Computing and Informatics, Vol. 29, 2010, 509–520

A MULTI-FACTOR CUSTOMER CLASSIFICATION EVALUATION MODEL

Qiaohong Zu, Ting Wu, Hui Wang

School of Logistics Engineering
Wuhan University of Technology
430 063 Wuhan, China
e-mail: zuqiaohong@whut.edu.cn, wt0208@gmail.com, sophie@whut.edu.cn

Revised manuscript received 24 March 2010

Abstract. Pervasive application of data mining technology is very important in analytical CRM software development when the distributed data warehouse is constructed. We propose a multi-factor customer classification evaluation model CLV/CL/CC which comprehensively considers customer lifetime value, customer loyalty and customer credit. It classifies clients with synthetic data mining algorithms. In this paper, we present an extended Bayes model which substitutes the primary attribute group with a new attribute group to improve the classification quality of naive Bayes.

Keywords: Classification model, extended Bayes model, customer classification prediction, weighted Bayes algorithm, lifetime value, customer loyalty degree, client capital credit, fuzzy neural network, Markov chain

1 INTRODUCTION

In the pervasive computing application, on-line analysis and data mining are critical technologies of analytical CRM system based on distributed data warehouses. Useful rules can be mined from a large volume of data through various models and algorithms, which can support the decision-making process [1].

Customer value is currently the major factor in customer classification. However, the impact of customer behavior on customer value should not be ignored. Therefore, combining customer value and customer behavior provides more accurate customer classification [2, 3].

The traditional classification is mainly based on client value that takes time, frequency and quantity of purchase into account. However, the overall degree of customer loyalty has a great effect on enterprise profits, and customer capital credit can be used to measure commercial risk [4].

Enterprises analyze clients and classify different customer bases in order to make decisions with the analytical CRM system.

2 CUSTOMER CLASSIFICATION MODEL BASED ON CUSTOMER LIFETIME VALUE/CUSTOMER LOYALTY DEGREE/CUSTOMER CAPITAL CREDIT (CLV/CL/CC)

The CLV/CL/CC customer classification model considers three factors – customer lifetime value, customer loyalty and customer credit. They are based on customer capital management theory, customer life cycle theory and customer loyalty theory. This model classifies clients by applying data mining algorithm synthetically. The calculation results of indexes of the factors are used in customer clustering with improved K-means to obtain customer clustering groups, which is the pre-treatment of classification forecasting with extended Bayes model. Combining the two algorithms can improve the effectiveness of customer classification [5, 6].

2.1 Customer Lifetime Value Model with the Expenditure Assignment

The customer life-long value is the key factor to carry out the customer relationship management plan successfully. However, its influencing factors are so complex that there is seldom effective quantitative analysis algorithm at present. In this paper, we propose a lifetime value model with expenditure assignment, which is a relative scientific quantitative calculation method.

The customer expenditure assignment refers to the percentage of the quantity a customer purchasing a brand from certain company in his/her total quantity has bought, which reflects the relation between the customer and the corporation. When we calculate customer life-long value, we should not ignore the customer expenditure assignment [6]. We study the quantitative calculation of the customer life-long value based on the purchase transformation of customers.

$$\mathrm{CLV} = \sum_{t=0}^{T} \left[(1+r)^{-t} \left(\sum_{i=1}^{F_t} S_{it} \right) R_t \right] \tag{1}$$

In Equation (1), CLV (customer lifetime value) means current value of profit contribution of a customer bringing to the corporation in some periods [4].

- t – period (a month, a quarter or a year)
- r – discount rate
- T – calculation time in an enterprise plan (how many time cycles)

- F_t – expected frequency of customers' purchasing in period t
- S_{it} – expected expenditure share in customers purchase of a brand in purchase cycle i and period t
- R_t – the average profit contribution that customers purchase in period t.

Even though the amount of itinerant customers is constant in the market, their expenditure assignments may not be the same in different purchase times. The brand choices of mobile clients can be affected by following factors: promotions, prices, services of the brand of this purchase time, but not by previous conditions. The "without aftereffect" characteristic of mobile client's expenditure assignment meets with the supposed conditions of Markov chain theory. Therefore, the Markov model can be used to predict the change of expenditure assignments of mobile clients [7].

Markov chain is a special random process, which shows that the state of things changes from past to present, and then changes from now to future. Its distinct characteristic is "without aftereffect". Generally speaking, the state at the end of period n is not related to the state before period $n-1$, but is related to both the variance of increasing and decreasing in period n and to the state at the end of period $n-1$.

Provided the system has r states altogether, and p_{ij} is the mutual transition probability among various states $(i, j = 1, 2, \ldots, r)$, then the following matrix is called the purchase transfer probability matrix.

Supposed that there are n brands clients to choose, the matrix of purchase conversion is as follows:

$$P = \begin{bmatrix} P_{11} & P_{12} & \ldots & P_{1j} & \ldots & P_{1n} \\ P_{21} & P_{22} & \ldots & P_{2j} & \ldots & P_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ P_{i1} & P_{i2} & \ldots & P_{ij} & \ldots & P_{in} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ P_{n1} & P_{n2} & \ldots & P_{nj} & \ldots & P_{nn} \end{bmatrix} \tag{2}$$

P_{ij} refers to the transfer possibility of customers buying brand i this time and buying brand j next time, and P_{ii} expresses the probability to maintain buying brand i, and $\sum_{j=1}^{n} P_{ij} = 1$.

In Equation (2), $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n$. The top row of this matrix represents the situation that one of the customers buys brand 1 this time, but may buy other brand next time. P_{11} means the possibility that the client still purchases brand 1 next time, P_{12} expresses the possibility that the client buys brand 2 next time, and so on. Therefore, we are able to calculate the customer expenditure assignment with the customer purchase transfer matrix, and we can calculate customer lifetime value with the expenditure assignment [3, 8].

2.2 Calculation of Customer Loyalty Degree Using Fuzzy Neural Networks

The reasons to predict customer loyalty using fuzzy neural network are as follows:

- First, customer loyalty model is complex nonlinear function, and different customer groups have different functions. Therefore, to simulate the change law of customer loyalty is almost impossible with linear system only. However, artificial neural network, which has the merits of non-linear imitation capability, pattern recognition ability and the approximate nonlinear mapping ability in arbitrary precision, is an ideal tool for customer loyalty simulation.
- Second, there are many subjective and quantitative influence factors of customer loyalty. However, there is no explicit and precise description of these indicators. In order to be consistent with the way of people's thinking in customer loyalty forecasting, these indicators should be fuzzy. Fuzzy neural network is an organic combination of fuzzy technology and neural network technology. We can construct a kind of neural network or self-adaptation fuzzy system, which can deal with fuzzy information "automatically". Therefore, the fuzzy neural network method is suitable for the research of customer loyalty evaluation prediction.

Customer loyalty degree is a comprehensive evaluation value according to trust, transaction frequency, service effect, satisfaction degree and the possibility of customers accepting the same enterpriser's service. We set up the index system of customer loyalty degree and propose the forecasting customer loyalty model based on fuzzy neural net.

We select 10 operable factors: purchase duration, purchase frequency, change trend of purchase frequency, wallet share, cross-selling, expenditure assignment, mention rate of product, price sensitivity, customer satisfaction, etc., as parameters to construct neural network model (with initial index weights) to predict customer loyalty degree (as shown in Figure 1).

The value of parameters needs to be standardized in the same range in neural network, otherwise some parameters whose value is too large will dominate the training process. There are two types of input data: quantitative data and qualitative data, which should be pre-treated by standardized fuzzy processing to be between 0 and 1. Because the dynamic range of recognizing data is reduced, the possibility of successful prediction will be improved.

After the attribute indexes of training samples are standardized, they will be input to the SQL Sever 2005 to perform model training with initial weights. When the final index weights are determined by training, we can obtain the final training model to predict quantitative loyalty degrees of other customers.

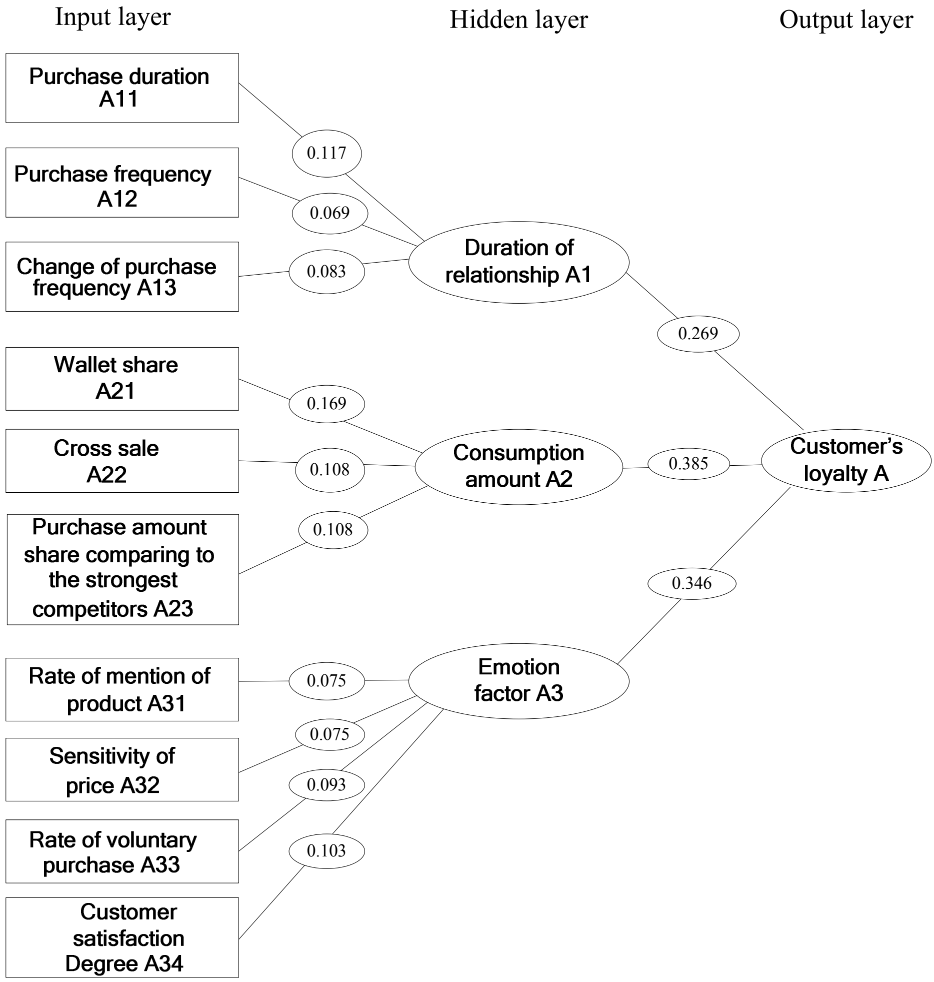

Fig. 1. Network model of the loyalty degree calculation

2.3 Calculation of Customer Capital Credit with Fuzzy Evaluation

Before we apply the theory of fuzzy appraisement to evaluate capital credit calculation of enterprise customers or individual customers, we should analyze the index affiliations influencing customer capital credit with fuzzy theory first.

Because the evaluation value of customer capital credit is fuzzy, and the index and grade of measuring customer capital credit are also fuzzy, customer capital credit should be measured with the fuzzy mathematics method, and the index weight and relevancy are determined by preferential comparison method [9].

Referring to the literature, personal credit score indicators are selected combined with the specific circumstances of the case. The process of calculating capital credit of individual customers by the fuzzy evaluation method is as follows:

- First, we set up 13 classic indicators in the four aspects being the characteristic variables of model. Adopting the opinions of domain exports and considering real circumstances of the case, we can determine the specific indicators.
- Second, we consolidate 13 related personnel's opinions, which include manufacturers, professionals, marketing personnel and long-term clients, to evaluate customer capital credit. Comparing the mutual importance of customer capital credit indexes, the index weights are calculated by fuzzy comparative method and system analysis method (as shown in Table 1).

index	first grade weight	second grade weight	comprehensive weight
sex	0.224	0.154	0.035
age	0.224	0.188	0.042
civil state	0.224	0.196	0.044
domicile	0.224	0.15	0.034
education	0.224	0.312	0.07
profession	0.186	0.538	0.1
working time	0.186	0.462	0.086

Table 1. Individual customer capital credit index weights

2.4 Determining Index Weights of the Three Factors

Attribute importance shows the effect of attribute on decision in rough set theory. Supposing $SGF(a_i, D)$ shows the relative importance of some attribute to decision attribute D. If the attribute has m condition attributes, then its weight can be calculated by help of Equation (3):

$$w_i = \frac{SGF(a_i, D)}{\frac{1}{m}\sum_{i=1}^{m} SGF(a_i, D)}. \tag{3}$$

The index weight can be determined applying the concept of attribute importance in rough set theory. The process of calculation is as follows:

1. The knowledge expression system of the father index attribute can be established, which is the decision attribute. The son indexes constitute condition attribute set C, $C = a_1, a_2, \dots a_i \dots, a_n$.

2. The knowledge expression system is processed numerically, then the repeat lines should be deleted.
3. The positive region $pos_{C-a_i}(D)$ is determined and the importance of every condition attribute $r_C(D) - r_{C-a_i}(D))$ is calculated.
4. Normalization processing is performed. Set $P_A = \sum_{i=1}^{n} P_i$; the weight of son index can be calculated by $W_{i0} = P_i/P$ [10, 11].

20 customer samples are selected, which can be divided into three grades according to their customer lifetime value, customer loyalty degree and customer capital credit (as shown in Table 2).

attribute	condition attribute C						decision attribute D	
sequence number	customer lifetime value		customer loyalty degree		customer capital credit degree		customer classification	
	real value	dispersing	value value	dispersing	real value	dispersing	real value	dispersing
1	500	2	0.75	2	82	1	silvery	2
2	647	1	0.74	2	86	1	golden	1
3	350	3	0.56	3	65	2	common	3
4	420	2	0.67	2	80	1	silvery	2
5	467	2	0.70	2	76	2	silvery	2
6	320	3	0.56	3	62	3	common	3
7	330	3	0.68	2	64	2	common	3
8	353	3	0.55	3	50	3	common	3
9	874	1	0.85	1	86	1	golden	1
10	356	3	0.56	3	52	3	common	3
11	368	3	0.52	3	53	3	common	3
12	510	2	0.39	3	65	2	common	3
13	821	1	0.75	1	75	2	golden	1
14	536	2	0.66	2	74	2	silvery	2
15	423	2	0.42	3	62	2	common	3
16	658	1	0.7	2	76	2	silvery	2
17	454	2	0.56	3	55	3	common	3
18	545	2	0.80	1	81	1	silvery	2
19	863	1	0.8	1	78	2	golden	1
20	230	3	0.4	3	46	3	common	3

Table 2. Discrete simplified table of customer classification indexes

Customer lifetime value: 1high (above 600), 2middle (400–600), 3low (below 400).

Customer loyalty degree: 1high (above 0.8), 2middle (0.6–0.8), 3low (below 0.6).

Customer capital credit: 1high (above 80), 2middle (60–80), 3low (below 60).

After simplifying Table 2, some formulas can be gained as follows:

$$U/\text{ind}(b,c) = \{(1,2),(3,8),(4,6,10),(5,11)(7),(9)\}$$

$$U/\text{ind}(a,b) = \{(1,4),(2,10),(3,5),(6),(7,9)(8,11)\}$$

$$U/\text{ind}(a,c) = \{(1),(2,7),(3,6),(4,8),(5),(9,10),(11)\}$$

$$U/d = \{(1,4,10),(2,7,9),(3,5,6,8,11)\}$$

$$U/C = \{(1),(2),(3),(4),(5),(6),(7),(8),(9),(10),(11)\}$$

then:

$$\mathrm{pos}\{b,c\}(d) = \{3,5,7,8,9,11\} = 6$$
$$\mathrm{pos}\{a,c\}(d) = \{1,2,3,5,6,7,11\} = 7$$
$$\mathrm{pos}\{a,b\}(d) = \{1,3,4,5,6,7,8,9,11\} = 9$$
$$\mathrm{posC}(d) = \{1,2,3,4,5,6,7,8,9,10,11\} = 11$$

and then: $rC(d) = 11/11 = 1$

$$r\{b,c\}(d) = 6/11 = 0.545$$
$$r\{a,c\}(d) = 7/11 = 0.636$$
$$r\{a,b\}(d) = 9/11 = 0.818$$
$$rC(d) - r\{b,c\}(d) = 1 - 0.545 = 0.455$$
$$rC(d) - r\{a,c\}(d) = 1 - 0.636 = 0.364$$
$$rC(d) - r\{a,b\}(d) = 1 - 0.818 = 0.182$$

The result shows the important degrees of the customer lifetime value, customer loyalty and customer capital credit to customer classification decision. The sequence of importance from high to low is: customer lifetime value, customer loyalty degree, and customer capital credit. After normalization processing, the calculation results of the attribute weights are 0.455, 0.364, 0.181.

3 CUSTOMER CLASSIFICATION PREDICTION BASED ON EXTENDED BAYES CLASSIFIER

Bayes theory was put forward by Thomas Bayes, an English priest and mathematician, in 1763. In Bayes theory, there are several assumptions. If the result of an event is uncertain, the only way to quantify the event result is to calculate its probability. If the frequency of the event occurred in past tests is known, the occurrence probability of the event in future can be predicted by mathematical method [12].

Naive Bayes classifier is a classification method based on Bayes theory, which shows excellent properties in many fields. Compared with other classifications theory, Bayes classifier has the lowest error rates. Actually, because the assumption of independent condition usually can not be realized and there is lack of some accurate probability distribution of events, the accuracy of Bayes classifier prediction will decrease. There are many references on improving the performance of naive Bayes classifier [6].

In naive Bayes classifier algorithm, the importance of every attribute is equal, and every weight is 1. Actually, the effect of every attribute on classification is different. When the weights are rearranged with comparing mathematical expect values of attribute weights, we obtain an extended Bayes classifier in which the new attribute set instead of the old one. It is an excellent classification model.

3.1 Classifying Prediction of Extended Bayes Classifier

The necessary steps of classification prediction based on extended Bayes classifier are as follows:

- First, each n-dimension eigenvector ($X = x_1, x_2, \ldots, x_n$) should meet the threshold measurement of the correlation analysis of attribute, which describes the properties of a customer.
- Second, given the customer values, which are classified as k levels and consistent with the groups of previous customers clustering, the probability of each new customer or potential customer which belongs to corresponding customer value level can be predicted, customers can be divided into different value levels according to their maximum probability. The probability of a customer in one customer value level can be calculated based on Bayes theorem by help of Equation (4).

$$P(C_i\,|X\,) = \frac{P(X\,|C_i\,)P(C_i)}{P(X)} \qquad \forall i \tag{4}$$

Because all customer value degrees $P(X)$ are constants, we just need to calculate its maximum value of $P(X|C_i)P(C_i)$. The prior probability of each customer value degree is $P(C_i) = s_i/s$. s_i is the training sample number of C_i, which is customer value degree in training set, and s is the general training sample number.

Third, if a training set contains many attributes, the calculation of $P(X|C_i)$ will be complex. In order to simplify the calculation process, we suppose the customer attributes are mutual condition independent, then:

$$P(X\,|C_i\,) = \prod_{k=1}^{n} p(\omega_k \cdot x_j\,|C_j\,) \qquad \forall i \tag{5}$$

$$p(\omega_k x_j/C_i) = \omega_k s_{ij}/s_i \tag{6}$$

where s_{ij} is the number of training samples when the value of attribute $A_j(A_j \in C_i)$ equals to x_j. s_i is the number of total training samples, ω_k is the weight of attribute A_j. If A_j is a continuous function, suppose it is Gauss distribution, then:

$$p(\omega_k \cdot x_j\,|C_i\,) = g(\omega_k x_j, \mu_{C_i}, \sigma_{C_i}) = \frac{1}{\sqrt{2\pi\sigma_{C_i}}} e^{\frac{(\omega_k x_j - \mu_{C_i})^2}{2\sigma_{C_i}{}^2}} \tag{7}$$

where $g(x_j, \mu_{C_i}, \sigma_{C_i})$ is the Gauss density function of attribute A_j. μ_{C_i}, σ_{C_i} are the mean and standard deviation after attribute A_j multiplies the weight ω_k [7].

The probabilities of a customer in every customer value degree can be calculated by the formula. After the calculation result is normalized, the customer can be categorized into the corresponding customer group according to the final probability.

3.2 Classifier Verification and Analysis

Comparing the effects of different Bayes algorithms (as shown in Figure 2), the top line represents the naive Bayes model, the middle line represents the improved extended Bayes model, and the bottom line represents the ideal model. Obviously, the extended Bayes raises the accuracy of classification prediction.

The extended Bayes model is proved valid by using the customer's personal detailed information and transaction information from the source database of the case. The source database from certain manufacturer contains 20181 customers' information. 12824 customers of the source database are chosen as testing samples whose login time is over 1 year.

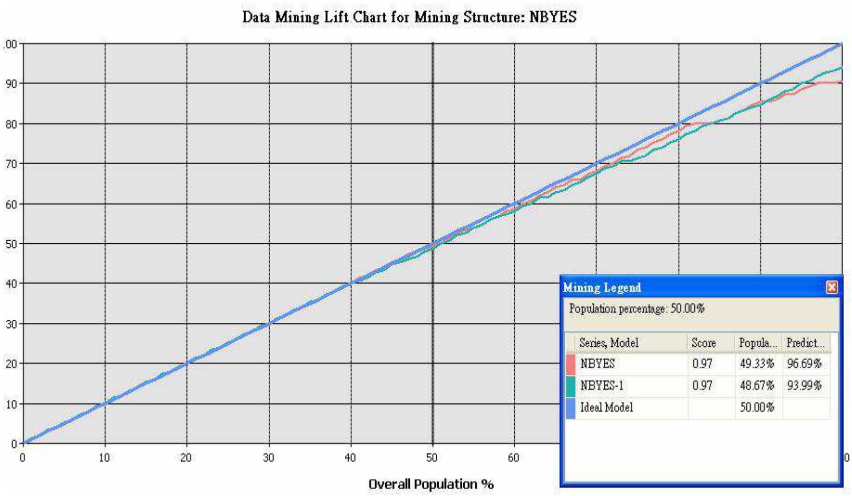

Fig. 2. Comparison of the precision of ideal model, naive Bayes model and extended Bayes model

4 CONCLUSIONS

We proposed an extended Bayes algorithm with weights. The algorithm organically combines a cluster preprocess with classification prediction. It can be applied to classification prediction of customers and to raise accuracy. In the end of this paper, we also provided an example that demonstrates the effectiveness of the multi-factor customer classification model.

REFERENCES

[1] Tang, B.—Wang, S.: Customer Relationship Management. Higher Education Publishing House, China, 2001.

[2] Chicco, G.—Napoli, R.—Piglione, F.: Comparisons Among Clustering Techniques for Electricity Customer Classification. IEEE Transactions on Power Systems, Vol. 21, May 2006, pp. 933–940.

[3] Qiaohong, Z.—Dingfang, Ch.: Customer Life-Long Value Computing Research Based on Expenditure Assignment. International Conference on Wireless Communications 2008, Networking and Mobile Computing, WiCOM 2008, Dalian China, 2008, pp. 468–472.

[4] Qiaohong, Z.—Wenfeng, L.: The Research of Customer Classification Based on Extended Bayes Model. 3rd International Conference on Pervasive Computing and Applications 2008, ICPCA 08, Alexandria Egypt, October 2008, Vol. 1, pp. 22–25.

[5] Qiaohong, Z.—Wenfeng, L.: Design and Development of Analytical CRM Based on OLAM Technology. Journal of Southeast University (English Edition), Vol. 23, Dec 2007, pp. 168–172.

[6] Jian-Guo, Zh.—Tao, B.: Machine Learning and Cybernetics: Customer Classification in Commercial Bank Based on Rough Set Theory and Fuzzy Support Vector Machine. International Conference 2008, Vol. 2, July 2008, pp. 1212–1217.

[7] Wang, S. L.—Wang, X. Z.: A Fuzzy Comprehensive Clustering Method: Advanced Data Mining and Applications. Third International Conference, ADMA 2007, Proceedings, Vol. 4632, 2007, pp. 488–499.

[8] Bing, F.—Xing, L. X.: Application of Reliability-Based Criteria to the Customer Classifications. Safety and Reliability for Managing Risk, Vol. 1 C3, 2006, pp. 2313–2318.

[9] Derrig, R. A.—Ostaszewski, K. M.: Fuzzy Techniques of Pattern Recognition in Risk and Claim Classification. The Journal of Risk and Insurance, Vol. 62, No. 3, 1994, pp. 447–482.

[10] Zhang, X. S.—Cheng, H. F.: Fuzzy Comprehensive Evaluation of Discipline Competitiveness based on AHP. Advancing Knowledge Discovery and Data Mining Technologies, Proceedings, 2009, pp. 56–59.

[11] Tsymbal, A.—Puuronen, S.: Ensemble Feature Selection With the Simple Bayesian Classification in Medical Diagnostics. 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002), 2002 Maribor Slovenia, 2002, pp. 225–230.

[12] Yang, S. X.: Mixed Clustering Algorithm and Electricity Customer Classification. International Seminar on Future Information Technology and Management Engineering 2008, proceedings, Leicestershire United kingdom, 2008, pp. 107–110.

[13] Cheng, C. H.—Chen, Y. S.: Classifying the Segmentation of Customer Value Via RFM model and RS Theory. Expert Systems with Applications, Vol. 36, 2009, pp. 4176–4184.

[14] Han, J. W.—Kamber, M.: Data Mining: Concepts and Techniques. Morgan Kaufmann Publisher, 2001.

Qiaohong Zu is the director of the Institute of Logistics Information and Digital Engineering of Wuhan University of Technology. She received her M. Sc. and Ph. D. degrees, both in mechanical manufacture and automation. She also gives lectures at Wuhan University of Technology and is a supervisor and consultant for master and bachelor studies. She is (co-)author of scientific books and numerous scientific papers. She has participated in the National Key Technology R & D Program (No. 2006BAH02A06) and the National Natural Science Foundation of China (No. 60475031), as well as in several national projects. Her research interests include artificial intelligence, knowledge management and data mining.

Ting Wu is a Master's candidate at the Institute of Logistics Informatics and Digital Engineering in Wuhan University of Technology. She is also a researcher at the Institution of Intelligent Manufacture and Control. During the study period, she has participated in some projects in the field of data mining, such as the project of the National Key Technology R & D Program (No. 2006BAH02A06) and so on. She is (co-)author of scientific books and numerous scientific papers. Her primary research interests are data warehouse, on-line data, artificial intelligence.

Hui Wang is a Master's candidate at the Institute of Logistics Informatics and Digital Engineering in Wuhan University of Technology. In 2008 she received her Bachelor's Degree in transportation from the Lanzhou Jiaotong University. She worked as a technical student at the Intelligent Control for Data warehouse/Data mining Research. She is the author and co-author of several scientific papers about online analytical processing (OLAP) and complex net. Her research interests include data warehouse, data mining, artificial intelligence. She also gives lectures at Huazhong University of Science and Technology.