Web Content-Mining

Web content-mining is the process that to obtain the potential, valuable knowledge or model from the document content , description information or search results on Web page, at the same time can also be used for obtaining valuable information from mining of Web organization structure and link relationship. The commonly used method of Web content-mining are WebOQL and Aboy.

### B. Web Structure-Mining

Web structure mining is the mining structure mode link potential Web, is a mining structure of Web page directory path hyperlink relationship, document structure and the directory path structure in URL. In the Web space, the useful information not only is contained in the content of the page, but also included in the structure of the page, through the analysis of the hyperlink structure of the, if hyperlinks pointing to it more, then the page is important, so you can change the search path on the basis of it.

### C. Web Usage-Mining

Web usage-mining is the process that discovering the access pattern by mining the Website log files and related data. Each server in the WWW has retained the transaction data that generated between user and network in the interactive process, namely the Web server log file. Through the analysis of these data can help you understand user behaviour, improve the structure of your website, to provide users with personalized service

#### IV. THE PROCESS OF WEB DATA MINING IN E-COMMERCE

The process of Web data mining in e-commerce consists of three main stages: data preparation and pretreatment, pattern discovery and pattern analysis:

##### A. Data preparation and pretreatment

Data preparation and pretreatment is a stage that to find the data source for pattern discovery. The raw data obtained from the data source of the e-commerce, is not only huge, but also there may be a lot of noise data, redundant data and incomplete data and so on, the direct mining is very difficult. So we need to extract the data in the database, and then certain cleaning, to eliminate the useless data, retain the useful data, this process mainly includes the data integration, data cleaning, user identification, user session identification, access path supplement and transaction identification and so on.

##### B. Pattern discovery

Use of useful data generated by the data preprocessing stage to modeling is the work of this stage. Firstly, according to the type of data and data scale to assume reasonably, establish the appropriate data model, and then select the appropriate tools and methods to verify the hypothesis. Finally find the valuable information.

##### C. Pattern analysis

Filter out less interesting rules and patterns that generated by the pattern discovery stage is the task of this stage. Through the analysis of these to find useful information, and through the verification of on-line information, combined with customer personal information, to find out the valuable market information or potential market value.

#### V. WEB DATA MINING TECHNOLOGY IN E-COMMERCE

The Web data mining technology in e-commerce divided into the following five:

##### A. Path analysis

Through the analysis of log information, determine the path which was the most frequently accessed in a Web site is called path analysis, there are a number of other relevant path information can be obtained, then you can change the design of the site structure by analysis these.

##### B. Association rules

Association rules are to find the correlation between client and access documents on the web. To find out the relationship between the customer purchase behavior, constructs the relational model, and then the website can be organize site better, so that the burden of users filtering information will be

reduced, and we can recommend to clients according to the users' current purchase behavior.

##### C. Classification rules

Classification rules are that Classification based on some attribute of the data. After classification, you can conduct business according to the characteristics of this kind of customer, provide the personalized service. For example these can be classified according to different regions or different age.

##### D. Sequence pattern

Sequence pattern and association rules is similar, both are relation between the data. But sequence pattern focus on the relationship between the data before and after the sequence. Finding the internal transaction mode that "Some items to follow some other items", through the discovery of sequential patterns, you can predict customer access mode, select the targeted page, to meet the specific needs of visitors.

##### E. Cluster analysis

According to specific rules to classify the records which are unclassified is called cluster analysis. It will put the whole database divide into different groups, there are obvious differences between groups, but the data of same group are similar. For example, some customers often visit domestic appliance, they may be customers those would be married. The service to them is different from "students", "white collar", "farmers" and so on.

#### VI. APPLICATION OF WEB DATA MINING IN E-COMMERCE

The appearance of e-commerce, changed the business philosophy, has brought the huge change for various industries. The combination of data mining technology and e-commerce can greatly improve the capability of enterprises to extract information and information management, enterprise information resource value to be fully reflected. Mainly in the following points:

##### A. Keep old customers, develop new customers

With the business philosophy that "customer-focus" are getting support among the people, analysis of customer, understand customer and guiding customers' needs has become an important topic for the enterprise management, through the Web data mining, e-commerce sites can understand user preferences, adjust the Web page, to provide personalized service for the users, improve customer satisfaction, prolonging the dwell time of customers at the site, to keep the old customers; through the analysis of the new user access, mining valuable potential customers, implement the strategies for these customers, to guide them to become a real user, reflected in the e-commerce is both of the order quantity and the economic benefit are increased.

### B. Enterprise resource optimization

Cost saving is the key to corporate earnings. Through the data mining technology, you can grasp the enterprise resource information in real-time, accurate, through the analysis of historical financial data, transaction data and other information, you can find the key point of enterprise resource and the ratio of output to input of the main activities, so that you can provide a basis for enterprises to optimize the allocation of resources. E-commerce enterprise can obtain reliable information, to analyze customer behavior in the future, for marketing and promotion through data mining; by understanding the customer's browsing, to decide the advertisement position, increase advertising investment rate of return, so as to reduce operating costs, improve the competitiveness of enterprises.

### C. Improve the site design

By mining the customers' behavior record and feedback, to improve the system design for designer. Using association rules in the Web, you can dynamically adjust the site structure for different customers, so that customers can easily access to the page they want to. This will left a good impression to customers, and increase the probability of the next visit.

### D. Clustering of customers

In e-commerce, customer clustering is the main component. By grouping similar browsing behavior of customers, and analysis of the common characteristics of customers in the same group, can help the e-commerce operator understand customer needs better, make appropriate adjustments to the Webpage content, and to meet the internal needs of customers in many ways, do the best to provide high quality, good service for customers.

## VII.CONCLUSION

Web data mining is a very fast development area. This paper introduces the application of Web data mining in e-commerce data, and related concepts. Web data mining technology will help enterprises to find valuable knowledge in the network information resources. It can predict the consumption trend of customers, help businesses find and retain customers, adjust market strategy, make correct decision, promote the development of electronic commerce. But there are still some problems in the combination of web data mining technology and electronic commerce, these problems remain to be further explored and solved.

## REFERENCES

- [1] Jinyong Liu. WEB data mining research applications in E-commerce[J]. Network security technology and Application, 2013 (9): 25-26
- [2] Ning Luo, Jungang Xu. Application of Web data mining in E-commerce[J]. Electronic technology. 2012, 4: 005.
- [3] Hongxia Zhao, Yumei Zhao. Application of Web data mining in E-commerce[J]. Financial computer of Southern China. 2004, 11: 96-99.
- [4] Xianchun Zou, Zhong Xie, Yanhui Zhou. El-commerce and Web data mining[J]. Computer application, 2001, 21(5): 21-23.
- [5] Li Kan, Hao Pan. Application of Web data mining technology in E-commerce[J]. Computer knowledge and technology, 2010, 4: 816-81.
- [6] Ru Nie, Senping Tian. Web data mining and its application in E-commerce[J]. Beijing: Computing technology and automation, 2006, 12: 237-239.