Third International Symposium on Intelligent Information Technology and Security Informatics

Design of Multi-mode E-commerce Recommendation System

Zhao Qiu
College of Information Science & Technology, University of Hainan
Haikou, China.
qiuzhao73@hotmail.com

Mingrui Chen
College of Information Science & Technology, University of Hainan
Haikou, China.
mrchen@hainu.edu.cn

Jun Huang
College of Economics & Management, University of Southwest
Chongqing, China.
john0227@hotmail.com

***Abstract*—At present, most e-commerce recommendation systems act only as just a single tool and provide a single recommendation model. However, due to the complexity of e-commerce system itself, different occasions and customers of different identities require different types of recommendation services. In the paper, we study the architecture of complex e-commerce which can collect multiple types of data, use a variety of recommendation algorithms order to meet the needs of different types of recommendation.**

Keywords-recommendation system; e-commerce; multi-mode.

I. INTRODUCTION

We have been brought into the era of information explosion by the rapid development of Internet Technology. Whereas, the presentation of mass information makes it difficult for users to look for the items they are interested in so that the usage rate of information is reduced; and on the other hand, it also turns some information which are rarely accessed by people into "rubbish" in network and cannot be accessed by common users. Meanwhile, there is also an overload problem in e-commerce. Although a large number of commodities are expanding user's choice space but it brings the difficulties in purchasing the commodities needed for customers as well. In order to make a deal, people have to visit a large number of useless pages. Although search engine can alleviate this problem, essentially however, it works on semantic matching, namely all the results matched with query will be feedback to users according to different matching priorities. No matter one single user or different users enters the same query, they will get the same results. As an important means of information filtering, recommendation system is a very promising method of resolving the problem of information overload currently [1].

At present, a non-formal concept of recommendation system which is widely cited is proposed by Resnick and Varian (1997): 'It takes advantages of e-commerce sites to provide customers with product information and recommendations, helps users to decide what to buy and simulates sales staff to help them to complete the procedure of purchase' [2]. A complete recommendation system is composed of three parts: behavior recording module in collecting user's information, model analysis module in analyzing users' preferences and recommendation algorithm module [3]. Behavior recording module collects user's explicit information (such as: registration information, ratings or comments and keywords in searching goods) and implicit information (such as: user's browsing behavior, shopping cart information and purchase records). Model analysis module carry out analysis based on the collected records of user's behavior and figure out their interest field. And according to the user's interest field, recommendation algorithm, with recommendation strategies, gives out relevant results.

Since the proposal of the concept of recommendation system, more and more researchers have been concerned that it has gradually become an important research content of e-commerce IT technology. E-commerce recommendation system has been advanced a great development both in theory and practice. But with the further development of e-commerce systems, recommendation system is also facing a series of challenges.

E-commerce system contains a great variety of different types of data. For example, the database can gather plenty of data about the user' purchases, web server maintains specific logs of user's visiting to e-commerce web sites, and at the same time we can also keep records of user registration information, user assessment information and user voting information in database. However, most of current e-commerce recommendation systems only used a small portion of the available information to generate recommendations. They act only as just a single tool and provide a single recommendation model. Meanwhile, e-commerce itself is very complicated. A single type of recommendation system does not fit in the entire e-commerce systems and different occasions of e-commerce needs different types of recommendation. So, in this article, we design a recommendation system which adopts multiple models to provide users more precise and well-rounded information via collecting a wide range of user's information and combining with a variety of recommendation algorithms which can compensate for each other, and summarizing integrated recommendations at end. This will not only save the transaction time, stimulate the purchases, and increase customer' satisfaction and loyalty, but also can refine the structure of web site, adjust the products and enhance the delivery efficiency of the advertising business etc.

This work is supported by NSF of Hainan Province in China (No. 807018); Haikou City Key Science and Technology Plan Program(N0. 2009-038); Research Centre of Rural Economics and Management of Southwest University(08jwsk281); 123 Academic Team-building Project Supported by School of Economics and Management of Southwest University; Central College Basic Scientific Research Project(SWU0909508);

978-0-7695-4020-7/10 $26.00 © 2010 IEEE
DOI 10.1109/IITSI.2010.108

II. SYSTEM ARCHITECTURE DESIGN

Mentioned in the previous section, there is no perfect recommendation algorithm and recommendation system, for each of them has its advantages and disadvantages. And multiple model e-commerce recommendation system in this paper is the very answer to how to integrate these recommendation algorithms organically in order to provide uniform services outside. The architecture diagram of this system is shown in Figure 1.

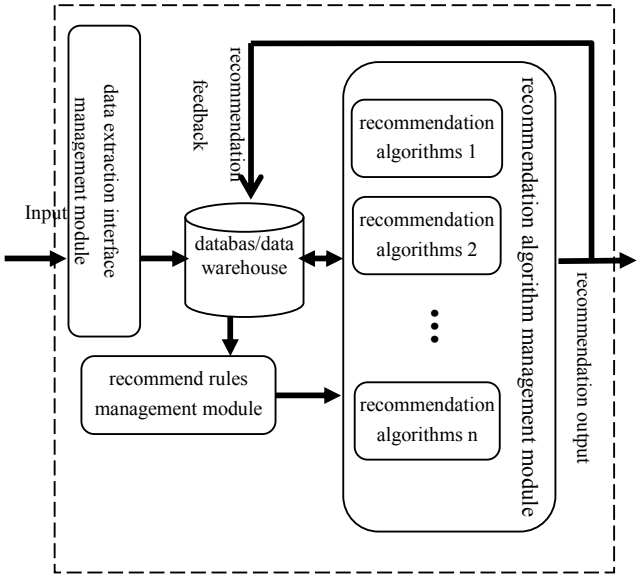

Figure 1. The architecture diagram of Multi-mode E-commerce Recommendation System

The system consists of data extraction interface management module, recommend rules management module, recommendation algorithm management module, database or data warehouse and various recommendation algorithms. Data extraction interface management module manages a variety of data extraction tools and takes charge of them to extract data from an external data source, or may be commercial databases or Web server logs in order to complete necessary cleaning and conversion work and it finally puts the data into database or data warehouse which can be used by recommendation algorithms. Although large amounts of data generated in the real world, most recommendation systems only use a small proportion of it to provide recommendation information. What's more, as the only access to basic data of the whole system, in order to meet the needs of various recommendation algorithms, data extraction interface management module should be able to use information as much as possible, collect various types of data and integrate them into the database effectively.

Recommend rules management module obtains information from database, such as the identity of current user and his browsing behavior. According to the pre-set rules and that information on which it matches, the most appropriate recommendation algorithm for the current user will be figured out. The recommend rules are usually set according to experiences. For example, if the current user hasn't been registered, logged on only as a visitor, the system can only recommend the commodities which resemble what the user is looking for, for the recommendation system has no information about the current user so that cannot figure out the interest field. The setting of recommend rules is crucial for the operation and accuracy of the system and it's also a key factor in integrating various recommendation models. The correct setting requires expertise, and based on the analysis of advantages and disadvantages of different recommendation algorithms and together with the screening and filtering of existing data, it can set out the optimal rules. Otherwise, the error set of recommended rules may lead absurd results.

Recommendation algorithm management module helps to register new recommendation algorithms into system or manage and call the existing algorithms. It receives the specific notifications of recommendation algorithm from recommend rules management module and consequently calls appropriate recommendation algorithms to produce results. And at the same time, it can renew the database according to the recommendation results, namely the feedback of recommendation results. The key technology of recommendation algorithm management module is to solve the problem of incompatibility between the data interface of recommendation algorithm and that of database and the method is usually to modify the interface of recommendation algorithm rather than that of database. There so many types and formats of data needed by different algorithm. If in order to make a new algorithm compatible we modify the interface of database, it will bring about compatibility issues within the original algorithms. When there are too many algorithms, it'll even be a nightmare. And if we only modify the interface of the algorithm newly registered, it won't bring such a problem. No changes will happen to basic database and original algorithms. The feedback of recommendation acts as data's update in database.

III. SYSTEM INPUT

Different e-commerce recommendation systems have different input mechanisms. Based on a variety of keywords, different e-commerce recommendation systems will generate different types of recommendation. Because multiple model recommendation system contains several recommendation algorithms and needs various kind of input information, in order to produce high-quality recommendations, the system designs in this paper should be able to receive multiple types of input information, including:

1) Personal information of user's registration: It may include gender, age, locations, educational level, occupation, income, hobbies and other information which could be helpful in classifying the users in order to predict the user's interest field.

2) Users browsing behavior records: There are some interest fields hidden in users browsing behaviors or habits when visiting e-commerce sites which even themselves are unaware. These behaviors include browsing the goods and the categories the goods belong to, browsing path, the time they stay in the commodity pages and selected goods

shopping cart, etc. Such information extraction is carried out without user's notice and the user simply can't feel their browsing behaviors of visiting web servers have been recording. Because these records are obtained under the user's consciousness, they are more objective and credible and at the same time it can prevent arbitrary or even deliberate misleading in user's registration as well.

3) Keywords in user's search: User's searching behavior is different from the browsing behavior which is much more arbitrary and random. Searching behavior shows that the user is finding his interested goods while the keywords reveal the goods or the properties that appeal to the user.

4) Users purchase records: User's buying behavior, in particular buying certain goods regularly, can better account for the user's preferences of commodities. Recommendation system generally considers user's purchase behavior as an implicit rating.

5) Users rating records: Users can rate a certain commodity with score which is generally a number in a confined scope. The value of such a number account for their preference of commodities. Example, a 5-number method, 1 indicates dislike very much, 2 indicates dislike, 3 indicates noncommittal, 4 indicates like and 5 indicates like very much.

6) Users text evaluation records: Users can use text items to evaluate some certain goods in a describing tone. Such as, commodity performance, whether the price is reasonable and if anything flawed etc. The recommendation system is hardly to deal with text information, so the system is noncommittal at this and only provide this to users when they are visiting.

IV. SYSTEM OUTPUT AND FEEDBACK

Different e-commerce recommendation systems have different outputs as well. Because multi-mode recommendation system adopts various algorithms, it should be able to support various kinds of outputs, including:

1) Personalize goods recommendations: According to predicted interest fields or goods sales, it is likely to recommend the users deliberately with preference list of goods that they may like which is the output form of most e-commerce recommendation systems.

2) Text evaluation output: Text evaluation includes other users' evaluating information about a single commodity or a certain type of commodities, especially the ones of top-ranking or high visiting frequency which represent the evaluation of the mass, and even the professional ones of experts in the field that gets the public identity. Such information is conducive to guide the user's purchase tendency, therefore the system can recommend users text evaluation of the goods they may like according to the interest fields.

3) Rating output: In order to intuitively provide users with an overall assessment of goods they may like, the system should be able to integrate ratings of other users to present to the current user. There are several ways of summarizing scores, such as average method, sub-proportion method etc. For example, when using 5 points method to assess commodities, sub-proportion method figures out the proportion of the various scores so that the current user can see an overall assessment of the products clearly.

V. FEEDBACK OF RECOMMENDATIONS

The output of present e-commerce recommendation systems is to make recommendations to people. In fact, these outputs contain prediction or reasoning of users or certain goods which has a certain degree of credibility and legitimacy. And it can guide data update in turn which we call the feedback of recommendations. This is an advantage of multi-modality recommendation system, namely recommendation algorithms can learn mutually from each other.

A. Commonly used recommending algorithms

Recommending algorithm is the core and critical part of recommendation system which largely decide whether the type and the performance of the system is good of not. However, there is no uniform standard of classifying the types of current recommendation system. Many scholars make different compartmentalization of recommending method from different perspectives [5-7]. Nevertheless from the algorithm point of view, the recommending method can be divided into three categories: knowledge-based recommendation, content-based recommendation and collaborative filtering recommendation.

Knowledge-based recommendation relies on the knowledge in project field or the knowledge in users (such as based on demographic characteristics) to carry out rule-based and case-based reasoning in order to get related recommendations. Content-based recommendation generates recommending results according to the correlation between the information content and user preferences which uses the similarity between resources and user profiles to make recommendations. Algorithms of this kind firstly build up user profile according to user basic personal information and user interest fields figured out by the system, then check the match degree between user profile and would-be predicted project and finally select the project with larger similarity to make recommendations. The basic idea of collaborative filtering is to make recommendations or predictions according to the behaviors of who have the similar points of view. The pattern of collaborative filtering works is like this: based on a user's rating on other projects and the rating records of the entire group to predict the current user's rating on a particular project which hasn't been rating. Collaborative filtering technology is the most successful one in recommendation system. So far, a large number of papers and studies have focused on this and most of the recommendation systems adopt collaborative filtering technology. Although the collaborative filtering technology has been a big success in personalized recommendation systems, with the site's complexity of structure and content, even the increasing number of users, it has also exposed some shortcomings in the same problems, such as:

1) Sparsity problem: in a number of recommendation systems, each user's information is so limited that there will be very sparse data in evaluation matrix and hard to find similar users set which lead to greatly reducing in recommending effects [9].

2) Cold start problem: namely the problems occur when adding new projects or new users. If a new project that no one to evaluate, then this project certainly cannot be recommended, and the recommendation system will lose its roll either especially in the systems which adopt collaborative technology.

B. feedback of recommendations

The multi-mode recommendation system mentioned in this paper is mainly based on collaborative filtering algorithm. Therefore, it also has the shortcomings of collaborative filtering technology talked above. This paper recommends using the output feedback to solve the sparse and cold-start problems, specific practices are as follows:

1) Regardless of what kind of recommending algorithms to be used for getting the results, it shows the prediction that the user thinks highly of the goods in the set of recommendations. We can use this predicting value to modify the rating set the database which is equivalent to save the results of each recommendation.

2) Carry out cluster analysis according to user personal information and integrate users with similar attributes into a group. Assume that the users in the same group have similar preferences for goods. In process of collaborative filtering analyzing, we only select the rating set of the same group in order to analyze. This rating set is the subset of entire the user ratings. No doubt, this will enhance the intensity of rating and overcome the shortcomings of sparse matrix on rating. What's more, because the size of rating set is reduced, the speed of collaborative filtering analysis can be improved. This method can solve the cold start problem as well. Because, new user has no evaluation on goods, so collaborative filtering is not available, but his personal information would remain. We can use personal information for clustering, predicting the rating according other member's rating in the group. In this way, cold start problem will be solved.

3) User personal information or ratings which have been made are usually needed for building user profile by means of content-based recommendation. That use user profile to predict the un-rating project can enhance the intensity of rating. We can use these rating set to carry out collaborative filtering in order to improve its performance.

4) As algorithm of References [11], firstly classify the users - project rating matrix and turn them into user –class rating matrix, secondly based on the measurement of project-class similarity, predict user's rating in order to enhance the intensity of rating, and finally get recommending results with collaborative filtering algorithm.

VI. CONCLUSION

With the rapid development of e-commerce and the explosive growth of the information on Internet, the demands of recommendation system is being more and more urgent. In the past decades years, the research and application of recommendation system has been made considerable progress both in academia and industry. However, there are still a number of same problems that need to be improved constantly, such as sparse problems, cold-start problem etc. In this article, we design a multi-mode e-commerce recommendation system. By means of collecting a wide range of user's information, combined with a variety of recommendation algorithms which can learn from each other and finally comprehensively integrate all information into an overall result to provide users goods of service recommendations more accurately and comprehensively.

ACKNOWLEDGMENT

This work is partially supported by NSF of Hainan Province of China (No. 807018); Supported by Haikou City Key Science and Technology Plan Program(N0. 2009-038); Supported by Research Centre of Rural Economics and Management of Southwest University(08jwsk281); 123 Academic Team-building Project Supported by School of Economics and Management of Southwest University; Supported by Central College Basic Scientific Research Project(SWU0909508);

REFERENCES

[1] XU Hai-Ling, WU xiao. Comparison Study of Internet Recommendation System. Journal of software, Vol.20, No.2, February 2009. pp350-362

[2] Resnick P, Varian HR. Recommender systems. Communications of the ACM, 1997, 40(3):56 一 58

[3] Liu J G, Zhou T, Wang B H. Research Progress of Personalized Recommendation System. Progress in Natural Science. Vol.19, No.1, Jan 2009. pp1-15

[4] Nichols DM. ImPlicit and filtering. In:Proc. of 5th DELOS Workshop on Filtering and Collaborative Filtering. BudaPaest: ERCIM, 1997.31-36.

[5] Balabanovic M, Shoham Y. Fab: Content-Based, collaborative recommendation. Commnunications of the ACM, 1997, 40(3):66-72.

[6] Schafer JB, Konstan J, Riedl. recommender systems in e-commerce. In: Proc. of the lst ACM Conf. on Electronie Commerce. NewYork: ACM Press, 1999.158-166.

[7] Terveen L, Hill W. Beyond recommender systems: Helping people help each other. In: Carroll J, ed. HCI in The New Millennium. NewYork: Addison-Wesley Publishing Co., 2001. 1-21

[8] Resnick P, Varian HR. Recommender systems.Communications of the ACM, 1997,40(3):56-58

[9] ZHAO L, HU N J. Algorithm Design for Personalization Recommendation Systems. Journal of Computer Research and Development, 2002. 39(8): pp986-991.

[10] YU Li, LIU L, LUO Z H. Comparison and Analysis on E-Commence Recommendation Method in China. Systems Engineering-theory & Practice, 2004. (08): pp96-101.

[11] Zhang Haipeng, Li Liebiao, Li Xian and Zhou Yafeng, Collaborative Filtering Recommendation Algorithm Based on Item-Class Rating. Journal Information, Vol.27, 2008.(02): pp218-223