

Application Research in E-Commerce Recommendation System of Web Mining Technology

Wu Zhihong

School of Economic and Management
Shengyang Aerospace University
Shengyang, China
Wuzh@sau.edu.cn

Wang Qingjun

School of Economic and Management
Shengyang Aerospace University
Shengyang, China
32582669@qq.com

Abstract—Web Data Mining refers to the process of extracting the implicit and previously unknown information and knowledge with potential values from the huge, incomplete, unpurified, unclear and random practical and applicable data on the Internet. From the business sense, Web data mining is a new technology of processing business information based on the Internet and the extraction, transformation, analysis and other modeling processes of huge business data from the commercial database on the Internet so as to extract the key data of auxiliary business decisions. Through the application of the Web data mining technology, browsing habits and buying preferences of a large number of users can be analyzed, thus precise marketing like directional recommendation can be carried, and this is the so-called e-commerce recommendation system.

Keywords- informatization; data mining; directional recommendation; big data analysis; precise marketing

I. INTRODUCTION

In the information age, there is a complex variety of information, which is invading our work, study and lives. And this uneven information is also interfering with our thinking and analyses with an inexhaustible number. In order to analyze these vast amounts of information so as to get the data results we want, we have developed the data mining technology.

Based upon actual needs and by combining theories and practices, this paper elaborates on the current application of Web data mining technology as well as the development of e-commerce recommendation system and the application after combining these two, which to some extent helps to solve practical problems.

II. OBJECTIVES AND SIGNIFICANCE OF RESEARCH IN E-COMMERCE RECOMMENDATION SYSTEM OF WEB MINING TECHNOLOGY

The E-commerce recommendation system analyzes users' buying habits based upon the users' information it has acquired, and then makes recommendation automatically, thus shortens users' time of searching goods and finally makes users begin to rely on it and the results. After users have taken likes for the convenience of this recommendation system, they will establish a long-term

and stable relationship with the shopper, so that we have achieved the goal of customer maintenance.

At present, the development of the recommendation system has reached the stage as how to combine with mining technology. It is of paramount importance for the data mining technologies especially the Web mining technology as how to deeply exploit potential users' information and find their potential needs. And this paper is just about the application of Web data mining technology in e-commerce recommendation system.

III. INTRODUCTION TO RELEVANT THEORIES OF WEB MINING TECHNOLOGY

Data mining refers to the finding process of extracting implicit and previously unknown relevant information and knowledge of potential value from a huge amount of unpurified, unclear and random applicable data.(as excerpted from Baidu Encyclopedia) Web mining, the application of data mining on the Web, refers to the process of extracting the useful models you have been interested in as well as implicit knowledge from Web-related data resources and then forming an integral research and application body of Web structure mining, Web content mining and Web usage mining by combining traditional data mining technologies with the Web and comprehensively using technologies in many fields, e.g. statistics, computer network, database & data warehouse, visualization, etc.

A. Operating Principles of Web Data Mining Technology

Generally speaking, Web mining technology refers to the process of pre-treating, model discovery and model analysis of target data sets of Webpage structure and content. Target data set refers to relevant data extracted from the Web data sources according to users' requirements and in Web mining, data will be extracted mainly from these data. Pretreatment is an important part to remove obviously wrong and redundant data from the target data, and then screen the selected data and turn them into effective form, making data mining more effective. Model discovery is to select appropriate models and parameters according to the mission requirements to establish the data mining algorithm and show the clear situation by using some methods (as shown in the Figure).

Model analysis is to explain and evaluate the discovered models, repeated extract them when necessary, and finally present to users the mind results in a simple and easy way.

B. Features of Web Data Mining Technology Structure

Web data mining technology can be divided into: Web structure mining, Web content mining and Web usage mining. As shown in Figure 1:

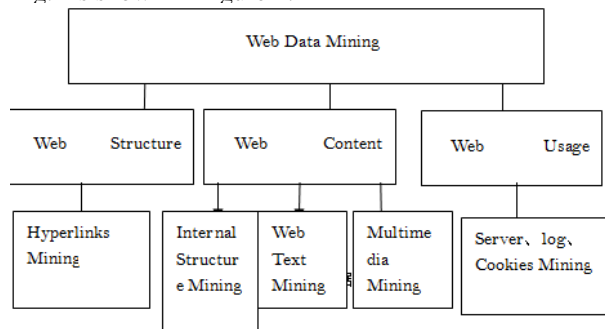


Figure 1. Web Data Mining Structure

1) Web Structure Mining

In Web structure mining, information is derived mainly from Web organizational structure and link relations. Web structure mining includes hyperlinks mining and internal structure mining. And in terms of structure, each Web page is not a completely flat structure, but has its own specific structure. The purpose of Web data mining is to find the page structure and Web structure, and by using the hidden information of these structures, we can find more valuable models or knowledge.

2) Web Content Mining

Web content mining refers to summarize, classify, cluster, make correlation analysis, and predict the trend using Web text, etc. of the large collection of documents on the Web and its is a process of extracting valuable knowledge from the description of Web texts. There are text data and multimedia data like sound, images, graphics, video, etc. on the Web and there is also free text without structure, plate-structure data tagged with HTML and structure data from the data, therefore, Web content mining also includes text mining and multimedia mining.

3) Web Usage Mining

Web usage mining is the model to extract the logs users are interested in from their browsing histories, and analyze these log data so as to get to understand users' behaviors and habits, thus provide personalized recommendation service for them. It mainly includes Server, Cookie and Log Mining. The Web data mining introduced in this paper mainly refers to Web usage mining.

IV. APPLICATION OF DATA MINING TECHNOLOGY IN E-COMMERCE RECOMMENDATION SYSTEM

A. Application of Web Mining Technology in Optimization of Web Design

With the application of Web mining technology in optimization of Web design, many data mining tools have emerged in the various integral mining technologies. And these tools are auxiliary integrated analysis tools in developing various websites, running on the users' database or data warehouse of the web. According to the

functional modules in the Web mining process design, many practical problems in the process of optimizing web design and operation can be solved. Application of this algorithm in optimizing web design is elaborated as follows:

1) Optimization of Site Link Structure

By mining users' Web log files, the optimization of Web site link structure can be considered from two aspects: one is to find the correlation of users' access and add links between closely related Web pages. Another is to find users' desired positions. If the access frequency of desired positions is higher than that of actual positions, then it could be considered to establish a navigation link between the two.

2) Determine Correlation between Requested Pages

There is a certain orderly relation of users' access to Web sites, which reflects a kind of users' access interest, that is to say access interest of a group users have a strong correlation with their access sequence, and this orderly relation can be mined out by the mining algorithm of interest-relation rules.

3) System Improvement

Users' satisfaction degree of web service is also related to system performance, including the server cache, network transmission, load balance and data distribution, etc. When a user stays longer on a page, it is maybe not because he has great interest in this page, but that the server's status as well as display of content transmission on the page may cause user wait for a long time when opening the page. Before users request dynamic pages, dynamic pages have already been produced so as to decrease the server's response delay time

4) Network Security

With the rapid development of Internet, website security has become the focus of the society, and at the same time become the main factor restricting applications based on the Internet. By using data mining technology, web designers no longer rely on qualitative guidance of experts to design a website, but modify and design website structure and appearance according to users' access models and browsing preference so as to make the site more in accordance with the habit of users and add users' interest of visiting websites.

B. Application of Web Data Mining Technology in Customer Classification Management of E-Commerce

In e-commerce system, the e-commerce environment built by analyzing users' data features in order to solve problems of T (delivery), Q (quality), C (cost), S (service) and E (environment), etc. has covered the whole process of enterprises' production and management, and it is an organic whole integrating personnel, technology, management and enterprises' logistics, cash flow and information flow. Like the systems of production management, inventory management and cost control, CRM is also an application system in e-commerce. On the whole, CRM provides comprehensive information about enterprise management based on customers. By investigating the enterprise management, the author thinks that comprehensive information about enterprise management has the following features:

1) The Data are Extensive and of Large Quantity

Comprehensive information about enterprise management includes financial management, staff management, energy management, material reserve management, order management, production schedule management and cost management and the source of this information is multifaceted: some from the database of each MIS and OA system within the enterprise, some from original documents and manual reports and others may come from the Internet or business database of business partners. Therefore, the large amount of data can be imagined.

2) *Various Data Forms*

There are not only traditional files and data in the database, but also Internet/business intranet data, and there may also be multimedia data like sound and images. And data of the same type (such as database data) may be heterogeneous with different storage formats.

3) *Geared to Different Business*

To providing different information according to different business, and usually the emphases of business demands for information are different. Both information directly related to management and a lot of auxiliary information is needed.

4) *Information Integration*

The application of data mining in CRM is actually to extract business data, marketing data and customer data from a enterprise's data warehouse as the model sample, and use various data mining methods to mine data, analyze and model the mining results and constantly optimize the model so as to provide guidance for the enterprise's marketing decision-making and market planning.

C. *Realization Process of Web Mining Technology in Network Marketing*

Network-marketing-related source data collected in the data collection process include the above customer data, product data, intention data and some external data, specifically including the server log data, customers' registration information, transaction data, etc. Herein, customer behaviors can be understood better by combining customer information and server log data; sales of products can be analyzed by combining customer information and transaction information. All these data can change at any time, which makes data collection with extremely vital significance.

D. *Application of Web Data Mining Technology in E-Commerce Network Marketing*

Customer information within the system mainly involves in network marketing include customer information and transaction information, etc. However, it is needed to combine external information like market and competitors to carry out marketing, and only in this way can we formulate feasible marketing strategy. When all these data are gathered together, they have not only a large number, but also various types. And it is difficult to effectively put some external data into the database, which has increased difficulty to the implementation of data mining. Therefore, it is necessary to classify these data before mining them so as to make preparation for data mining. According to their different sources, data can be divided into the following categories:

Customer data: The biggest change network marketing has brought to traditional marketing methods is the transition from the previous mass marketing to the "one to one" marketing of enterprises to individuals, which is also an embodiment of the "customer-centered" business philosophy.

Commodity data: Commodity data are put into the backstage database of e-commerce system by relevant personnel at the beginning of sales. They are relatively stable in network marketing, and generally will not change.

Intention data: In order to communicate more with customers, many e-commerce sites will set a chat room and enterprise BBS, etc.

E. *Application of Data Mining in E-Commerce Marketing*

From the classification of data in network marketing and implementation of data mining in network marketing, we can see that in the application of data mining in e-commerce marketing, a large amount of data are analyzed and from the many seemingly unrelated data, meaningful information to network marketing and hidden in these data as well as the links between them can be found by data mining

1) *Cross-selling*

Cross-selling is a marketing strategy. Although it has been classified as an aspect of customer relationship management in many analyses, it is however essentially a combined sales strategy which customers are interested in and they may buy products. The key to the success of cross-selling is to ensure that the cross-sold products are indeed what customers are interested in, and the implementation of this strategy can help enterprises to maintain a good and lasting relationship with customers. Data mining in cross-selling is to use this data analysis technology to find the most appropriate matching for product sales.

2) *To Establish Personalized Service System*

Nowadays, as competition between websites has been becoming more and more intense, it has become the primary task of network marketing as how to satisfy customers, retain former customers and attract more new customers. This requires the enterprise marketing sites to quickly and accurately find the information customers needed, provide different services for different customers, allow customers to make customized pages according to their own needs, and provide customers with information like products' marketing strategies, etc.

3) *To Establish Marketing Intelligence System*

Marketing intelligence is the information related to the daily development of all kinds of marketing environments. Marketing activities are based on the understanding and analysis of the market, which needs to collect and sort out a lot of marketing information. Marketing information has a strong timeliness and is constantly changing, which requires enterprises' Marketing Departments to continually collect all kinds of information timely so as to get to know the new situation, research into new problems and take the initiative in marketing. In modern marketing environment, it has become enterprises' Marketing Departments' work to collect marketing environment trends, industry trends, competitors' trends, policies and regulations as well as international market information in network marketing by

using Web data mining technology so as to establish marketing intelligence system.

V. CONCLUSION

The current e-commerce recommendation system basically has the functions all businesses need and can recommend to users any information they want to see at any time, however, it is still imperfect in intelligent screening of information. For example, when I search a noun at Baidu or Google, then all sorts of advertisements will appear at the advertising columns of website alliance and browsers, although I do not need them. A mature recommendation system must get to know the user's purpose of searching this word, whether he just wants to look up the dictionary for the explanation and source of this word or want to buy the relevant products. If developers do not understand this, the advertisement recommendation will only make users feel sick and click rate and display will begin to fall.

Functions of e-commerce recommendation system are still in perfection, and all the problems we have encountered will no longer exist in the near future. After eliminating technical hurdles, what we will study is how to get a better user experience and recommend the information users really need more intelligently. There is a long, long way to go and all the researchers need to work together.

REFERENCES

- [1] Shi Jiansheng, Wu Weiguo, etc. Research of Mining Users' Browsing Models in Web Log [J], Journal of Xi'an Jiaotong University, 2011,3-5.
- [2] Han Jiawei, Meng Xiaofeng, Wang Jing, etc. Research of Web Mining [J], Computer Technology Application, 2012,3-8.
- [3] Pang Zhiying. Application of Data Mining Technology in E-Commerce [M], Shaanxi: Xi'an Jiaotong University Press 2011,236-240.
- [4] Lin Ruijuan, Hou Dewen. Research of Web Mining and Its Application in E-Commerce [J], Humanities and Social Science, Chinese Academy of Sciences, 2013(8), 23-27.
- [5] Deng Xiuqin, Jiang Lianhua. Research of E-Commerce Recommendation System [J], Journal of Eastern Liaoning University 2014(4), 38-42.
- [6] Tan Chunhui, Wang Quanli. Combination of Personalized Service and Web Data Mining Technology [J], Shaanxi: Xi'an Jiaotong University Press 2013. 08 82-85.
- [7] Victor•Kenneth Big Data [M], Zhejiang: Zhejiang People's Publishing House 2013.
- [8] Feng Yongping Application of Data Mining Technology in E-Commerce [D], Zhejiang: University of Electronic Science and Technology 2012,03, 34-46.
- [9] Lin Li, Pan Hao. Application of Data Mining Technology in E-Commerce[J], Computer Knowledge and Technology, 2014, (4) :816-818.
- [10] JPerner,Petra.AdvancesinDataMining.MedicalApplications,E-Commerce,Marketing,andTheoreticalAspects[M] , Springer-Verlag,2012;20-80;
- [11] Shen Hongchao, Research of Application of Data Mining Technology in E-Commerce [D], Wuxi: Jiangnan University, 2013;
- [12] Zhu Zhiguo, Kong Liping. Research and Application of Web Mining Technology in E-Commerce [J], Journal of Hunan Post and Telecommunication College, 2013, (1): 32-37;
- [13] Chen An, Chen Ning, Zhou Longxiang, etc. Data Mining Technology and the Application [M], Beijing: Science Press, 2012: 54-55;
- [14] Zhu Xiaoyun. Research of Data Mining as well as Users in Personalized Service. Intelligence Magazine [M], 2014 ; (2) ;
- [15] Bao Jue, Huang Guoxing, Zhang Zhao. Optimization Methods of Network Structure Based on Web Log Mining [J], Computer Engineering, 2013(12) 83-85.
- [16] Yu Bo, Zhu Donghua, Liu Zhuojun, Application of Weighted Association Rule Mining Algorithm in E-Commerce[J] , Computer Engineering and Application, 2014.44 (17): 128-130;
- [17] Dong Ning. Application of Data Mining Technology in CRM[J], Computer Engineering and Design, 2013.28 (6): 1429-1431
- [18] Fuhrmann S, Krumke S O, Wirth H C. Multiple Hotlink Assignment[C] , In: Proceedings of the TWenty-Seventh International Workshop on Graph-Theoretic Concepts in Computer Science,Rostock,Germany,2005
- [19] Wang Keliang, Yang Li, Zha Pugeng. Refinement Marketing Research Based on RBF Neural Network [J], Business Research. 2008.03:64-67
- [20] Xia Minjie, Zhang Jinge. Research of Applying Clustering to Improve Site Structure in Web Log Mining[J] , Journal of Zhongyuan University of Technology, 2005,16 (6) :39-41.