

Using Fuzzy Association Rules to Design E-commerce Personalized Recommendation System

Guofang Kuang*, Yuanchen Li

College of Information Technology, Luoyang Normal University
Henan LuoYang, 471022, China

*Corresponding author, e-mail: kuanguofang2012@163.com

Abstract

In order to improve the efficiency of fuzzy association rule mining, the paper defines the redundant fuzzy association rules, and strong fuzzy association rules redundant nature. As much as possible for more information in the e-commerce environment, and in the right form is a prerequisite for personalized recommendation. Personalized recommendation technology is a core issue of e-commerce automated recommendation system. Higher complexity than ordinary association rules algorithm fuzzy association rules, the low efficiency become a bottleneck in the practical application of fuzzy association rules algorithm. The paper presents using fuzzy association rules to design E-commerce personalized recommendation system. The experimental results show that the new algorithm to improve the efficiency of the implementation.

Keywords: e-commerce, personalized recommendation system, fuzzy association rule

Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

Recommended in the e-commerce system is the use of data mining techniques to analyze the access behavior of visitors to e-commerce sites produce the results of the Recommended product information to help access the customer access to interested e-commerce system already included in the planning and construction of the database system establishment, not particularly high technical content of the e-commerce recommendation system is Added in the original database system on the use of data mining technology on the dynamic customer access returned data to analyze and tune out the customers may be interested in the product directory.

In practical work, the evaluation of a thing, often involving multiple factors or indicators, which requires a number of factors to make a total evaluation of things, that a comprehensive evaluation. Things are often not a simple good and bad, instead of using vague language is divided into different levels of review [1]. For this type offuzzy evaluation object, the fuzzy mathematics method can be used to conduct a comprehensive evaluation. The fuzzy comprehensive evaluation is based on fuzzy mathematics, the application of fuzzy relationship synthetic principle, some boundary.

With the rapid development of the Internet to bring people is in order to the information technology world and the network economy era. E-commerce has been steadily progressing; the expansion and the implementation of all aspects of had a profound impact on the development of business and personal life. Virtual Internet businesses do not like the traditional environment entities construction and investment, businesses with customers, suppliers, etc. to establish a more direct link, e-business models to provide more and better opportunities for the development of businesses; same time, the development of the Internet is also a constant surprise to people, people are not home through a virtual network can be free to get their goods. But in fact, the technology of e-commerce there are a lot of issues need to research and analysis, this hardware device is not upgrading the user has not changed the concept of virtual online shopping, but also the credibility of the business and the lack of a good after-sales service and other reasons.

The biggest advantages of the role of e-commerce recommendation system for e-commerce personalized recommendation system lies in its ability to collect user interest data and users take the initiative to make personalized recommendations based on user interest

preferences, that is, each time you enter a user name and password when the user-commerce sites, recommendation system in accordance with the level recommended by users of the target user preference favorite N products, and the recommended system gives real-time updates, that is, when the system of product and user interest information change, given the recommended sequence will automatically change greatly facilitate the users, but also improve the level of service. In this paper, the definition of redundant fuzzy association rules about nature and propose a new algorithm to remove redundant fuzzy association rules to improve the efficiency of the implementation. The paper presents using fuzzy association rules to design E-commerce personalized recommendation system.

2. The Research of Fuzzy Association Rule Mining

These frequent itemsets treated after fuzzy association rules, the final output of the people interested in the rules. First a fuzzy attributes into fuzzy membership function value of the variable range of numerical quantity mapped to the corresponding fuzzy set theory in the domain. Calculated the weight of all transaction data items (or attributes) corresponding fuzzy sets. fuzzy set with maximum weights for each project (or attributes) using only the following mining process to ensure the number of items (or attributes) with the original consistent algorithm mining focus on the most important fuzzy project (or fuzzy attributes), reduces the time complexity [2]. Fuzzy concept used in the mining process will be able to find fuzzy association rules.

Association rules generated by the frequent fuzzy pattern are similar to the traditional algorithm of longer narrative. Determine the membership function is a key step in the mining fuzzy association rules. For the determination of the membership function, statistical School and non-statistical School held metric pruning strategy of the index space called support-based pruning. This metric pruning strategy relies on the support of a critical nature, that is, a set of support will not exceed its support of the subset, as is shown by Equation (1).

$$\begin{aligned} X(m+1) &:= [x^T(m+1,1), x^T(m+1,2), \dots, x^T(m+1,M)]^T \\ &:= [x^T(mM+1), x^T(mM+2), \dots, x^T(mM+M)]^T \end{aligned} \quad (1)$$

If the candidate items set membership values greater than or equal to the minimum support, L r +1 equivalence class items set on the candidate set. Mining fuzzy similar association rules is found to meet the greater than the user-specified minimum support, minimum confidence and minimum similarity rules. Set equivalence class items set contains $s = (s_1, s_2, \dots, s_q)$, $q \geq 2$ so \min_sup , \min_sup minimum support threshold, items not constitute $C_k + 1$ elements and other items set. Apriori algorithm used to compress the nature of the search space is also used in the algorithm: all non-empty subset of the nature of a frequent itemset must frequent item sets [3].

The association rule mining problem can be formalized is described as follows: $i\}$ is a collection of all the items. All Service D m collection (database), each transaction T is a collection of some of the projects, T is contained in I, each transaction can be identified by the unique identifier TI. A D set T. Association rules is shaped like a $B \Rightarrow I$ Rules $ABBI$, BI , and A is an item set the things T contains A if and only if AB of implications, as is shown by Equation (2).

$$K_E(x) = \begin{cases} \frac{1}{2} c_d^{-1} (d+2)(1 - \|x\|^2), & \text{if } \|x\| < 1 \\ 0, & \text{if } \|x\| \geq 1 \end{cases} \quad (2)$$

A priori algorithm actually there is two premises assuming: 1) the same as the importance of each element of the database; 2) the purpose of the database distribution is uniform, i.e. the frequency is the same or similar. However, in the real world database is often not the case. Database projects uneven distribution, frequency of occurrence is quite different, it will lead to the support can not reasonably be set, if set high, the discovered association rules

may not involve lower frequency items; while low , you will find too much does not make sense even false association rules [4].

Data mining is today's database research, development and application of the most active branches of the association rules and data mining in one of the many other features, the most important and one of the most widely used association rule discovery can help many business decision making, such as classification design, cross-shopping, and fire sale analysis, and so on of the most famous and most widely used association rule algorithm is proposed by R.Agrawal, Apriori algorithm. Association rules algorithm is shifted by the deterministic fuzzy quantity-oriented type, Boolean turning traditional algorithms to solve the "sharp boundary problem, as well as mining, association rules more in line with the thinking habits, as is shown by Figure 1.

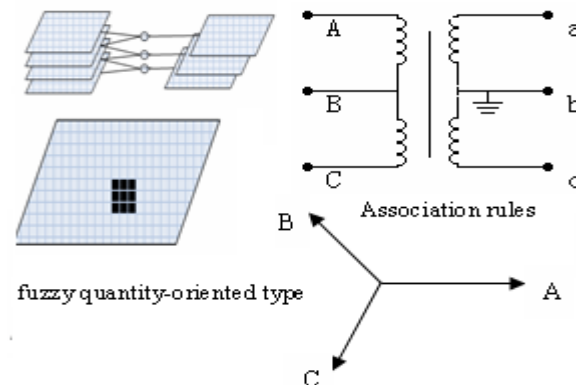


Figure 1. The Fuzzy Association Rules Architecture

Algorithm thinking: basic steps A priori algorithm is similar to first identify all weighted support is not less than the user-specified minimum support frequent itemsets, then use the frequent itemsets to generate all the rules to meet the minimum confidence. Input: transaction database D; minimum support threshold m in sup Output: D in the frequent item sets L method: $W(ij) = S \text{ can}(D) // \text{scan the database D to obtain the probability of project properties, and calculate its rights values the c reat D ataBaseFD; // based on the original database D, created the fuzzy databases FD L 1 = f ind_frequent_1 } \forall \text{ item sets (FD); these algorithms with VC, set support.}$

Determine the membership function, and value converted into the language of fuzzy sets, and statistical items each fuzzy membership value of the property and, the maximum fuzzy attribute for the items; equivalence class of fuzzy attribute take maximum items in the price class of membership as the equivalence class of fuzzy attributes; determine the candidate set membership values to the degree of membership of the candidate in each transaction set items value of the intersection that take the candidate set items the minimum membership, and add them together, the candidate set membership values, as is shown by Equation (3) [5].

$$f(t) = f_L(t) + w_L(t) + \dots + w_{M-2}(t) + w_{M-1}(t) = f_L(t) + \sum_{l=L}^{M-1} w_l \quad (3)$$

Candidate fuzzy rule set redundancy association rules to is strong fuzzy association rules, the dollars need to calculate its implication. Advantage of this nature, before scanning the database, based on the strong fuzzy association rules has been excavated, removed from the candidate rule set fuzzy association rules out redundancy, thereby improving efficiency.

Taking into account the usual method, fuzzy association rule mining may have redundant useless rules exist in the real world items similarity mining fuzzy the similar association rule can be found in more versatility and knowledge of the importance of data exclude a large number of redundant rules. This article gives a similar relationship between goods and the similarity of the concept [6]. On the basis of the above to determine the

membership function of the sample data, based on the concept of similarity, combined with the literature the proposed extract association rules the Apriori algorithm, the paper proposes a new mining fuzzy similar association rules algorithm has versatility and rules applicable to all kinds of data redundancy.

Fuzzy set is different from the concept of traditional set an important aspect of the element x in the fuzzy set A , not absolute \forall belong $\#$ or $\#$ a set \forall does not belong, but there is a fuzzy value $A(x)$ (usually at 0, 1 of room) representative of the elements of the degree of membership of this collection $\forall \# A$ set up in things set D has support s wherein s is in D things contains A, B (i.e. the percentage of both A and B), it is the probability $(A! B)$. the percentage of rules A with B c , it is the conditional probability $P(A | B)$. i.e. things D : $\text{supp ort}(AB) = P(A! B) B) = P(A | B)$ having a degree of confidence c , D contains A 's transaction also includes: $x \exists [0, 1$, called the degree of membership, as is shown by Equation (4).

$$\Phi(m) = \left[\left(\prod_{i=0}^0 A(mM+i) \right)^T, \left(\prod_{i=1}^0 A(mM+i) \right)^T, \dots, \left(\prod_{i=M-1}^0 A(mM+i) \right)^T \right]^T \quad (4)$$

A priori algorithm actually there is two premises assuming: 1) the same as the importance of each element of the database; 2) the purpose of the database distribution is uniform, i.e. the frequency is the same or similar. However, in the real world database is often not the case.

3. The Development of E-commerce Personalized Recommendation System

In the increasingly fierce competitive environment, the e-commerce recommendation system can effectively keep the user, to prevent churn, improve the sales of the e-commerce system. Recommended system in the e-commerce system has good prospects for the development and application has gradually become an important research content of the e-commerce IT technology, the more and more attention of researchers [7]. E-commerce recommendation system has been great development in the theory and practice. However, with the further expansion of e-commerce systems, e-commerce recommendation system is also facing a series of challenges. Recommendation algorithm design, and the recommended system architecture and other key technologies in e-commerce recommendation system for e-commerce recommendation system faces major challenges, the beneficial exploration and research.

Application of e-commerce there is a very broad space for development, from the new development model of the business services a lot of requirements, including the distribution of goods delivery, buy product quality, Ease of purchase, after-sales maintenance, and returned, which can be compared to see that the problem is more prominent recommended information problems of e-commerce sites. With the rapid development of the logistics industry, the improvement of the supply chain, businesses set up shop by the virtual network platform to provide users with a wide variety of goods [8], as is shown by Equation (5).

$$\mu_{ik} = \frac{\exp(-\frac{d_{ik}^2}{2\sigma^2})}{\sum_{j=1}^c \exp(-\frac{d_{jk}^2}{2\sigma^2})} \quad i = 1, 2, \dots, c; k = 1, 2, \dots, n; \sum_{i=1}^c \mu_{ik} = 1 \quad (5)$$

Utility-based recommendation technology and knowledge-based recommended. Do not generalize the user's long-description. And the level of user preference Recommend to a user's favorite N products. Given above is based on the rating of the user needs and the match between the optional set by calculating the goods to the recommendation is updated in real time. That is, when the effectiveness of the Product in the system and user interest owned user to make recommendations. The focus of the problem is how to record material change for every user. Given the recommended sequence will automatically change greatly facilitate the users to build a suitable utility function, is as follows:

$$W_{2^j} f(n) = 2^{-j/2} \int_{-\infty}^{\infty} f(x) \psi(2^{-j} x - k) dx \tag{6}$$

The main purpose of the e-commerce recommendation system is to browse the website's users into buyers. Login customers some not determine purchase intent, only commodity information website. If the site can be easily and accurately to users the goods they need to stimulate the desire to buy, then they will be added to the ranks of the purchaser. According to a survey, buy experience recommended users to post information, will be increased so that other users of the desire to buy. (2) Recommended system can effectively enhance the cross-selling of the site.

The input section is updated and collect user information when the user enters the website query; it is recommended that the system will collect these data [9]. Input can be divided into two parts of individuals and groups of customers, including customers personal input is the target user, the target user is recommended that they should be on some project analysis and evaluation, so that you can show the user's interests, in order to Recommended system recommended effective and convenient. The input mode includes the explicit input implicit input, keyword input and user purchase history. Groups enter the form of groups to evaluate the project, text evaluation, rating scores, as is shown by Figure 2.

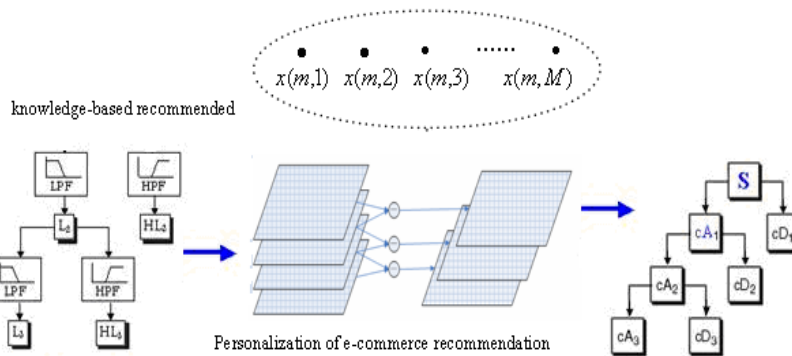


Figure 2. The Design of E-commerce Recommendation System

Study the contents of the e-commerce personalized recommendation system of e-commerce personalized recommendation four areas: First, the source of information recommendation system to solve the problem - Recommended system based on user interest data and information, how the e-commerce environment. For more information as much as possible, and in the right form is a prerequisite for personalized recommendation; second, to achieve customer acceptance and recognition of personalized recommendation, accurate design, high efficiency of the personalized recommendation algorithm is the core.

In addition, let recommendation system is accepted by the majority of users, recommended system must make an objective, comprehensive evaluation, with particular attention to the aspects of the evaluation accuracy rate, personalization, security, user satisfaction; recommendation system application final study of the end result, recommendation system not only provides users with completely personalized shopping environment should support decision-making for the sales and customer relationship management, as is shown by Equation (7).

$$\begin{aligned} P^{(z)}(m+1|m) &= \Phi(m)P^{(z)}(m,M)\Phi^T(m) + \bar{Q}(m) \\ &= \Phi(m)P^{(\alpha)}(m,M)\Phi^T(m) + \bar{Q}(m) \\ &= P^{(\alpha)}(m+1|m) \end{aligned} \tag{7}$$

Commodities browsing are information. Also mentioned is the high level of service enterprises. Overall, e-commerce is restricted to meet the technology to determine the best match. The role of the Service recommendation system mainly in three aspects of the e-commerce website viewers into purchasers of e-commerce sites to improve the cross-selling (rs-sln) cos el g i 3 increase customer loyalties to the e-commerce site degrees based on utility recommended need to project the characteristics of attribute data [10]. This recommendation technology must first get the user function described the effectiveness of the project. Sort all the items and then are use the utility function. Take the top N items as recommended to the target user.

Recommended system face the user (user), the task is to provide users with the project (item) recommended. User is the user recommendation system, that is, customers in e-commerce activities. The project is being recommended object refers to the customers to choose the products and services provided in the e-commerce activities, final recommendation is returned to the user recommended content, as is shown by Figure 3.

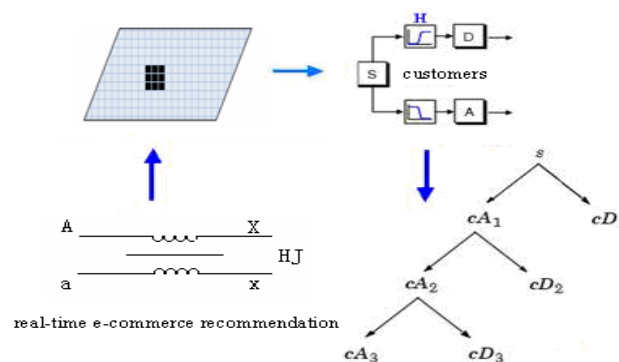


Figure 3. System Personalized Recommendation System Design and Recommend Strategies Structure

Based on the content analysis of customer preferences characteristics, and other customer preferences characterized, it is collaborative filtering on the similarity between the customers, the final recommendations and target customer preference characteristics similar to high or strong correlation of the product as well as its neighbors Testimonials products [11]. Recommended both to avoid the recommended results similar and single, and also makes the product can be recommended unselected. Overall, e-commerce personalized recommendation system has played a very important role in the development of electronic commerce process.

Personalized recommendation system design and recommend strategies should highlight customer needs this theme. The different characteristics of the customer's need are to use different recommendation strategies and recommendation system. According to the master customer the amount of information, the number of customers can be divided into low amount of information customers and the amount of information customers. The low amount of information customers, such customers, including all non-registered customers registered customers of account activity low. Some of these accidental accesses to customers is likely a potential customer, recommended early access products in line with his needs, is likely to be potential customers into valuable customer.

4. The Design of E-commerce Personalized Recommendation System Based on Fuzzy Association Rules

In order to improve the efficiency of fuzzy association rule mining, the definition of a redundant fuzzy association rules, and analyzes the redundant nature of strong fuzzy association rules, the new algorithm to improve the efficiency of mining fuzzy association rules by removing redundant. Addition, support for the use of degree and implication is defined the strong fuzzy association rule mining problem, delete redundant fuzzy association rules and do

not remove redundant fuzzy association rules calculation results with the experimental results were compared and results showed that, when the database project a large number of deleted the redundant fuzzy association rules can improve the efficiency of mining.

Basic database tables must be in the e-commerce website, such as product information, user information, website information, external, should also include: Table of parameters used to initialize the data set, only to Customer Product rating rated commodity Recommended worktable, customer merchandise purchase record table, commodity clustering table, customer clustering table, the recommended list of goods with the popularity of the Internet and e-commerce development, e-commerce system to provide users with more and more choice, its structure become more complex, users often get lost in a large number of goods space, unable to find the goods they need [12]. E-commerce recommendation system directly interact with the user, the simulated shop sales workers to provide goods recommended to help users find the necessary goods to the successful completion of the purchase process, as is shown by Equation (8).

$$d[(x_1, y_1), (x_2, y_2)] = \begin{cases} a_0|x_1 - x_2| + a_1|y_1 - y_2| & \text{if } |x_1 - x_2| > |y_1 - y_2| \\ a_0|y_1 - y_2| + a_1|x_1 - x_2| & \text{else others} \end{cases} \quad (8)$$

Fuzzy association rule mining algorithm is based on the Apriori framework, and therefore its computational complexity by the following factors: the average width of the holding degrees threshold support value, number of items, the number of transactions and affairs [13]. Count on support in the production of L for each transaction, the need to update each item in the transaction database. Assuming average width called for the transaction, the operation time $O(N)$ where N is the total number of affairs. N^2 , set to generate candidate k a need to merge the frequent (k-1) itemsets determine if they are at least k 2. Each merge operation can take up to k-2 equal comparison. Equal to the case, each time to produce a viable candidate k a set; in the worst case, the algorithm must merge the last iteration of each pair found a 1 a set of frequent k.

Data analysis is designed to establish customer behavior model, forming a customer file. According to the recommendation algorithm needs, these patterns can be customer classification or clustering, but also some association rules [14]. These models should be able to answer questions such as high, low consumer has what characteristic, and they are what their preferences and buying habits and behavior problems. Data mining, neural network, model algorithm and other means of information processing in data analysis has been widely used. These methods each have advantages and disadvantages, recommendation system needs both accuracy and real-time performance, a good system is likely to be a variety of methods or technologies. This paper produces a product to recommend list: according to the collection of client information, through the analysis of the data may be obtained from the customers the most interested in the list, recommended to the customer.

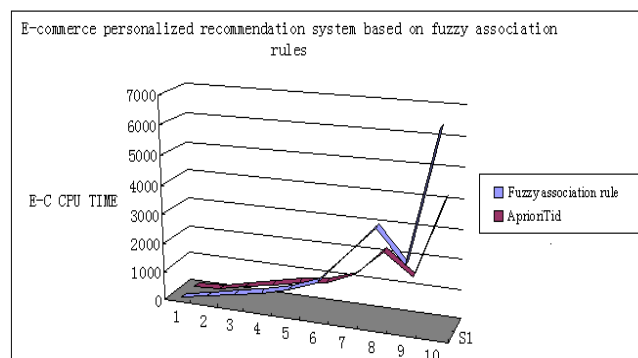


Figure 4. Compare of E-commerce Personalized Recommendation System Based on Fuzzy Association Rules with AprioriTid

Relatively simple since the data in the experiment 2, the data is more complicated in Experiment 1, the methods of this article than in a single, easy to determine the conclusion of the method used herein consistent rules obtained herein obtained rules credibility higher than the proposed method. This shows that the proposed method has validity, can reduce redundant rules. Method, as well as a new algorithm, based on the method and the concept of similarity mining fuzzy similar association rules to avoid the existing algorithms need to determine the membership function of the sample data difficult subjective, but also reduces the redundant rules, has the versatility, concise advantages calculation is simple and the rules applicable to the general transaction data sets fuzzy similar association rules mining database transaction growth, especially C2 scale, especially in the massive database, the algorithm will significantly improve the frequent itemset generation efficiency, as is shown by Figure 4.

The paper presents using fuzzy association rules to design E-commerce personalized recommendation system. Experimental results and analysis on a comparison of the proposed algorithm, the 367 data sets were tested. 231 data sets are: "BodyFat txt (61 numeric items, 362 data);"Computer Activity databases" (42 numeric items, 81 pieces of data); "pyrim arff (15 numeric items, 34 data). algorithm running environment: PIII 866 and 384 MHz, the Windows bait Visual C++ traditional representation of the uncertainty associated rules and application scope of the relationship between the number of the number of fuzzy association is more suitable for the habit of thinking and reasoning, we can notice that the items of 55 few weights is 2.1, if the use the Apriori or AprioriTid but algorithms.

This paper gives the definition of redundant fuzzy association rules, analysis of the relevant properties. And put forward by removing redundant fuzzy association rules to improve execution efficiency of the new algorithm. At the same time, the analysis of experiment results compared the use of support degree and implication degree to solve the problem of fuzzy association rule mining methods.

5. Conclusion

The purpose of this function is to enable consumers shopping in the virtual network can quickly find goods that meet their requirements. In this environment, users need a way to clear and thorough understanding of the analysis service purchasing needs and personal preferences, recommendation system (Recommender Systems) came into being. E-commerce recommendation system can be an effective and convenient recommended, so that users can easily find the merchandise they need, also known as personalized recommendation system. The paper presents the design of E-commerce personalized recommendation system based on fuzzy association rules. The obvious drawback of this method, that is easy to overlook or too much emphasis on the partition near the threshold point partition using fuzzy concept to abstraction, generalization, and makes the final mining rules out natural, easy to expert understanding. Attributes is the fuzzy set theory, fuzzy concept not clear connotation and denotation subordinate role, this division is a hard division.

Acknowledgements

This paper is supported by The Education Department of Henan Province Natural Science Fund Project (No.2010B520018).

References

- [1] Yaling Dou. A Multi-Criteria Fuzzy Decision-Making Method to QoS Multi-Constraints Routing Based on Vague Sets. *IJACT*. 2012; 4(5): 58-67.
- [2] Lijuan Zhou, Hui Wang, Wenbo Wang. Parallel Implementation of Classification Algorithms Based on Cloud Computing Environment. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(5): 1087-1092.
- [3] Liu Hua-ling, Zheng Jian-guo, Sun Cihai. The Association Rule Theory and its Application on the Complete Lattice. *JCIT*. 2012; 7(9): 1-5.
- [4] ZHANG Yao, FENG Yu-qiang. Hybrid Recommendation method IN Sparse Datasets: Combining content analysis and collaborative filtering. *JDCTA*. 2012; 6(10): 52-60.
- [5] He Yueshun, Du Ping. The Research of Landslide Monitoring and Pre-warning Based on Association Rules Mining. *JCIT*. 2011; 6(9): 89-95.

- [6] Jin Huting, Ding Yong. Improved Image Retrieval Performance Based on Fuzzy C-means Clustering Algorithm. *IJACT*. 2012; 4(8): 52-57.
- [7] Yue Hou, Yuemei Mai. Chaotic Prediction for Traffic Flow of Improved BP Neural Network. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(3): 1682-1690.
- [8] Yun-Peng Cai. Construction Personalized Recommendation System Of E-Commerce Based On Variable Precision Rough Set. *IJACT*. 2012; 4(11): 125-131.
- [9] Gao Ren, Tang Long, Wu Juebo. A Novel Recommender System Based on Fuzzy Set and Rough Set Theory. *A/SS*. 2011; 3(4): 100-109.
- [10] Yiqun Li, Cong Shu, Li Xiong. Information Services Platform of International Trade Based on E-commerce. *A/SS*. 2011; 3(1): 78-86.
- [11] Song YQ, Zhu YQ, Sun ZH, Chen G. An algorithm and its updating algorithm based on FP-Tree for mining maximum frequent itemsets. *Journal of Software*. 2003; 14(9): 1586-1592.
- [12] Rashad Yazdanifard, Baome Mahmoud Baruani, Shahriar Mohseni. Review of Electronic Data Interchange in Business to Business E-Commerce in a Competitive Global Market. *RNIS*. 2012; 9: 48~53.
- [13] Guangjun Yang. Mining Association Rules Based on Cultural Clonal Selection Algorithm. *IJACT*. 2012; 4(10): 141-147.
- [14] Theeramunkong, Thanaruk. Applying passage in Web text mining. *Journal of Computers*. 2004; 19(1-2): 149-158.