

## The Measurement of Online Customers' Product Loyalty Status and Its Application

Horng-Jinh Chang<sup>1</sup>, Lun-Ping Hung<sup>2</sup> and Chia-Ling Ho<sup>3,4\*</sup>

<sup>1</sup>Department of Business Administration, Asia University,  
Taichung, Taiwan 413, R.O.C.

<sup>2</sup>Department of Information Management, National Taipei College of Nursing,  
Taipei, Taiwan 112, R.O.C.

<sup>3</sup>Graduate Institute of Management Sciences, Tamkang University,  
Tamsui, Taiwan 251, R.O.C.

<sup>4</sup>Department of International Trade, Technology and Science Institute of Northern Taiwan,  
Taipei, Taiwan 112, R.O.C.

### Abstract

Comparing with the traditional store, the online store can keep the track of customers' purchasing records and personal information. By analyzing these customers' records, online store can have a better understanding of their customers' profile and purchasing behavior. In this paper, we define a standard product loyalty status, or *SPLS*, using customers' purchasing records to evaluate each customer's loyalty to a certain product. *SPLS* is incorporated with loyal customers' personal backgrounds as the input of cluster analysis that divides loyal customers into different groups. Loyal customers in the same groups have similar purchasing behavior and personal backgrounds. Similarity analysis measures the similarity of backgrounds between a non-loyal customer and groups of loyal customers in order to find this customer's belonged group. Then, an expected *SPLS* value is assigned to this non-loyal customer to estimate his/her probability of purchasing a certain product. Customers who have expected *SPLS* value larger than a threshold are regarded as potential customers. Marketing specialists should recommend the product to potential customers. An experimental result shows that there are more than 50 percent of potential customers who actually purchase the product and become a "real" customer. A prototype of our proposed model is used by a fast growing online retailer in Taiwan and is still in the experimental period.

**Key Words:** *SPLS* (Standard Product Loyalty Status), Loyal Customer, Potential Customer

### 1. Introduction

This paper focuses on locating potential customers of a certain product. Since an online marketplace may offer diverse product lines to online shoppers, and it would take too long to analyze entire product lines and all customers, we concentrate only on leading or fast growing products and potential customers. Furthermore, online shoppers are intangible because of their concern with privacy: infor-

mation reported online may be incomplete or inaccurate. Thus, we use only a limited set of personal information such as gender, age, and monthly income as the input to our analysis. This paper presents a way to detect potential customers who are likely to purchase the leading or profitable product and recommend them the product.

The rest of this paper is organized as follows. In section 2, we briefly describe three approaches and the data upon which we based our model. The process of locating potential customers is described in section 3. Section 4 contains our conclusions.

---

\*Corresponding author. E-mail: clho@mail.tsint.edu.tw

## 2. Background

Cluster analysis and similarity analysis are both widely adopted by marketing analysts to build up customer profile. In our proposed method, we use cluster analysis to establish loyal customers' profile and similarity analysis to estimate non-loyal customers' expected product loyalty status. They are described below.

### 2.1 Cluster Analysis

Cluster analysis is widely used to establish object profiles on the basis of objects' variables. Objects can be customers, web documents, web users, or facilities. Hruschka [1], Ozer [2], Weber [3], Espinoza et al. [4], Smirnov et al. [5], Tsai and Chiu [6], Rosaci [7], Xu et al. [8] all used cluster techniques to segment customers and markets. Weng and Liu [9] used a two-stage cluster technique to find customers with interests similar to those of target customers. Buccafurri et al. [10] proposed an agent-based hierarchical cluster technique that operated on user profiles and e-commerce sites.

A  $K$ -means cluster algorithm is widely used by many researchers. It chooses  $K$  initial cluster centers randomly from  $N$  observations and assigns the remaining  $N-K$  observations to the cluster that is nearest based on Euclidean distance. Then, the center of each cluster is recalculated from the observations assigned to each cluster. In subsequent steps, observations are reassigned to the nearest cluster and the center of each cluster is recalculated; this sorting continues until no observation is reassigned to a new cluster. Examples of  $K$ -means clustering are found in Kuo et al. [11], and Shin and Sohn [12]. Ray et al. [13] used  $K$ -means clustering to sort customers. Kim and Ahn [14] used GA-based  $K$ -means to determine the segments of an online shopping market. In this study, we cluster loyal customers who have similar personal backgrounds and purchasing behavior.

### 2.2 Similarity Analysis

The similarity measure is simply that: the measure of similarity between objects. Objects can be documents, people, products, even human DNA. Similarity measure is used in a variety of fields, including psychology, image processing, biotechnology, marketing, e-commerce, and others. Methods of measuring similarity include Euclidean distance, city block (Manhattan) distance, Cheby-

shev distance, Minkowski distance, Canberra distance, Bray Curtis distance, Angular separation, and correlation coefficient. In this paper, to measure the similarity between potential customers who have never before purchased the leading product and loyal customers, we use squared Euclidean distance. Euclidean distance is the most commonly-used method to calculate distance between two points; it takes the square root of the square of the difference between the coordinates of a pair of objects. If a potential customer falls in the range of a group of loyal customers, his or her purchasing behavior is very likely to be similar to the purchasing behavior of that group of loyal customers. Based on Euclidean distance, we develop a process of finding potential customers who are similar to loyal customers. In the field of Internet technology, Niu and Shiu [15], Yi et al. [16], Zeng et al. [17], and Velasquez et al. [18] applied similarity analysis to various problems.

## 3. The Process of Locating Potential Customers

The goal of this paper is to use cluster analysis, similarity analysis, to locate potential customers likely to purchase a certain product at an online store. A potential customer is defined as a customer who has never before purchased the leading product but is very likely to purchase it by analyzing his/her background information and purchasing records. There are three steps in the process of finding potential customers. First, carefully choosing one product that contributes most to a company's sales or shows a dramatic growth in sales is the first and most important step. Second, cluster analysis is performed on customers' personal information and purchasing behavior. All customers in the same cluster of loyal customers will have similar personal backgrounds and purchasing behavior. Finally, we look at the backgrounds of non-loyal customers who have never purchased the leading product, to see whether they are similar to the backgrounds found in any group of loyal customers. An expected  $SPLS$  value is assigned to each non-loyal customer to estimate his/her probability of purchasing a certain product. Non-loyal customers who have expected  $SPLS$  value larger than a threshold are regarded as potential customers.

### 3.1 Product Loyalty Status ( $PLS$ )

To measure the degree of loyalty shown by cus-

tomers who have purchased a certain product before, we define a product loyalty status, or *PLS*, to identify each customer's loyalty to this product. According to Oliver's definition of loyalty (Oliver [19]), loyalty is a deeply-held commitment to again buy a preferred product or again patronize a preferred service in the future, even in spite of situational influences and marketing efforts that could otherwise bring someone to switch. Based on Oliver's definition of loyalty, we take *PLS* to be the degree of a customer's commitment to again in the future buy a preferred product. *PLS* is affected by four factors: total quantity of purchasing (*TQ*), frequency of purchasing (*pp*), quantity purchased in each period (*q*), and the time frame of purchasing (*n*). When any three of the factors remain unchanged, the relationship between *PLS* and the fourth factor should satisfy these requirements:

1. Total quantity of purchasing (*TQ*) is positively related to *PLS*, meaning the relationship between *TQ* and *PLS* is  $TQ \uparrow \Rightarrow PLS \uparrow$ .
2. Frequency of purchasing, or purchasing times (*pp*), is positively related to *PLS*, meaning the relationship between *pp* and *PLS* is  $pp \uparrow \Rightarrow PLS \uparrow$ .
3. Quantity purchased in each period (*q*) is positively related to *PLS*, meaning the relationship between *q* and *PLS* is  $q \uparrow \Rightarrow PLS \uparrow$ .
4. The time frame of purchasing (*n*) is inversely related to *PLS*: if two customers show identical *TQ*, *pp*, and *q*, the one with the shorter time frame of purchasing will be considered more loyal to the product. The relationship between *n* and *PLS* is  $n \downarrow \Rightarrow PLS \uparrow$ .

Thus, we measure *PLS* with this formula:

$$SPLS = \left( \frac{pp}{n} \right) \times \sum_{j=1}^n \left( q_j \times \frac{j}{n \times (1+n)/2} \right) \quad (1)$$

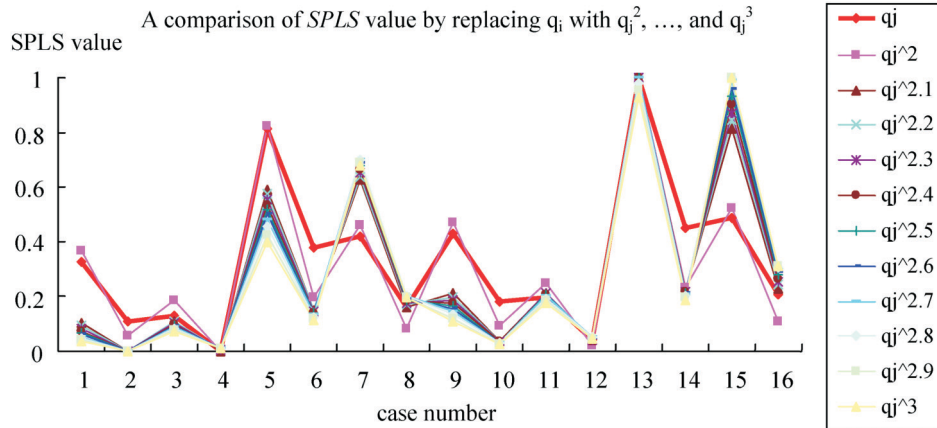
where *pp* represents frequency of purchasing, *q<sub>j</sub>* represents quantity purchased in the *j*th time period, and *n* represents the time frame of purchasing. The result is a non-negative real number. *PLS* is calculated for each customer. Customer *i*'s *PLS* is expressed as *PLS<sub>i</sub>* and it indicates the degree of customer *i*'s loyalty to the product. The length of a period is a product's turnover period, which is determined by the system designer depending on each product's unique features (the turnover

period for consumer goods is shorter than that for durable goods). *n* is found by dividing the time period between a customer's first and last purchases by the turnover period. For example, a customer has purchased a product several times over two years, and the system designer decides to set one month as the turnover period for this product. For this product, customer *i*'s value of *n* is 24. A high *n* shows that a customer has been purchasing the product over a long period of time.  $\sum q$  represents the total quantity purchased by the customer. Despite taking account of the quantity purchased in each period (*q*), total quantity of purchasing (*TQ*) also significantly influences *PLS*. Moreover, the quantity purchased more recently has more influence on the calculation of *PLS*. In other words, data closer to the last purchasing record is weighted more heavily than the older data closer to the first purchasing record. *pp* is the number of periods in which customer *i* purchased the product; a high (*pp*/*n*) value indicates intensive purchasing activity. *PLS* depends on two factors: intensity and strength of purchasing. (*pp*/*n*) represents customer *i*'s intensity of purchasing, and the section included in  $\sum$  represents customer *i*'s strength of purchasing. In order to normalize *PLS* value to the interval [0, 1], *PLS* will be replaced by standard *PLS* (*SPLS*) as calculated by formula 2:

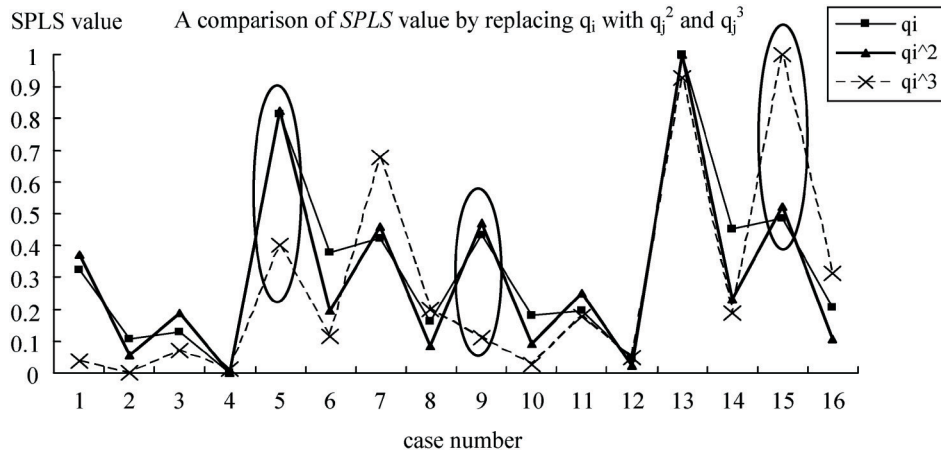
$$SPLS = \frac{\left( \frac{pp}{n} \right) \times \sum_{j=1}^n \left( q_j \times \frac{j}{n \times (1+n)/2} \right) - PLS_{min}}{PLS_{max} - PLS_{min}} \quad (2)$$

where *PLS<sub>max</sub>* is the maximum *PLS* value found among all loyal customers and *PLS<sub>min</sub>* is the minimum *PLS* value found among all loyal customers. *PLS<sub>i</sub>* lies between these two bounds:  $PLS_{min} \leq PLS_i \leq PLS_{max}$ . The *i*th loyal customer's *SPLS* is denoted *SPLS<sub>i</sub>*. A higher *SPLS* value implies a higher loyalty.

The *SPLS* value is further tested by replacing *q<sub>i</sub>* with *q<sub>i</sub><sup>2</sup>*, *q<sub>i</sub><sup>2.1</sup>*, *q<sub>i</sub><sup>2.1</sup>*... and *q<sub>i</sub><sup>3</sup>* to choose the best representation of the value. The result is depicted in Figure 1. In this figure, it is obvious that the *SPLS* values of using *q<sub>i</sub><sup>2.1</sup>*... and *q<sub>i</sub><sup>3</sup>* are almost the same in most cases and show slight difference in case number 5 and 15. Thus, in order to simplify and clarify the result, Figure 2 is a simplified form of Figure 1 by eliminating the *SPLS* values of using *q<sub>i</sub><sup>2.1</sup>*... and *q<sub>i</sub><sup>2.9</sup>*. In Figure 2, case number 5, 9, 15 show a significant difference between the *SPLS* values of using



**Figure 1.** The graph of different *SPLS* values by replacing  $q_i$  with  $q_i^2$ ,  $q_i^{2.1}$ , ...,  $q_i^3$ .



**Figure 2.** The graph of different *SPLS* values by replacing  $q_i$  with  $q_i^2$  and  $q_i^3$ .

$q_i^2$  and  $q_i^3$  and the *SPLS* values of  $q_i^3$  in these cases change the desired result. For example, the *SPLS* value in case 15 is larger than the *SPLS* value in case 13 that should be the largest value among all cases. Thus, we decide to replace  $q_i$  with  $q_i^2$  and the preliminary experimental evaluation that we conduct using  $q_i^2$  in the calculation of *SPLS* value has shown a desirable result.

### 3.2 Using K-means Clustering to Build a Profile of Loyal Customers

By using the *SPLS* value of each loyal customer, we can use cluster analysis to group loyal customers of similar personal backgrounds and purchasing behavior as measured by loyalty status. Personal background is expressed in categorical data and *SPLS* is expressed in numerical data. We adopt K-means analysis to perform the

task. The K-means clustering method can accept as input both numerical and analytical data. The number of clusters is determined by the Cubic Clustering Criterion (CCC). Using 85 loyal customers who are active buyers of a leading product at an online shop located in Taiwan as the input data of cluster analysis, the statistics package SPSS produces the clusters illustrated in Figure 3. The meaning of interval data of input attributes is listed below:

- AGE: under 20, 20–30, 31–40, above 40 (1/2/3/4)
- INCOME (monthly income): less than \$20,000, \$20,000–\$30,000, \$30,001–\$40,000, above \$40,000 (1/2/3/4)
- EDU (education): under high school, high school, college or above (1/2/3)
- MARITAL (marital status): yes, no (0/1)
- GENDER: male, female (0/1)

Cluster Center				
	Cluster			
	1	2	3	4
AGE	2	3	2	1
INCOME	2	4	2	1
EDU	3	3	3	2
MARITAL	1	0	1	1
GENDER	1	1	0	1
SPLS	.68	.88	.56	.66

(1)

Distance between cluster centers				
cluster	1	2	3	4
1		1.411	1.059	2.412
2	1.411		2.215	3.535
3	1.059	2.215		2.228
4	2.412	3.535	2.228	

(2)

Observation points in each cluster	
cluster	
1	11.000
2	27.000
3	5.000
4	42.000
Valid points	85.000
Invalid points	.000

(3)

Figure 3. The result of K-means cluster analysis.

K-means clustering calculates the Squared Euclidean Distance (SED) between every seed point and each cluster center and assigns seed points to the closest cluster. Next, it re-calculates each cluster center. Then it re-calculates the SED between each seed point and each cluster center, re-assigning seed points if they are closer to a new center. Re-calculation of centers and re-assignment of seed points are repeated until no seed point is assigned to a new cluster. Thus, the degree of similarity between two loyal customers is measured by SED. Table (1) in Figure 3 shows the value of input attributes of each cluster center. In this case there are four clusters.

- Cluster 1: Their product loyalty status is middling.
- Cluster 2: They are among the most loyal of all customers.
- Cluster 3: They are among the least loyal of all customers.
- Cluster 4: Their product loyalty status is middling.

Table (2) shows the SED between cluster centers. The distance between cluster 1 and cluster 3 is 1.059, the shortest distance between two clusters. The distance between cluster 2 and cluster 4 is 3.535, the longest distance between two clusters. According to Table (1), the personal backgrounds of clusters 1 and cluster 3 are mostly identical, except for gender, and *SPLS* values in

these two clusters are quite close. As for cluster 2 and cluster 4, the personal backgrounds of these two clusters are very different, except for gender, and the difference between *SPLS* values among them is large. Table (3) in Figure 3 shows the number of seed points in each cluster. Cluster 4 is the largest group, cluster 3 the smallest. The cluster of the most loyal customers, cluster 2, has 27 customers, which is more than 30 percent of all customers.

### 3.3 Using Similarity Analysis to Estimate the Standard Product Loyal Status (*SPLS*) Value of a Potential Customer

The result of K-means cluster analysis is a profile of loyal customers. Loyal customers in each group have unique characteristics. However, potential customers who have never purchased the analyzed product before have no purchasing record with which to calculate an *SPLS* value. Thus, we use similarity analysis to compare the personal background of each potential customer to the backgrounds in groups of loyal customers in order to estimate an *SPLS* value. In short, the goal of similarity analysis is to estimate a potential customer's *SPLS* value. Each potential customer has *m* characteristics but doesn't have an *SPLS* value.

These are the steps to calculate a potential customer's estimated *SPLS* value:

1. In each cluster, calculate the SED between the center and every seed point. The calculation of  $SED$  takes account of every characteristic except  $SPLS$ .
2. Find  $SED(k)_{max}$ , where  $k$  represents the  $k$ th cluster and  $SED(k)_{max}$  is the maximum  $SED$  value between the center and any seed point of the  $k$ th cluster.  $SED(k)$  lies between two boundaries,  $SED(k)_{min} \leq SED(k) \leq SED(k)_{max}$ . Thus,  $SED(1)_{max}$  is the maximum SED value in cluster 1, etc.
3. For each customer who has never purchased the analyzed product, calculate the  $SED$  between this customer and every cluster center:  $SED_{p1}, SED_{p2}, \dots, SED_{pk}$ , where  $p$  represents the  $p$ th customer and  $k$  represents the  $k$ th cluster.  $SED_{p1}$  denotes the SED between customer  $p$  and the center of cluster 1.  $SED_{pk}$  is the SED between the  $p$ th potential customer and the  $k$ th cluster center. The calculation of SED takes account of every characteristic except  $SPLS$ .
4. If a customer  $p$  has the condition of  $SED_{pk} \leq SED(k)_{max}$ , it means this customer falls into the  $k$ th cluster and is a potential customer. This customer's estimated  $SPLS$  value is the  $SPLS$  value of the center of the  $k$ th cluster.
5. If a customer  $p$  has the condition of  $SED_{pk} > SED(k)_{max}$ , it means this customer does not belong to the  $k$ th cluster. If this customer does not belong to any cluster, he can not be identified as a potential customer, and thus he'll be assigned no  $SPLS$  value.
6. If a potential customer is included in more than one cluster, this customer's estimated  $SPLS$  value is found by averaging the  $SPLS$  values of the centers of all clusters that he is included in.

Table 1 gives an example of the calculation of a customer's estimated  $SPLS$  value. Table (1) in Table 1 shows the value of input characteristics of cluster centers of loyal customers, obtained via cluster analysis. The column  $SED(k)_{max}$  is the maximum  $SED$  between the center of the  $k$ th cluster and all seed points in that cluster. This number is the distance between the cluster center and the farthest seed point in that cluster. If a seed point's  $SED$  is less than  $SED(k)_{max}$ , this seed point is very likely to fall into this cluster. If the seed point's  $SED$  is greater

**Table 1.** An example of the calculation of the estimated  $SPLS$  values of potential customers

Group ( $k$ )	Age	Income	Education	Marital status	Gender	$SPLS$	$SED(k)_{max}$
1	2	2	3	1	1	0.68	3
2	3	4	3	0	1	0.88	4
3	2	2	3	1	0	0.56	4
4	1	1	2	1	1	0.66	2

(1) The value of input characteristics of centers of clusters of loyal customers.

Non-purchase customer ( $p$ )	Age	Income	Education	Marital status	Gender	$SED_{p1}$	$SED_{p2}$	$SED_{p3}$	$SED_{p4}$
A	3	4	3	0	1	6	0	7	15
B	3	4	2	1	1	6	2	7	13
C	2	2	2	1	0	2	8	1	3
D	2	4	1	1	0	9	7	8	12
E	1	2	1	0	0	7	13	6	4

(2)  $SED$  between potential customers and centers of clusters.

Customer ( $p$ )	$SED_{p1}$	$SED(1)_{max}$	$SED_{p2}$	$SED(2)_{max}$	$SED_{p3}$	$SED(3)_{max}$	$SED_{p4}$	$SED(4)_{max}$	Estimated $SPLS$
A	6	3	<u>0</u>	4	7	4	15	2	<b>0.88</b>
B	6	3	<u>2</u>	4	7	4	13	2	<b>0.88</b>
C	<u>2</u>	3	8	4	<u>1</u>	4	3	2	<b>0.62</b>
D	9	3	7	4	8	4	12	2	<b>0</b>
E	7	3	13	4	6	4	4	2	<b>0</b>

(3) Calculation of potential customers' estimated  $SPLS$  values.



than  $SED(k)_{max}$ , this seed point does not belong to this cluster. Table (2) in Table 1 shows five potential customers' input characteristics and the  $SED$  to four cluster center.  $SED_{p1}$  is the  $SED$  between the  $p$ th potential customer and the center of cluster 1, etc. Table (3) combines Table (1) and Table (2). The last column in Table (3) is the estimated  $SPLS$  value of a potential customer. The  $SED$  value of potential customer A is less than the maximum  $SED$  value of cluster 2 only. Thus, potential customer A falls into cluster 2. For the same reason, potential customer B falls into cluster 2. Potential customer C falls into clusters 1 and 3. The estimated  $SPLS$  value of potential customer C is calculated by the average  $SPLS$  value of these two clusters. Customer D and customer E do not belong to any cluster.

#### 4. Conclusion

Spurred by the rapid development of Internet technology, online shopping is becoming more and more popular. Online marketplace expands their services and product lines to satisfy customers' quickly-changing needs, and spend copious resources developing customer relations. For any online store, providing customers with products they are interested in is the key to success. This paper provides a procedure to discover potential customers who are very likely to purchase a certain product, so that they may be presented with that product.

This paper uses two methods of analysis: cluster analysis and similarity analysis. Cluster analysis is used to cluster loyal customers who have similar personal backgrounds and purchasing behavior. Similarity analysis is used to find the similarity between loyal customers and potential customers. We expect this procedure to boost the sales at an online store. Eventually, this targeting would help grow the base of loyal customers. An experimental result shows that there are more than 50 percent of potential customers who actually purchase the product and become a "real" customer. A prototype of our proposed model is used by a fast growing online retailer in Taiwan and is still in the experimental period.

#### References

- [1] Hruschka, H., "Market Definition and Segmentation Using Fuzzy Clustering Methods," *International Journal of Research in Marketing*, Vol. 3, pp. 117–134 (1986).
- [2] Ozer, M., "User Segmentation of Online Music Service Using Fuzzy Clustering," *Omega*, Vol. 29, pp. 193–206 (2001).
- [3] Weber, R., *Customer Segmentation for Banks and Insurance Groups with Fuzzy Clustering Techniques*, In: J. F. Baldwin (ed.): *Fuzzy Logic*, John Wiley and Sons, Chichester, pp. 187–196 (1996).
- [4] Espinoza, M., Joye, C., Belmans, R. and De Moor, B., "Short-Term Load Forecasting, Profile Identification, and Customer Segmentation: A Methodology Based on Periodic Time Series," *IEEE Transactions on Power Systems*, Vol. 20, pp. 1622–1630 (2005).
- [5] Smirnov, A., Pashkin, M., Chilov, N., Levashova, T., Krizhanovsky, A. and Kashevnik, A., "Ontology-Based Users and Requests Clustering in Customer Service Management System," *Autonomous Intelligent Systems: Agents and Data Mining*, International Workshop, St. Petersburg, Russia, June, pp. 231–246 (2005).
- [6] Tsai, C. Y. and Chiu, C. C., "A Purchase-Based Market Segmentation Methodology," *Expert Systems with Applications*, Vol. 27, pp. 265–276 (2004).
- [7] Rosaci, D., "An Ontology-Based Two-Level Clustering for Supporting E-Commerce Agents' Activities," *Proc. of the 6th International Conference on E-Commerce and Web Technologies*, Copenhagen, Denmark, Aug., pp. 31–40 (2005).
- [8] Xu, B., Zhang, M. M., Pan, Z. G. and Yang, H. W., "Content-Based Recommendation in E-Commerce," *International Conference on Computational Science and Its Applications*, Singapore, May, pp. 946–955 (2005).
- [9] Weng, S. S. and Liu, M. J., "Feature-Based Recommendations for One-to-One Marketing," *Expert Systems with Applications*, Vol. 26, pp. 493–508 (2004).
- [10] Buccafurri, F., Rosaci, D., Sarne, G. M. L. and Ursino, D., "An Agent-Based Hierarchical Clustering Approach for E-Commerce Environments," *Proc. of the Third International Conference on E-Commerce and Web Technologies*, Aix-en-Provence, France, Sep., pp. 109–118 (2002).
- [11] Kuo, R. J., Liao, J. L. and Tu, C., "Integration of ART2 Neural Network and Genetic K-Means Algorithm for

- Analyzing Web Browsing Paths in Electronic Commerce,” *Decision Support Systems*, Vol. 40, pp. 355–374 (2005).
- [12] Shin, H. W. and Sohn, S. Y., “Segmentation of Stock Trading Customers According to Potential Value,” *Expert Systems with Applications*, Vol. 27, pp. 27–33 (2004).
- [13] Ray, P. S., Aiyappan, H., Elam, M. E. and Merritt, T. W., “Application of Cluster Analysis in Marketing Management,” *International Journal of Industrial Engineering - Theory Applications and Practice*, Vol. 12, pp. 127–133 (2005).
- [14] Kim, K. J. and Ahn, H., “Using a Clustering Genetic Algorithm to Support Customer Segmentation for Personalized Recommender Systems,” *13th International Conference on AI, Simulation, and Planning in High Autonomy Systems*, AIS 2004, Jeju Island, Korea, October, pp. 409–415 (2004).
- [15] Niu, B. and Shiu, C. K., “Using Similarity Measure to Enhance the Robustness of Web Access Prediction Model,” *Knowledge-Based Intelligent Information and Engineering Systems, 9th International Conference*, Melbourne, Australia, Sep., pp. 107–111 (2005).
- [16] Yi, S. Z., Huang, B. and Chan, W. T., “XML Application Schema Matching Using Similarity Measure and Relaxation Labeling,” *Information Sciences*, Vol. 169, pp. 27–46 (2005).
- [17] Zeng, C., Xing, C. X., Zhou, L. Z. and Zheng X. H., “Similarity Measure and Instance Selection for Collaborative Filtering,” *International Journal of Electronic Commerce*, Vol. 8, pp. 115–129 (2004).
- [18] Velasquez, J. D., Yasuda, H., Aoki, T. and Weber, R., “A New Similarity Measure to Understand Visitor Behavior in a Web Site,” *The Institute of Electronics, Information and Communication Engineers Transactions on Information and Systems*, Vol. E87-D, pp. 389–396 (2004).
- [19] Richard L. Oliver, *Satisfaction: A Behavioral Perspective on the Consumer*, New York: McGraw-Hill (1997).

**Manuscript Received: Nov. 28, 2007**

**Accepted: Jan. 3, 2008**