Personalized Product Recommendation in E-Commerce

Sung-Shun Weng
Dept. of Information Management
Fu Jen University, Taipei, Taiwan
im1032@mails.fju.edu.t

Mei-Ju Liu
Dept. of Information Management
Fu Jen University, Taipei, Taiwan
beer@im.fju.edu.tw

Abstract

The purpose of this paper is 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 derived in order to recommend customers products that have potential attraction with customers. The approach of this paper has its strength to be able to recommend to customers brand new products or rarely purchased products as long as they fit customer interest profiles. This research also derives customers' interest profiles that can explain recommendation results. The interests on particular features of products can be referenced for product development.

1. Introduction

The popularity of the Internet is a great advantage for the growing trend of customer personalization. New enterprises can manage a particular customer's Web experiences by customer personalization and retain the communication or interaction with the customer. Such understanding of customers can be applied to transform customer information into quality services or products. Most enterprises use product oriented strategies rather than customer oriented ones to sell products. To improve customer satisfaction, feedback rate, loyalty, Web sales, and reputation, one-to-one marketing is seen as the most effective approach for customer relationship management. However, with the great number of customers, how do enterprises identify their interests? The answer to this question is to build personalized Internet services. The purpose of personalization is to adjust strategies of promotion and advertisement to fit customer interests [13].

First, 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 this research is aimed to not only promote visiting rate of web stores, but also increase opportunities of sales, and even advertisement revenue, which shall increase a website's profitability. The purposes of this research are listed as the following:

(1). Provide personalized product recommendation

This research filters useful product information from websites in order to provide or recommend appropriate products to target customers or even potential customers. Thus, the focus of this paper aims to provide personalized recommendations in order to increase customers' revisiting rate or repurchasing rate.

(2). Increase the accuracy of recommendations

Through customers' purchasing histories, the product relevance, such as brand, material, size, color, appearance, price, quality, etc., can be studied to understand customers' preferences toward particular product features. When a customer browses products online, relevant products can be recommended. Meanwhile, this can help companies to understand a consumer's purchasing behaviors.

2. Research background

The purpose of this research is to build a personalized recommendation system based on product features. In the following, we will discuss issues of personalization and recommendation systems.

2.1. Personalization

Personalization is that a website can respond to a user's unique and particular needs. Mobasher et al.

Proceedings of the 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE'04)
0-7695-2073-1/04 $20.00 © 2004 **IEEE**

defined Web personalization as an act of response according to the individual user's interest and hobby on Internet usage [9]. Adomavicius and Tuzhilin proposed 5 stages of personalization while customer responses are measured to monitor other stages [2]. They are: (1) collect customer information, (2) profile customers, (3) compare similarity, (4) deliver and present personalized information, and (5) measure customer responses.

2.2. Recommendation systems

E-commerce websites can predict a customer's future purchasing behaviors through the information collected from a customer's past purchasing behaviors and demographic data of that customer. Therefore personalized products can be recommended to customers and might achieve the effects of transforming people who just browse on the Web into consumers, increasing customer loyalty and enhancing cross selling [12].

Current common approaches for personalized recommendation systems are the content-based approach and collaborative filtering approach [3, 4, 6, 7, 8, 11].

(1). Content-based approach

A content-based system relies mainly on content and relevant profiles as main recommendations to customers. For example, a customer's past purchasing history may be traced and relevant products can be recommended to customers. The advantages include: (a) the system can recommend customers with products which match their specific interests or tastes, while these products may not be of interest to others, (b) when customers accept the recommendations and feedback to the system, collaborative filtering can then support the recommendation system, and (c) the reasons of recommendations may also be given. On the other hand, the disadvantages include: (a) unique or different products may not be delivered to customers since the system recommends products according to customers' past browsing records or buying history, and (b) multimedia information, such as images, pictures, and sounds, cannot be analyzed by the system since contents of multimedia are difficult to define.

(2) Collaborative filtering approach

The collaborative filtering approach groups members of similar characteristics. Their shared interests are analyzed and items of the same interest can be recommended to the group members in need. The advantages include: (a) the system recommends products according to past purchasing history of group members who share the same interests. Therefore product natures do not have to be analyzed, causing no problem even if recommended products are of different properties, and (b) the recommended products could be quite different from customers' past preferences. Therefore, it is possible to derive potential needs and interests of customers [5]. The disadvantages of this approach include: (a) products not yet purchased and rated could not be recommended to customers [4], and (b) the system must perform mass product comparisons prior to finding similar groups to target customers,

The content-based approach recommends products to customers according to the associations among products. On the other hand, collaborative filtering methods cluster customers based on the similarity of their profiles. Through group members' hobbies, behaviors, or browsing paths, recommendations are delivered to target customers of a particular community based on the groups with similar profiles. Each of the two approaches has its own strengths and weaknesses, and the best way is to combine both approaches to promote system accuracy for better results of recommendation.

3. Research methodologies

The main idea of this research is based on product features for recommendations. Through customers' transaction records, their preferences toward specific product features can be learned. Thus the problems generated by traditional market basket analysis or collaborative filtering approaches, such as that new products or rarely purchased products could not be recommended, can be solved. Customer interest profiles can be defined and products with potential attraction can be recommended. Meanwhile, a two-stage clustering approach can be applied to find similar customers who may share the same interests as target customers, therefore potential needs of customers can be derived.

3.1. Research design

This research first analyzes customers' transaction records for preference analysis. Then grouping customers with collaborative filtering concepts is conducted in order to find similar groups of target customers so that potential needs or interests of target customers can be recommended which shall improve recommendations that only target right on a customer's preference.

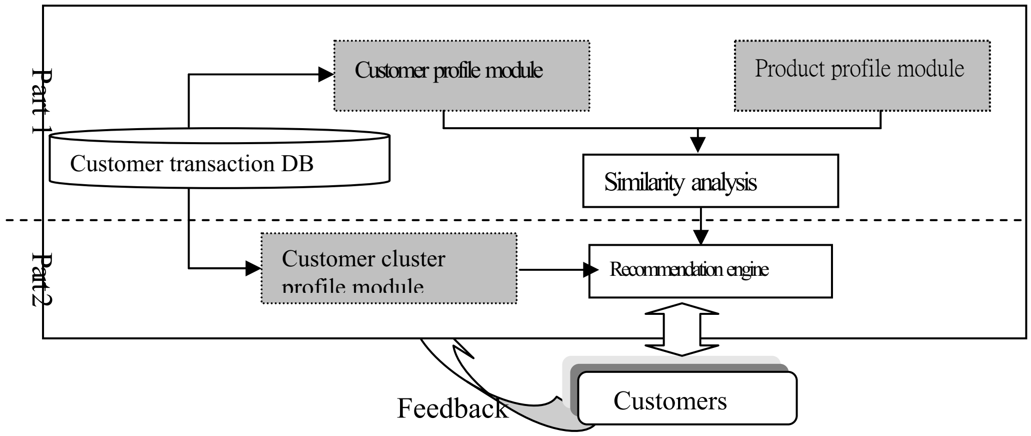

Figure 1. System architecture

The system architecture of this research is shown in Figure 1. For finding customer interests, the procedures of this research are divided into product profile module and customer profile module. Product profile module mainly finds product features in the product database, where features of products form the profile. In customer profile module, product features from customers' past purchasing records are derived for customer profile. Then customer profile and product profile are later analyzed for similarity, and candidate products of recommendation are thus generated. In the recommendation modules, products not yet purchased or rarely purchased can still be recommended to customers with accuracy.

For customer clustering module, customer behaviors are adopted as variables for clustering. Customers with similar purchasing habits and interests are grouped together in order to recommend products with surprises to target customers.

Lastly, the candidate recommendations in the two major parts are presented to target customers through a recommendation engine. With the subsequent customers' feedback responses, the items of recommendations in the system are adjusted.

3.2. Product profile module

This research is established on 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 product database, such as price and brands. Then the product feature database can be built to enable the description of product feature profiles, which can be expressed by vectors, as expressed in formula (1):

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

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 for 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 format, i.e., if the product has the feature, the weight of feature value is 1, otherwise, the value is 0.

3.3. Customer profile module

This module 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 listed in the following:

(1). Calculate level of interest of a customer in product features

The ratio of features among the products purchased by customers, $CTI_{(n)}^{ij}$, is calculated as formula (2). $CTI_{(n)}^{ij}$ means customer interest which is customer n's interest in product feature ij.

$$CTI_{(n)}^{ij} = \sum f_{ij} / T_{(m)} \quad (2)$$

In formula (2), 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. As shown in the following, formula (3) calculates 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. $CTRI_{(n)}^{ij}$, 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.

$$CTRI_{(n)}^{ij} = \frac{CTI_{(n)}^{ij}}{1/n\sum_{n=1}^{N} CTI_{(n)}^{ij}}, N = 1, \ldots, n \qquad (3)$$

(2). Adjust customer profile weight

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.

$$C\tilde{T}I_{(n)}^{ij} = \begin{cases} \dfrac{\sum(f_{ij} \times \alpha)}{T_{(m)}} & where\, \alpha > 1 \text{ and if } purchased\_day < k \\ \dfrac{\sum f_{ij}}{T_{(m)}} & otherwise \end{cases} \qquad (4)$$

In formula (4), α is a self-set weight greater than 1. When a customer purchased the product recently (when $purchased\_day < k$), the product's weight of the product feature is higher and better suits customers' present consumption interest. After time is included for consideration, each customer's $C\tilde{T}I_{(n)}^{ij}$ also changes with it. $C\tilde{T}RI_{(n)}^{ij}$ is calculated the same as formula (3) in principle but follows the adjusted $C\tilde{T}I_{(n)}^{ij}$. Customer profile $C_{(n)}$ is based on adjusted $C\tilde{T}RI_{(n)}^{ij}$, which is expressed in formula (5) as follows:

$$C_{(n)} = \left(C\tilde{T}RI_{(n)}^{11}, \ldots, C\tilde{T}RI_{(n)}^{1j}, \ldots, C\tilde{T}RI_{(n)}^{i1}, \ldots, C\tilde{T}RI_{(n)}^{ij}\right) \qquad (5)$$

(3). Recommendation stage

For recommendation, Euclidean distance is used to calculate the similarity between customer profile and product profile, shown in formula (6), in order to find the top N products with highest customer profile. 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.

$$R_{score}(n) = \begin{cases} \sqrt{\sum \left(C_{(n)} - P_{(m)}\right)^2} & if\ |m| = 0 \\ 0 & otherwise \end{cases} \qquad (6)$$

In formula (6), $C_{(n)}$ is the customer profile and $P_{(m)}$ is the product profile. $|m|$ is the number of transactions of product m. If the product has been purchased by the customer, then the recommendation score it receives is 0, and will not be recommended to the customer.

(4). Customer feedback

This part 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. As shown in formula (7), β must be higher than 1 to move the order of product recommendation backwards. When the product's $\tilde{R}_{score}$ is higher than other candidate products, it is removed from the list of recommendation and is replaced by a product with the next lower $\tilde{R}_{score}$.

$$\tilde{R}_{score}(i) = R_{score}(i) \times \beta \quad , where \quad \beta > 1 \qquad (7)$$

(5). Customer cluster profile module

Customers having purchasing behaviors similar as target customers' ones may have interests that are also potential interests of target customers. Thus this research takes into consideration such potentials and integrates customer cluster profile module with collaborative filtering recommendation approach. In this task, through clustering analysis of data mining, two-stage clustering approach is used to effectively find customers with similar interests, preferences and behaviors as target customers. Each cluster's purchasing records are then analyzed, thus products purchased by other members in the cluster are recommended to target customers who have never purchased the items before, since such items could be of potential demand to the customer. The procedures of customer cluster profile module include: (1) data preprocessing, (2) two-stage clustering which clusters customers mainly by their purchasing behaviors, (3) cluster analysis, (4) adjustment of clustering profile weight, and (5) feedback analysis of customer cluster.

According to Abidi and Ong [1], two-stage clustering approach integrates Self-Organization Map (SOM) and K-means for clustering analysis. The number of clusters generated by SOM can be used as the value of K in K-means clustering analysis. This improves the repeated trial and error tests that K-means analysis has to undergo for the value of K. The first stage of clustering obtains data's cluster numbers and the cluster center of each cluster by clustering data by SOM. The second stage uses the number of clusters gained by SOM as the value of K in K-means. The members of clusters are adjusted in this stage to solve the problems of vague partition caused by SOM, thus the intra-group similarity becomes higher and the inter-group relationship becomes less similar.

The cost of products purchased by each group of customers is analyzed for the level of interests in each product category. Formula (8) calculates the ratio of purchased products to all products by the particular group of customers. This represents customers' interests in the particular product relative to other products. As shown in formula (8), $Amount\ (P_{(t)})$ represents the total amount spent on purchases of product category *t*. $\sum_{n \in cluster} total\_Amount\ (P_{(n)})$ is the total transaction amount spent on all products by the group, and CLI_t stands for money ratio spent on product category *t* out of the total purchasing amount. For example, if CLI_t is 0.8, then 80% of the group's total purchase is on product category *t*.

$$CLI_t = \frac{Amount\ (P_{(t)})}{\sum_{n \in cluster} total\_Amount\ (P_{(n)})}, \text{ where } t \in n \qquad (8)$$

Formula (9) calculates the group customers' interests toward a particular product category compared to average interest of all groups.

$$CRI_t = \frac{CLI_t}{(1/N)\sum CLI_t} \qquad (9)$$

In above formula, $(1/N)\sum CLI_t$ is the average transaction amount of all groups of customers spent on product *t*. N is the number of groups, and CRI_t (Cluster Relative Interest) stands for the particular group of customers' average spending on product *t* relative to the all groups' average spending on product *t*.

In formula (10), $|P_{tu}|$ is the number of a particular product's transaction. Customers' number of transactions of a particular product would affect the level of interests, thus this research includes "number of transactions" as a factor for consideration. The number of transactions is used to adjust the weight of the recommendation score, therefore the sum of money spent is still our main consideration, and getting a log from transaction number would prevent it from having too much influence. $R^n_{score}(tu)$ is the recommendation score of product *tu* among group *n*. However, if a customer has already purchased the product, then such product would not be recommended again.

$$R^n_{score}(tu) = CRI_t \times (1 + \log|P_{tu}|) \qquad (10)$$

For the feedback analysis of customer cluster, when a customer clicks on a product recommended by a customer cluster profile, the system will automatically adjust the recommendation score of that product toward that group. The more people click on it, the more people recommend it. Thus the product's recommendation score gets higher, which would have a higher probability of recommendation, as demonstrated in formula (11):

$$\widetilde{R}_{score}(i) = R_{score}(i) \times W \quad , where\ W > 1 \qquad (11)$$

Above, *i* is the product's number, *W* is the value of weight of product's recommendation score, which must be greater than 1.

4. Experiment results

This research considers the popularity of movie watching in real life, and consequently it would have a higher feasibility in obtaining the data for experiments. In the mean time, the online experiments would likely receive more sample data. Therefore, the experiments of this research are designed based on movie rentals. An online video/audio rental store and its recommendation system are constructed to evaluate the results of this research and the follow-up effects.

The data collection of this research is based on movie information of 586 movies from year 1997 to year 2003, as seen in Table 1. Table 2 is the six major movie features classified in this research.

Table 1. Classification of movies

Serial Number	Category	Quantity
1001~1032	Romance	31
2001~2042	Animation	41
3001~3075	Comedy	74
4001~4172	Science Fiction	171
5001~5139	Drama	138
6001~6013	War	12
7001~7120	Mystery/Horror	119

Table 2. Movie features

Features		Feature values	Remarks
Casts	Actors	ar1~ar88	Total of 88 Actors
	Actresses	as1~as57	Total of 57 Actresses
Director		dr1~dr52	Total of 52 directors
Price		Price	NT$40 ~ NT$100
Release		sub1~sub3	sub1:Disney sub2:Dream Works
Area		Ln1~Ln3	Ln1:Asia Ln2:America Ln3:Europe
Year of release		Year	1997 ~ 2003

For collecting customer data, this research gathered 140 movie lovers who have experiences of video rental. Prior to the experiment, a simple survey was conducted to acquire each person's 5 rented movie titles and a total of 1065 data were obtained. The transaction records later became the recommendation source of customer profile module and customer cluster profile module.

4.1. Evaluation measures of experiment

Precision and recall are used in this research as measures to evaluate the effects of the recommendation system; F1 is also used to represent the effects of combining precision and recall.

$$precision = \frac{No.\ of\ \ purchased\ \ items \cap recommende\ d\ \ items}{No.\ of\ \ recommend\ \ items} \tag{12}$$

$$recall = \frac{No.\ of\ \ purchased\ \ items \cap recommende\ d\ \ items}{No.\ of\ \ purchased\ \ items} \tag{13}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{14}$$

4.2. Experiment procedures

The purpose of the experiment is to test the effects of the recommendation system in this research. First it compares the results between feature recommendation and non-feature recommendation. Then it tests whether the effects of combining customer profile and customer cluster profile is higher than any other types of recommendation module. Meanwhile, it also compares the system with non-personalized recommendation.

There are 4 procedures in the experiment. Procedure 1 tests whether the recommendation based on product features works better than the mechanism that does not use product features. Thus, during procedure 1, this experiment randomly chooses 30 testing data out of 140 data to be the testing objects of procedure 1 and procedure 2. The movies seen by testing objects and movies not yet seen but interested by testing objects are used as testing data. Each testing object provides more than 30 of such movies. A total of 928 data are obtained. Procedure 1 tests the precision of the recommendation system based on these data. The experiment group tests the precision of the recommendation system in procedure 1. The experiment group is the recommendation model that is based on product features listed in Table 2. Meanwhile, the comparison group uses a non-feature recommendation, which is only based on movie categories (as the 7 categories listed in Table 1). Both of the experiment group and comparison group recommend top 10 movies to customers and precision is used as the measure index of the procedure.

Procedure 2 aims to test the feedback weight β in formula (7). When a customer does not click on the recommended movie, the movie's recommendation score will be multiplied by this weight to serve as the system feedback. When the customer enters the system again, the movie not clicked before will be ranked lower. Since the feedback weight setup is different for different business, it would also affect the precision and learning effect of recommendation systems. Therefore procedure 2 will test and find the better feedback weight. This part of the experiment mainly focuses on the repeated testing of different weight β, therefore the 928 data from the 30 testing objects in the previous procedure will be used again to simulate their behaviors through the feedback weight on β=1.1, β=1.2, β=1.3, β=1.4 and β=1.5. Each testing object and each β is tested 3 times to observe the system's learning ability at different feedback weight. Calculation of recall is expressed by treating the data provided by the testing objects as the denominator to represent the transaction records. The numerator is the number of movies that are recommended by both the testing objects and the recommendation system. For instance, if customer 1001 has 30 movies on transaction records, and 20 of the movies are the same as recommended by the system, then the recall is 20/30=0.6667.

Procedure 3 of the experiment tests whether modules that combine customer profiles and customer cluster profiles work better than the one with sole profile. The experiment group is the module that combines customer profile and customer cluster profile while 2 comparison groups are formed. Comparison group 1 only uses customer profile, and comparison group 2 only uses customer cluster profile, where 140 testing objects are divided into 3 groups. The test is conducted online, user behaviors are observed from

05/11/2003 to 05/18/2003. The system will determine for users which recommendation module to enter according to user account. In addition, since user interests may change with time, the user's group may be changed accordingly. To make the clustering results closer to the most recent behavioral pattern, while considering online clustering analysis may lead to customer's impatience due to the long wait, the system batches customer clustering to achieve the optimal customer clustering results.

For experiment group, the system recommends the top 5 movies each for customer profile module and customer cluster profile module, total of 10 movies. To have the same number of recommended movies, each comparison group recommends top 10 movies. Precision and recall are treated as evaluation indices in this experiment. Precision is calculated by dividing the number of movies that customers click in the recommendation area by the number of recommended movies. For instance, in an experiment group, if users click on 3 movies out of the 5 recommended movies, the precision is 0.6. If a customer clicks 4 movies, then the precision is 0.8. The average precision would be (0.6+0.8)/2=0.7. Recall is calculated by recommended movies purchased in customers' transaction history. If customer 1001's transaction has 10 movies, and 8 of them are recommended by the system, then recall is 0.8. Finally, F1 is used as an index considering the precision and recall, which is $2\times0.7\times0.8/0.7+0.8 = 0.7467$.

4.3. Experiment results

This experiment is based on 30 users' behaviors. Feature-based recommendations for movies achieve precision of 0.6104 while non feature-based recommendations only achieve 0.3067; about 30% lower than the experiment group. Since the comparison group only considers the classes of potential interests and neglects users' interests in other features (such as casts and directors), the recommended movies are restricted within certain types, and the precision is unable to increase. On the other hand, this research considers other features of movies and adopts them as bases of recommendation. The multiple considerations match user interests better.

According to the result, as β increases, the recall also gets higher (see Figure 2). As shown in Figure 2, when β gets higher, the recall and system precision does not get significantly higher. The higher value of β, on the contrary, would filter out too many movies. If a customer has a significant change in his transaction behavior, the system would omit some recommended movies due to a change of the customer's interest.

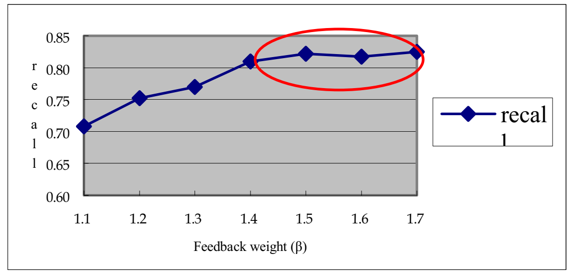

Figure 2. Recall of each weight (β)

When the value of β is small, the system feedback does not perform very well. This is possibly due to the fact that when movies are used as experiment data, a user's non-clicking of a particular movie implies his lack of interest in that movie. This would then result in a slim chance for that user to click the same movie when the system recommends the movie again. However, a user's interests would not change greatly within a short period, thus the system recall will increase with β.

Procedure 2 tests better β of customer feedback. Online experiments are conducted in procedure 3 with $\beta = 1.5$ in a hope to effectively filter out users' non-interested movies without neglecting users' potential interests by filtering out too many movies. In the experiment group, each of the 47 testing objects has an average of 2.6 visits to the website. 32 persons have visited the website 3 times or more, which is 68.1% of the total testing objects. Six people have visited only once, which is 12.8% of total. The remaining 9 testing objects have visited the website twice. In comparison group 1, the customer profile module, 47 testing objects' average visit is 2.4, and the 29 testing objects have visited 3 times or more, which is 61.7% of total. 10 people have only visited once, which is 21.3% in total. The remaining people have visited twice. In comparison group 2, customer cluster profile module: 35 people have visited the website 3 times or more, and 4 people have visited only once, yet the remaining 7 persons visited twice.

As Figure 3 shows, when the number of visits increases, the system precision also increases. It demonstrates that the system has the mechanism to learn customer behaviors. Such a learning mechanism is more obvious in the results on experiment group, since it combines two recommendation modules and allows the system to learn user interests and behaviors more effectively. When a testing object visits the website for the second time, experiment group's precision is significantly higher than those of comparison groups. When a user visits the website for the third time, the precision becomes as high as

0.7857, which is 31.9% higher than that of comparison group 1 and 30% higher than that of comparison group 2.

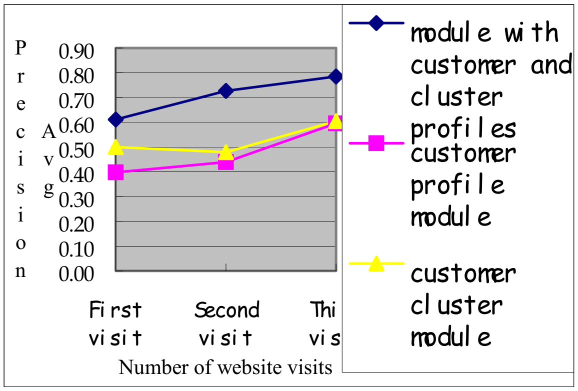

Figure 3. Average precision comparison

5. Conclusions

Many companies have constructed websites and started Internet businesses. To save customers' the effort to search for a product, as well as increase customers' loyalty, the recommendation system is able to serve as a mechanism to retain customers. The recommendation system of this research is constructed on a web-based system, mainly combining two recommendation modules, 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 an input variable of clustering 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 based on movie data, and the result shows that the combined module works better than just one module. In addition, the result also demonstrates that two-stage clustering can really cluster similar customers effectively, and thus bring recommendation items closer to target customers' interests.

In general, the concept of an RFM model is applied in this research. The transaction recency, frequency and monetary value are all considered, which allows the system to rapidly learn customers' behavior changes. The applied membership mechanism enables the system to minimize deviations in recommendations. There are three contributions in this research: (1) explainable recommendation results, (2) serve as references for developing new or potential products, and (3) bring benefits to enterprises from one-to-one marketing.

6. References

[1] S.S.R. Abidi, and J. Ong, "Automated data clustering based on a synergy between self-organizing neural networks and k-means clustering techniques", *Proceedings of IEEE TENCON*, Kuala Lumpur, 2000, pp. 568-573.

[2] G. Adomavicius, and A. Tuzhilin, "Using data mining methods to build customer profiles", *IEEE Computer*, 34(2), 2001, pp. 74-82.

[3] H.C. Chen, and A.L.P. Chen, "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*, Atlanta, 2001, pp. 231-238.

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

[5] J.L. Herlocker, J.A. Konstan, and J. Riedl, "Explaining collaborative filtering recommendations", *Proceedings on the ACM 2000 Conference on Computer Supported Cooperative Work*, Philadelphia, 2000, pp. 241-250.

[6] B. Krulwich, and C. Burkey, "The InfoFinder agent: learning user interests through heuristic phrase extraction", *IEEE Expert*, 12(5), 1997, pp. 22-27.

[7] K. Lang, "NewsWeeder: Learning to filter netnews", *Proceedings of the Twelfth International Conference on Machine Learning*, Tahoe City, 1995, pp. 331-339.

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

[9] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa, "Discovery and evaluation of aggregate usage profiles for web personalization", *Data Mining and Knowledge Discovery*, 6(1), 2002, pp. 61-82.

[10] D. Peppers, M. Rogers, and B. Dorf, "Is your company ready for one-to-one marketing?", *Harvard Business Review*, 77(1), 1999, pp. 151-156.

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

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

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