A FEATURE-BASED APPROACH TO MAKING RECOMMENDATIONS

Sung-Shun Weng, Shyue-Ping Chi, Dept. of Information Management, Fu Jen University, 510, Chung-Cheng Road, Hsin-Chuang, Taipei, Taiwan, R.O.C., 886-2-29031111, im1032@mails.fju.edu.tw

ABSTRACT

Most recommendation systems cannot recommend brand new products because new products often lack ratings and comments, even the recommendation systems can find customer groups that have similar interests as target customers. Similarly, products that are less often purchased have fewer records of ratings. Therefore the chances of being recommended are often lower. This research attempts to analyze customers' purchasing behaviors based on product features from transaction records and product feature databases. Customers' preferences toward particular features of products are analyzed and then rules of customer interest profiles are thus drawn in order to recommend customers products that have potential attraction with customers. The advantage of this research is its capability of recommending to customers brand new products or rarely purchased products as long as they fit customer interest profiles, a deduction which traditional market basket analysis and collaborative filtering methods are unable to do.

INTRODUCCTION

The emergence of Internet has brought many business opportunities for enterprises and companies. However, the overload of information has brought to the notice of E-commerce operators the necessity of providing more convenient or more intuitive product information. The information enables consumers to find the products they want, thus the consumer loyalty and desires of purchasing may be increased. Therefore, personalized services of recommendations which provide services and personalization in order to retain customers have prospered.

Understanding of customers can be applied to transform customer information into quality services or products. However, with the great number of customers, how do we identify their interests? The answer to this question is to build personalized Internet services. The purpose of personalization is to adjust promotion and advertisement to fit customer interest [6]. Therefore, it is necessary to understand customer interests and preferences and then provide suitable products or services at an adequate time. A good recommendation system must be relied on in order to recommend products or services of interests. The mechanism of recommendation not only promotes the rate of visiting, but also increases opportunities of sales, and even advertisement revenue, which shall increase a website's profitability.

RESEARCH BACKGROUND

E-commerce websites can predict a customer's future purchasing behavior through the information provided by suppliers, customer demographic data and a customer's past purchasing behaviors. Therefore personalized products can be recommended to customers and achieve effects of transforming people who just browse on the Web into consumers, increasing customer loyalty and enhancing cross selling [4]. Current common approaches for personalized recommendation systems are the content-based approach and collaborative filtering analysis [1] [2] [3] [5]. A content-based system relies mainly on content and relevant profiles as main recommendations to customers. For instance, a customer's past purchase may be traced and relevant products can be recommended to customers. On the other hand, Collaborative filtering analysis groups members of similar characteristics. Their shared interests and hobbies are analyzed and items of the same interest can be recommended to the community members in need. The core of content-based recommendation focuses on the correlation between products prior to

recommending them to users. Yet collaborative filtering methods group users of a community based on the similarity of their profiles. Through community members' hobbies, behaviors, or browsing paths, target customers of a particular community are recommended based on these groups. Each of the two approaches have their own strengths and weaknesses, and the best way is to combine both approaches in content-based orientation and collaborative filtering to promote system accuracy for better results of recommendation.

THE RESEARCH METHODOLOGIES

This research is established on objective and recognizable features of products in the product database. For example, a table may be made of wood, dark brown in color, suitable as a dining table, etc. Such characteristics are matched with basic properties recorded in the database, such as price and brands. Then the product feature database can be built to enable the description of product feature profile, which can be expressed by vectors, as expressed in formula (1):

$$P_{(m)} = \left(f_{11},\ldots,\quad f_{1k},\ldots,\quad f_{i1},\ldots,\quad f_{ij}\right),\ m = 1,\ldots,\mathrm{M} \qquad (1)$$

In formula (1), M stands for total number of products with m being the assigned number of a particular product. Also, i is the number of the product's features while j is the j^{th} feature value in the i^{th} feature. The parameter k means that the feature domain of every product feature is not fixed, yet each feature value is weighted in a binary manner, i.e., if the product has the feature, the weight of feature value is 1, otherwise, the value is 0.

The customer profile module in this research analyzes customers' transaction history and finds whether a customer has any particular interests through the products he has previously purchased. Product features that have potential influence on customers are analyzed as system bases of recommendation. Stages of customer interest profile are: (1) Calculate a customer's level of interest in product features, (2) Adjust customer profile weight, (3) Recommendation, (4) Customer feedback, and (5) Customer clustering.

To calculate a customer's level of interest in product features, the following formula is applied:

$$CTI^{ij}_{(n)} = \frac{\sum f_{ij}}{T_{(m)}} \qquad (2)$$

In formula (2), $CTI^{ij}_{(n)}$ (Customer Interest) is the ratio of features among the products purchased by customers, which is customer n's interest in product feature ij. i is the feature of the product and j stands for the j^{th} feature in product feature i, yet $T_{(m)}$ is the product quantity purchased by a customer.

$$CTRI^{ij}_{(n)} = \frac{CTI^{ij}_{(n)}}{\frac{1}{N}\sum_{n=1}^{N} CTI^{ij}_{(n)}},\ N=1,\ldots,n \qquad (3)$$

Moreover, the average level of interest toward a particular feature based on customers' transaction history, which means target customers' levels of interest become quantified relevant to general customers is calculated by formula (3). $CTRI^{ij}_{(n)}$, meaning customer relative interest, is the target

customers' level of interest toward a particular product feature relevant to general customers, while *N* is the number of customers.

A customer's interests and preferences change with time, thus recent product purchase can better reflect a customer's present interest. Therefore time is a factor taken into consideration in this research, and products recently purchased are weighted more. After then the similarity between customer profile and product profile is calculated by Euclidean distance, the shorter the distances, the higher the similarity. When a customer enters a system, categories of products a customer browses are observed, and products of the highest similarity to customer profiles are then found, of which the top five products would be recommended to the customer by the system. On the part of customer feedback, it mainly traces whether a customer clicks on recommended products. When a customer does not click, it shows that the products are not of his interest. Therefore, the system would automatically adjust the recommendation score. Since customer profile module uses Euclidean distance to calculate similarity prior to generating scores, and the closer distance demonstrates the higher similarity. The higher the recommendation score of a product is, the lower the chance that it would be recommended.

CONCLUSION

The recommendation system of this research is constructed on a web-based system, mainly combining two recommendation systems, individual customer profile module and customer cluster module. Customer profile module mainly aims to discover a customer's potential interests based on product features, and customer cluster module finds peers of the same interests as target customers. The total amount spent by customers on a particular product is used as a clustering input variable in order to eliminate the disadvantages resulted by over-detailed spending amount, including too many or too scattered dimensions and slow speed. The research is evaluated based on movie data, and the result shows that the combined module functions better than just one module. There are three contributions in this research: (1) Explainable recommendation result, (2) Serve as references for developing new or potential products and (3) Bring benefits to enterprises from one-to-one marketing.

REFERENCES

[1] Chen, H.C., & Chen, A.L.P. A music recommendation system based on music data grouping and user interests. *Proceedings of the 2001 ACM CIKM International Conference on Information and Knowledge Management*, 2001, 231-238, Atlanta.

[2] Ghani, R., & Fano, A. Building recommender systems using a knowledge base of product semantics. *Proceedings of Workshop on Recommendation and Personalization in eCommerce*, 2002, Malaga.

[3] Sarwar, B.M., Karypis, G., Konstan, J.A., & Riedl, J. Analysis of recommendation algorithms for e-commerce. *Proceedings of the second ACM Conference on Electronic Commerce 2000*, 2000, 158-167, Minneapolis.

[4] Schafer, J.B., Konstan, J.A., & Riedl, J. Recommender systems in e-commerce. *Proceedings of the First ACM Conference on Electronic Commerce*, 1999, 158-166, Denver.

[5] Lawrence, R.D., Almasi, G..S., Kotlyar, V., Viveros, M.S., & Duri, S, Personalization of supermarket product recommendations. *Data Mining and Knowledge Discovery*, 2001, 5(1/2), 11-32.

[6] Yu, P.S. Data mining and personalization technologies. *Proceedings of the Sixth International Conference on Database Systems for Advanced Applications,* 1999, 6-13, Hsinchu, Taiwan.