Fuzzy Logic Based Product Filtering for Web Personalization in E-Commerce

Bin Hua, Kok Wai Wong, Chun Che Fung
Murdoch University
School of Information Technology
South St, Murdoch, Western Australia 6150

***Abstract*- As Internet becomes popular, web personalisation for E-commerce has gained much attention. The most important strategy of web personalization is to provide the customers with appropriate information or services based on the knowledge about the customers' preferences. This paper proposes an alternate approach based on fuzzy logic for product filtering. Experiment results and simulation have shown that this is a viable approach.**

I. INTRODUCTION

Due to the popularity of Internet and the low cost involved in using Internet for marketing and delivery, E-Commerce poses as an important channel in most modern business models. In many cases, when a customer visits a retail or a catalog-based E-Commerce site, it will take the customer some time to browse the products. Due to the excessive amount of online product information available to the customers, and the need to better modeling the behaviors of the customers, web personalization for E-Commerce has become an important research area for the past few years [1,2]. Due to the huge amount of product information, customer may be confused or distracted while browsing or navigating in the site. It has been shown that a customer may not continue his/her engagement with a particular site if the customer needs to spend too much time and effort to find the desired product. Therefore, there is a need to incorporate certain techniques to filter the information and present only those products that fit the customers' needs. Typical techniques are statistical and pattern matching. Recently, advanced intelligent methods have also been in use. This is the main objective of web personalization, which is one of the significant components of Customer Relationship Management (CRM) using web technology.

As stated in Mobasher et al. [3], product filtering, which is one of the important components in web personalization is crucial for such E-Commerce site to be successful in selling their merchandise. Product filtering normally performs the filtering based on the customer profile of that particular customer by selecting information which most probably matches their interests. In order to successfully implement an appropriate product filtering technique, customers' preferences need to be learnt.

Despite some product filtering techniques have already been introduced, it is still a challenging task to generate the most relevant information for the customers. Some of the previous works include neural networks [2, 4], statistical learning [5] and query refinement [6]. Even though most of them have provided promising results in improving customers' online shopping experience, challenges in this field still exist. One of the challenges is to design a system that can handle the uncertainty of customers' shopping behaviors and allow experts to incorporate some human knowledge and experience into the system. Considering the uncertainty and impreciseness of customer's online behavior, fuzzy logic is selected in this paper as an alternative to build the product filtering system. The proposed fuzzy logic based system will analyze the collected customer data and learn customers' shopping behavior. By using fuzzy sets and linguistics fuzzy rules, the proposed system can handle uncertainties involved and at the same time provide human understandable fuzzy rules to allow human expert to update the rules.

II. WEB PERSONALIZATION

The term "personalization" is widely used today in both conventional and electronic commerce. This has formed one of the important areas known as Customer Relationship Management (CRM). Web personalization, which is mostly used for online businesses, is defined as any set of actions that can tailor the web experience to a particular user or customer [1]. Web personalization can improve the E-Commerce experience by helping companies to maximise their customers' perceived value, which may translate to the increase of profit. Ho and Tam [7] investigated the effects of web personalization in terms of the user's decision making stages. Their study showed that users are more willing to look at the personalized content at the early stage of their decision making. Most E-commerce sites which make use of personalization technology when dealing with customers reported a significant increase in revenue after they have implemented it [8].

Web personalization is used to provide personalized information, tailoring user's web experience, updating users' interests, etc. In Wu et al. [9] defined personalization as a system development approach for designing information systems that change configurations based on each user's needs and preferences. Kramer et al. [10] described web personalization as a toolbox of technologies and application features used in the design of an end-user experience. In this paper, we consider web personalization as a set of processes that uses different methods and techniques to modify web environment which will target at the fulfillment of satisfying a personal requirement and enhance user's web experience.

Murthi and Sarkar [11] viewed the personalization process itself as consisting of three main stages: learning, matching and evaluation. In the learning stage, the E-Commerce website collects data from the customers and analyzes collected data to learn about the customer's preference. The website then uses customers' preference to perform product

1-4244-1272-2/07/$25.00 ©2007 IEEE

filtering that best reflects the customers' needs. In the evaluation stage, it needs to develop appropriate metric to assess the effectiveness of a personalization model.

Wong et al. [2] used a flow chart to present the process of web personalization, which is shown in Figure 1. In a typical online shopping scenario with web personalization, a customer first logs into the web site using the unique ID and password. The system will then retrieve the customer's profile from a database and understand the customer's interest. Based on the generated preferences, only those relevant products with higher ranking will be selected. Eventually, the selected products information is assembled together to produce a personalized web page for the customer.

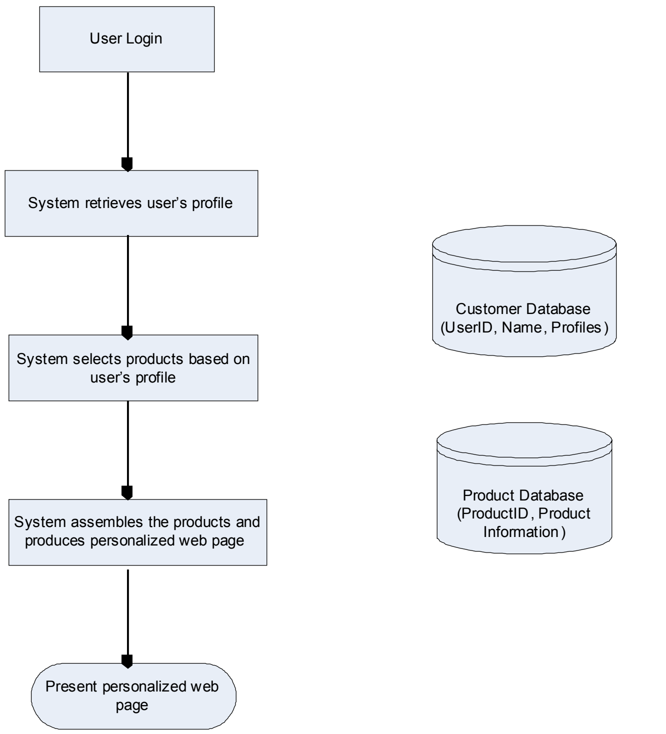

Figure 1. Web Personalization Process.

Data collection is the basic element for web personalization. Without it, web personalization is impossible to be put into practice. Collecting data is the same as observing and recording customer's online behaviors. When a customer interacts with the commerce system, he may browse the product catalogue, view the detailed description for the products, places some products into his shopping basket and finally makes purchase. These customer activities are collected and well stored in database for personalization purpose.

There are several ways to collect customer information. The two most commonly used methods are explicit profiling and implicit profiling. Explicit profiling is done by asking the customer to provide information directly, such as using a web form or survey for each individual customer. Implicit profiling is the collection of data without customer's awareness. To date, a number of approaches have been developed to deal with inexplicit profiling. Wang's customer behavior recording method makes use of a data structure to record customer's online behavior [12]. It has been proved to be an efficient and straight forward approach for collecting customer information. This data structure will be employed in our study to record customer's online activities.

Product filtering is the main focus of this research. It is used to provide recommendations to customers based on the collected customer information. With the recommendations, customer the customer can locate their "preferred" products and hopefully make purchases. This may increase the perceived value of the E-Commerce site. There are four common ways for online product filtering which is shown in Figure 2: rule-based filtering [13], simple filtering [13], content-based filtering [14] and collaborative filtering [14].

The content-based product filtering will be used in this research. It refers to the analysis of the attributes, such as price of a product to give hints on whether the customer will be interested in this product. Usually the hints are based on whether the attributes of the product match the preferences of the customer.

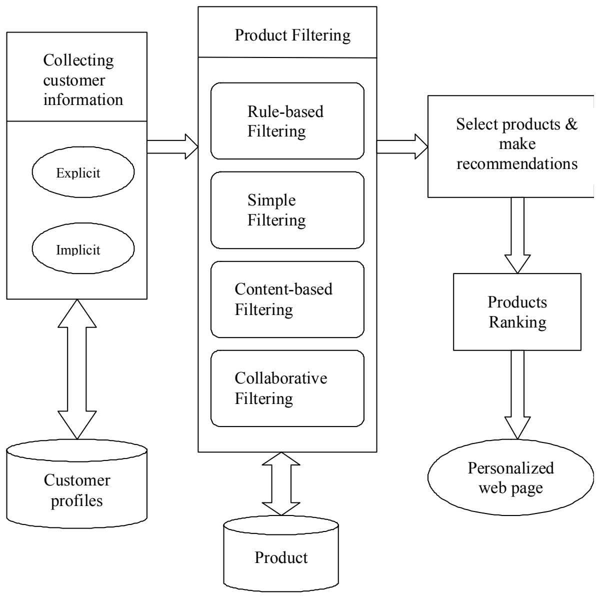

Figure 2. Online Product Filtering.

III. FUZZY LOGIC

Fuzzy logic theory was first introduced by Zadeh in 1965 [15]. The proposed theory allows vague, inexact and poorly defined concepts to be defined by a mathematical process. Fuzzy logic is different from classical two-value sets and logic. Whereas classical logic holds that everything can be expressed in binary terms (0 or 1), fuzzy logic replaces Boolean truth with a continuous range of truth-values in the interval [0, 1]. Fuzzy sets are generalized as a mathematical way to represent and deal with vagueness and uncertain in computing field.

The way to represent fuzzy sets is based on membership functions. The membership function defines how each point in the input space is mapped to a membership value between 0 and 1. The function can be linear, either descending or ascending; and it can be normal, triangular or S-shaped function as well. Generating suitable membership functions for fuzzy sets is one of the most challenging issues in fuzzy systems design. This is mainly due to the reason that it directly affects the accuracy of the established fuzzy system. Normally, membership functions are designed based on expert experience and optimized based on the expected outputs.

Fuzzy rules are the essence of knowledge representation in fuzzy rules based systems. Fuzzy rules are used to translate fuzzy inputs into the actual outputs. Fuzzy rules usually take the form of IF-THEN rule. Fuzzy rules are designed and optimized from the domain expertise, or sometimes can be generated from some well developed tools, such as the Fuzzy Rule Interpolation Matlab Toolbox (FRI Toolbox) [16]. The number of rules involved is related to the complexity of the problem.

Nowadays, fuzzy logic has been widely used to control dynamic systems, such as equipment that must adjust to constantly changing settings and conditions. Some successful industrial examples include: oil recovery, vacuum cleaner, washing machine and fuzzy lift control system. In commercial area, fuzzy systems are also well known as a powerful tool for knowledge support, especially in knowledge discovery.

IV. PROPOSED FUZZY LOGIC BASED SYSTEM

There are two data modules and two processing modules in the proposed system as shown in Figure 3. The two data modules are customer information and product information. The two processing modules include customer preference learning and product filtering, which are the key components in the web personalization model. The preference learning is assisted by the fuzzy logic system which deals with the uncertain information from the customer's online behavior. The proposed system presents an alternative approach for web personalization that integrates fuzzy logic for customers' preferences measurement.

The product filtering component will use a ranking function to match customer preferences with the products' attributes, filters out the unmatched ones and outputs the matched items in a rank order. Eventually, the personalized web page with the most relevant products will be presented to the customer in the next login. By turning the collected data into meaningful information, the E-Commerce site now can provide their customers with a different online experience.

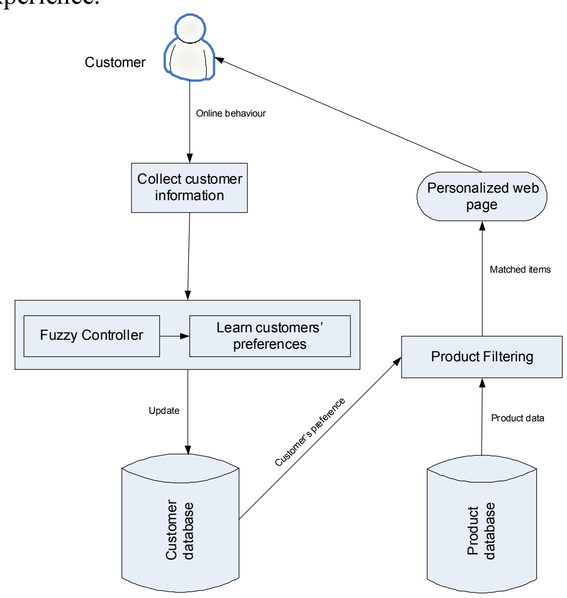

Figure 3. Proposed Fuzzy Logic Based Product Filtering.

The proposed system employs deal with the uncertainty of customers' online behavior. The proposed system will be able to generate the *Preference Significance* value based on the *access time* and the *behavior ranking result*. One of the input variables, access time, is the calculation of how many times the customer actually requested a certain item's URL through the shopping site. For example, if a customer clicked the product item *A* three times before logging out, the access time for item A is assigned as 3. Another input variable, behavior ranking is analyzed and stored using Wang's customers' behavior recording method [12]. Triangular membership function is used to design the fuzzy sets due to its small memory requirement and fast processing. The input and output values are evenly distributed into 5 ranges. The input value is determined based on the process knowledge, and the range of fuzzy's output, *Preference Significance*, is optimized by a 2^k factorial design. All the final membership functions are shown in Figure 4.

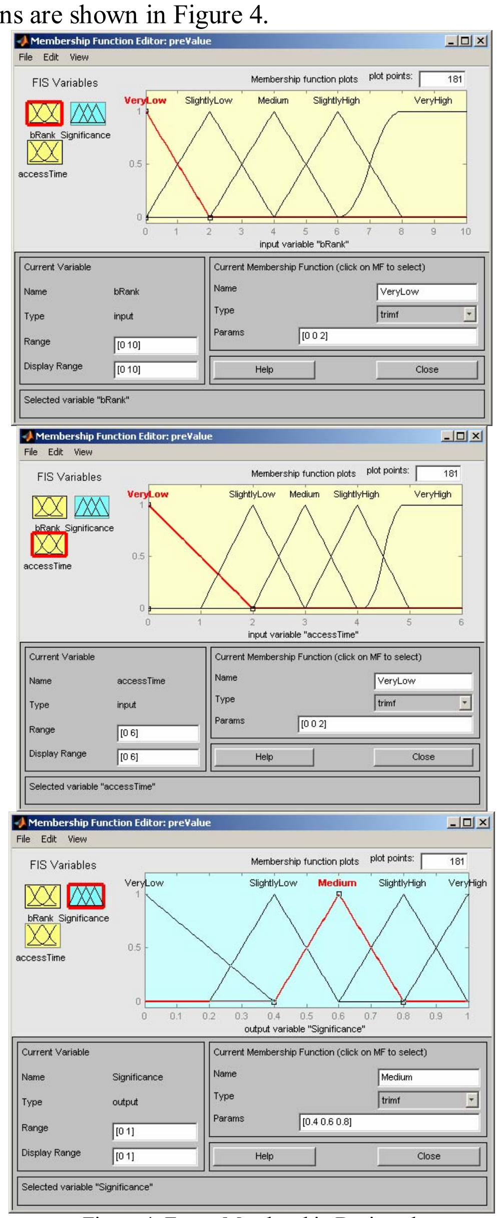

Figure 4. Fuzzy Membership Designed.

Based on Wan's Computational geometry approach [12], the proposed fuzzy rules base has been designed and optimized. Table 1 shows the fuzzy rule base represented by a matrix with *Behavior Ranking* in columns and *Access Time* in rows. The characters in Italic represent the output value, *Preference Significance*. The symbols equivalent as: VL: Very Low; SH: Slightly High; SL: Slightly Low; VH: Very High; M: Medium.

TABLE I
FUZZY RULES BASED

Ranking →	VL	SL	M	SH	VH
Access time ↓					
VL	*VL*	*SL*	*SL*	*M*	*M*
SL	*SL*	*SL*	*M*	*SH*	*SH*
M	*M*	*M*	*SH*	*SH*	*VH*
SH	*M*	*M*	*SH*	*SH*	*VH*
VH	*SH*	*SH*	*SH*	*VH*	*VH*

V. EXPERIMENTS AND SIMULATED DATA

In order to test the effectiveness of the proposed system, it is evaluated based on experiments using simulated data. A simulator is created to simulate the customer interaction with the E-Commerce site. The evaluation was carried out by simulating users' interactions with the system. In this research, in order to show the feasibility of using fuzzy system, only two attributes, price and perceived quality, are taken into account. However, in the real world, the perceived product's quality is subjective and may be related to the price. A product with higher price usually has a better perceived quality than a product with lower price. Therefore, just by using the randomly generated price and quality value is not appropriate. However, for the purpose of this research, we will use the developed product simulation from Wang's methodology [12].

The value of *ndpm* is a metric used widely for evaluating Information Retrieval systems [2]. It is used here to measure the effectiveness of the product filtering algorithm because of the similarity between the product filtering and the information retrieval. In the context of this research, ndpm is used to measure the difference between the customer's choices and the system's choices. Assume the symbol $\succ$ denotes a preference ranking, where a $\succ$ b means a is preferred to b. Let $\succ_c$ denotes customer ranking and $\succ_s$ denotes system ranking. Then the distance of the two rankings for a pair of product is calculated in the following way:

$$dist(a,b)=\begin{cases}1 & if\,(a\succ_s b\, and\, b\succ_c a)\, or\,(b\succ_s a\, and\, a\succ_c b)\\ 0 & otherwise\end{cases} \quad (1)$$

The *ndpm* is the average distance of all pair of products.

$$ndpm=\frac{\sum_{a,b\in T} dist(a,b)}{|T|} \quad (2)$$

where *T* is the set of products and |*T*| is the number of pairs of the products.

From the Figure 5, it can be observed that in all cases, the *ndpm* drops greatly after the customer's first login. In the customer's first login, due to the initial null value of the customer's online behavior, the filtering process is not executed. This also means the system performs effectively after reading the customer's online behavior. With no noise or small noise (noise = 5 and 10), the *ndpm* could reach near to zero after few logins. However, with large noise (noise = 20 and 30), the *ndpm* value is increasing and the amplitude of vibration is also getting larger. In both large-noise cases, *ndpm* can still keep less than 0.2 at most of the time, which means the system performance is still acceptable even in a large noise environment. This shows that fuzzy logic based filtering system can be an alternative used in web personalization.

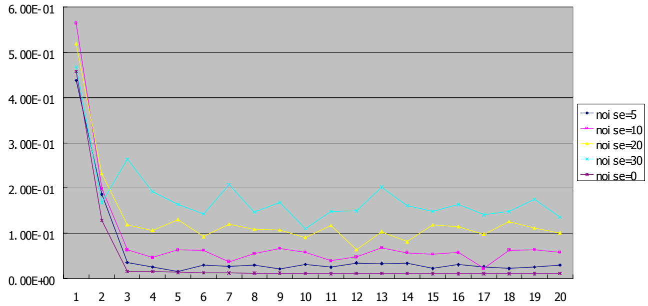

Figure 5. The *ndpm* of the simulated experiment.

REFERENCES

[1] M. Eirinaki and M. Vazirgiannis, "Web Mining for Web Personalization", *ACM Transactions on Internet Technology*, Vol. 3, No. 1, Feb 2003

[2] K.W. Wong, C.C. Fung, X. Xiao, K.P. Wong, "Intelligent Customer Relationship Management on the Web", *Proceedings of the IEEE International Region 10 Conference*, Melbourne, Australia, 5pp. Nov 2005.

[3] B. Mobasher, H. Dai, T. Luo and M. Nakagawa, "Discovery and evaluation of aggregate usage profiles for web personalization", *Data Mining and Knowledge Discovery*, Vol.6, Issue.1, pp. 61-83, 2002.

[4] S. Korper and J. Ellis, *The E-Commerce Book: building the E-Empire, Second Edition*, pp.39-41, Academic Press, 2000.

[5] M.H. Christiansen and S.L. Curtin, "The power of statistical learning: No need for algebraic rules" Available at: http://cnl.psych.cornell.edu/papers/cogsci99.pdf (Accessed 15/09/2006)

[6] K. C. Laudon and C.G. Traver, *E-commerce: Business Technology, Society*, pp.370, Addison Wesley, 2001.

[7] S. Y. Ho and K. Y. Tam, "An Empirical Examination of the Effects of Web Personalization at Different Stages of Decision Making", *International Journal of Human-Computer Interaction*, Vol. 19, No. 1, pp. 95-112, 2005.

[8] B. Beth, "Behind the Numbers: E-business communication gets personal", *Information Week*, August 2000

[9] H. Wu, MD. Gordon, K. DeMaagd and N. Bos, "Link analysis for collaborative knowledge building", *Proceedings of the fourteenth ACM conference on Hypertext and Hypermedis*, Nottingham, UK, pp. 216-217, 2003.

[10] J. Kramer, S. Noronha and J. Vergo, "A User-Catered Design Approach to Personalization", *Communications of the ACM*, pp. 45-48, August 2000.

[11] B. P. S. Murthi, S. Sarkar, ""The role of the management sciences in research on personalization", Management Science, Vol. 49(10), pp.1344-1362, 2003.

[12] Z. Wang and C. K. Siew, "A novel approach to Personalization product filtering in Electronic Commerce", *Australia Journal of Intelligent Information Processing Systems*, Vol. 7, pp. 81-95, Spring-Summer 2001.

[13] S. Murugesan and A. Ramanathan, *Web Personalization: An Overview*, Active Media Technology, pp. 65-76, 2001.

[14] B. M. Kim, Q. Li, C. S. Park, S. G. Kim and J. Y. Kim, "A new approach for combining content-based and collaborative filters", *Journal of Intelligent Information System*, Vol. 27, Issue. 1, pp. 79-85, 2006.

[15] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338-353, June 1965.

[16] K.W. Wong, Z. C. Johanyak, D. Tikk and S. Kovacs, "Fuzzy Rule Interpolation Matlab Toolbox – FRI Toolbox", *Proceedings of IEEE International Conference on Fuzzy Systems 2006*, Vancouver, Canada, pp. 1427-1433, July 2006.