An Improved Particle Swarm Algorithm for Online Trading Customer Classification

Hankun Ye

School of international Trade and Economics, Jiangxi University of Finance and Economics, Nanchang, 330013 China

Abstract

Correctly and effectively customer classification according to their characteristics and behaviors will be the most important resource for electronic marketing and online trading of network enterprises. Aiming at the shortages of the existing particle swarm and K-means algorithm for customer classification, this paper advances a new customer classification algorithm through improving the existing particle swarm algorithm and combining it with K-means algorithm. First the paper designs 21 customer classification indicators based on consumer characteristics and behaviors analysis, including customer characteristics type variables and customer behaviors type variables; Second, limitation of particle swarm algorithm and K-means algorithm are analyzed; Then corresponding improvements for particle swarm algorithm are advanced including improvement of the speed update formula of particle and , improvement of balancing the development and detection capability of particle of the algorithm; Thirdly, the online trading customer classification algorithm combining the improved particle swarm algorithm and K-means algorithm is advanced. Finally the experimental results verify that the new algorithm can improve effectiveness and validity of customer classification when used for classifying network trading customers practically.

Keywords

Particle Swarm algorithm, K_means algorithm, Electronic marketing, Customer classification.

1. Introduction

Customer relations management is one of the core problems of modern enterprises, whose customer oriented thought requires CRM system to be able to effectively obtain various kinds of information of customers, identify all the relations between the customers and enterprises and understand the transaction relation between customers and enterprises; meanwhile, deeply analyze customers' consuming behavior, find customers' consumption characteristics, providing personalized service for customers, supporting the decisions of enterprises. The three basic problems CRM needs to solve are how to get customers, how to keep customers and how to maximize customer value, among which maximizing customer value is the ultimate purpose, getting customers and keeping customers are both the means for realizing the purpose. The core of analyzing the three problems CRM needs to solve is to classify customers. "Getting Customers" and "Retaining Customers" need to ascertain which customers are attainable, which customers need to be kept, which customers are kept for a long term and which customers are kept for a short term, therefore, customer classification is needed. It is the same case with "Maximizing Customer Value". Due to different values of different customers, "Maximum Customer Value" of different customers should be distinguished. Thus, the core problem of enterprises to correctly implement CRM is to adopt effective method to reasonably classify customers, find customer value, focus on high-value customers with enterprises' limited resources, provide better service for them, keep "High-value" customers for loss prevention; also, establish corresponding customer service system through classification, carry out differential customer service management. Hence, customer classification is becoming a more and more popular research hotspot, also a research difficulty, becoming one of the urgent problems of CRM[1].

2. Summarization of Customer Classification Methods

The widely-used methods of enterprises for customer classification at present are mainly qualitative method and quantitative method. As the qualitative method for customer classification is just to classify all the target customers of enterprises in the macroscopic level, customer classification is carried out according to different value emphasis of different customers. The formation of customer value is simply expressed as: Value = Benefit — Cost. Qualitative classification method classifies customers in a simple way, only offering guidance for customer classification of enterprise in the macroscopic level, unable to provide specific and credible basis for enterprise decisions; furthermore, as there is no strict process of argumentation, the method depends on decider's subjective inference, there may be certain deviations in the analysis process, easily resulting in faulty decisions. For this reason, to truly provide customer classification information beneficial to enterprises should depend on quantitative technology for customer classification.

Quantitative classification method is to apply quantitative analysis technology to conduct customer classification on the basis of some specific customer variables (credit level of customers, purchasing power of customers, characteristics of demand of customers, etc.). Currently, there are mainly two categories of data mining for quantitative customer classification research, which are traditional statistical method and non-statistical method. The former mainly includes cluster analysis, Bayesian Classification, factor analysis method, etc.; this statistics-based method is unable to process a great deal of sophisticated customer data, and there are some problems on the accuracy of customer classification results, so to fundamentally solve the problem of customer classification needs to rely on non-statistical customer classification method, which mainly includes neural network, fuzzy set method, association rules, genetic algorithm, etc. The classification technology based on neural network is combined with certain information technology, which is a kind of mathematical method applicable to complex variables and multi influencing factors calculation, so it is more effective in solving complex customer classification problems with better classification accuracy, however, the convergence problem of the function itself greatly limits its application value in specific project practice. Secondly, classification is mainly based on such mathematical methods as fuzzy clustering, rough set, association rules, etc., although these methods offer classification reason explanation in a relatively clear way with better classification results under the circumstances of satisfactory data conditions, the modeling process needs to provide specific mathematical equations. As a result, these methods are limited by data conditions in specific application, always having problems like insufficient classification accuracy or poor "robustness", limiting the application in customer classification. Due to lots of influencing factors related to customer classification, more often than not, the complicated relations are difficult to be expressed in mathematical equations[1-6].

Particle Swarm Optimization, PSO, is a new method used in customer classification because it can obtain global optimal solution and get high classification accuracy when it combined with K-means algorithm. But PSO algorithm has slow convergence rate and is easy to fall into local extreme point, so the paper tries to present some improvements to overcome the limitations of the algorithm and takes advantage of powerful classification ability of the algorithm to classify online trading customer.

3. Selection of Customer Classification Indicators

The selection of reasonable classification variables is the basis of correct and effective customer classification, namely establishing scientific and reasonable classification indicators system. In view of the nature of trading and own characteristics of online trading, this Paper adopts customer characteristics type variable and customer behaviors type variable in the specific selection of customer classification variables[2].

(1) Selection of Customer Characteristics Type Variable

Customer characteristics type variable is mainly used for getting the information of customers' basic attributes. Such variable indicators as geographical position, age, sex, income of individual customer play a key role in determining the members of some market segment. This kind of variables mainly comes from customers' registration information and customers' basic information collected from the management system of banks, the contents of which mostly indicate the static data of customers' basic attributes, the advantage of which is that most of the contents of variables are easy to collect. But some of the basic customer-described contents of variables are lack of differences at times.

Based on analyzing and summarizing existing literatures, the customer characteristics type variables designed in this Paper include: Customer No., Post Code, Date of Birth, Sex, Educational Background, Occupation, Monthly Income, Time of First Website Browsing, and Marital Status.

(2) Selection of Customer Behaviors Type Variables

Customer behaviors type variables mainly indicate a series of variable indicators related to customer transacting behavior and relation with banks, which are used to define the orientation which enterprises should strive for in some market segment, and are the key factors for ascertaining target market. Customer behaviors type variables include the records of customers buying services or products, records of customer service or production consumption, contact records between customers and enterprises, as well as customers' consuming behaviors, preferences, life style, and other relevant information.

Based on analyzing and summarizing existing literatures, the customer behaviors type variables designed in this paper include Monthly Frequency of Website Login, Monthly Website Staying Time, Monthly Times of Purchasing, Monthly Amount of Purchasing, Type of Consumer Products Purchased, Times of Service Feedback, Service Satisfaction, Customer Profitability, Customer Profit, Repeat Purchases, Recommended Number of Customers, Purchasing Growth Rate.

4. Customer Classification Algorithm Based on PSO

4.1 K-means algorithm

Steps for K-means clustering algorithm are[7-9] (see Fig.1):

(1) Select n objects as the initial cluster seeds on principle;

(2) Repeat (3) and (4) until no change in each cluster;

(3) Reassign each object to the most similar cluster in terms of the value of the cluster seeds; (4) Update the cluster seeds, i.e., recompute the mean value of the object in each cluster, and take the mean value points of the objects as new cluster seeds.

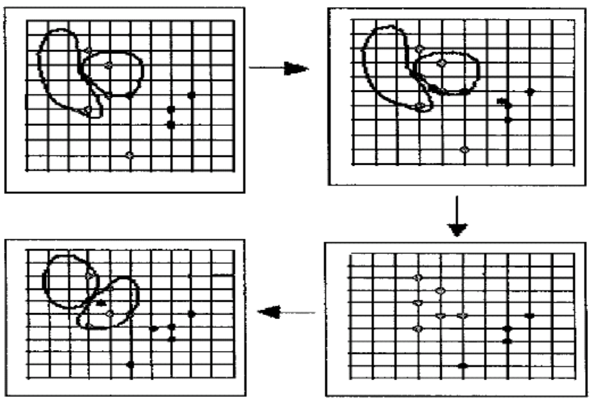

Figure 1. K-means Algorithm Procedures

When K-means algorithm is used to cluster data, the stability of the clustering results is still not good enough, sometimes, the clustering effect is very good (when the data distribution is convex-shaped or spherical), while sometimes, the clustering results have obvious deviation and errors, which lies in the data analysis. It is unavoidable for the clustered data to have isolated points, referring to the situation that a few data deviate from the high-dense data intensive zone. The clustering mean point (geometrical central point of all data in the category) is used as a new clustering seed for the K-means clustering calculation to carry out the next turn of clustering calculation, while under such a situation, the new clustering seed might deviate from the true data intensive zone and further cause the deviation of the clustering results. Therefore, it is found that using K-means algorithm to process the data of isolated points has a great limitation.

4.2 Particle Swarm Optimization

PSO is an evolutionary computation technology based on swarm intelligence, mainly used for obtaining global optimal solution, which was firstly developed from Kennedy and Eberhart simulating flying behavior of birds, the basic concept of which is that the potential solution of each optimization problem is a bird in search space, called “particle”; all the particles have an adaptive value determined by optimization function, and each particle has a speed vector determining their flying direction and distance; then particles follow the search of current optimal particle in solution space. PSO is initialized as a group of random particles and then finds optimal solution through iteration. In every iteration, particle updates itself through tracking two extreme values, among which one is the best solution particle itself finds till the current moment, called individual best value; the other is the best solution the entire group find till the current moment, which is global best value[10-12].

Suppose that there are m particles forming a group in D-dimensional target search space, among which the i th particle is expressed as a D-dimensional vector, $\vec{x}_i = (x_{i1}, x_{i2},......x_{id})$, i.e. the location of the i th particle in D-dimensional search space being the location of each particle is a potential solution. To substitute $\vec{x}_i$ into a target function can get its adaptive value, measuring the strength and weakness of $\vec{x}_i$ according to the size of adaptive value. The flying speed of the i th particle is also adimensional vector, recorded as $\vec{v} = (v_{i1}, x_{i2},......v_{id})$, recording the optimal location searched till now of the i th particle as $\vec{p}_{id} = (p_{i1}, p_{i2},......p_{id})$, and the optimal location searched till now of the entire particle swarm as $\vec{p}_{gd} = (p_{g1}, p_{g2},......p_{gd})$. Traditional PSO carries out the operation on particles according to formula 1 and formula 2.

$$v_{id}^{t+1} = \varpi v_{id}^{t} + c_1 r_1 (p_{id} - x_{id}^{t}) + c_2 r_2 (p_{gd} - x_{id}^{t}) \tag{1}$$

$$x_{id}^{t+1} = x_{id}^{t} + \alpha v_{id} \tag{2}$$

Among which $i = 1,2,3.....m, d = 1,2,3.....D$, learning factors c_1 and c_2 are nonnegative constants; r_1 and r_2 are random numbers, obeying to the uniform distribution of [0, 1]. In the formula, ϖ is nonnegative number, called momentum coefficient, controlling the influence of previous speed on current speed; the greater ϖ is, the bigger the influence of previous speed is, the stronger the global searching ability is; the smaller ϖ is, the less the influence of previous speed is, the stronger the global searching ability is, the local minimum able to be jumped out through adjusting the size of ϖ. α is called constraints factor, aiming to control the weight of speed. Condition for terminating iteration is, according to specific problems, generally selected as maximum iterations or optimal location particle swarm searched up till now meeting desired minimum adaptive threshold.

4.3 The First Improvement of Particle Swarm Algorithm

Two shortcomings of basic particle swarm optimization: one is that PSO is slow in convergence rate; the other is that PSO is easy to fall into local extreme point, causing that the algorithm cannot obtain global optimal solution.

PSO based on growth stage of particle mainly improves it on the speed update formula of particle. The core concept of algorithm improvement is that in the early phase of iteration of particle swarm algorithm, while particle is falling into local extreme point, adopt mutation mechanism to make it jump out of local extreme value and dynamically adjust the proportion of "individual cognition part ($pbest$)" and "social cognition part ($gbest$)" of particle in different evolution phases, thus improving optimal performance of algorithm. As the setting of inertia weight ϖ influences the balance between global searching ability and local searching ability of particle, it plays a critical role in the success of algorithm. Algorithm adopting larger inertial weight in the early phase can make particle swarm have strong global searching ability; adopting smaller inertial weight in the later phase can improve the local searching ability of particle swarm. Therefore, linear adjustment of inertial weight is adopted in this thesis.

Through the learning of parameters of PSO and large quantity of simulation experiment, parameters ϖ , c_1 and c_2 are expressed with mathematical expression as shown in formula (3), formula (4) and formula (5).

$$\varpi = \varpi_{\max} - (\varpi_{\max} - \varpi_{\min}) t / T \tag{3}$$

$$c_1 = \begin{cases} 1.4962 & 0 \prec t \prec 2T/3 \\ 1.4962 - t/T & 2T/3 \le t \le T \end{cases} \tag{4}$$

$$c_2 = \begin{cases} t/T & 0 \prec t \prec T/3 \\ 0.4926 & T/3 \le t \prec 2T/3 \\ 1.4926 + t/T & 2T/3 \le t \le T \end{cases} \tag{5}$$

In the above formulas, t is the current iterations of particles, and T is the maximum iterations of particle swarm. The calculation steps of PSO1 are as follows. Inputs are particle swarm scale N and particle dimension D, inertial weight parameters of particle swarm $\varpi_{\max}$, $\varpi_{\min}$ and target function $fitness(x)$; outputs are optimal location of particle swarm $gbest$; optimal solution $fitness(gbest)$.

Initialize particle location x and particle speed v ;

$t = 1$

Calculate the parameters of iterative formula through formula (3), formula (4) and formula (5);

If $gbest$ fails to be updated or is subtle in change in recent iterations and particle hasn't mutated in recent iterations, carry out mutation on the particle

(5) Update particle speed and location

(6) According to fitness function $fitness(x)$

(7) If the current location of particle is better than the optimal location experienced, $pbest$=current location of particle

(8) Return to the second step until $t = T$;

4.4 The Second Improvement of Particle Swarm Algorithm

Improvement 2 is particle swarm algorithm based on double mutation, the core concept of which is that on the one hand, adopt non-linear decreasing strategy to dynamically adjust inertial weight to control the behavior of particle, so as to make it better balance the development and detection capability of particle; on the other hand, while particle falling into local extreme point, adopt self-adaptive mutation mechanism to make it jump out of local extreme value, and while particle exceeding hunting zone, mutate it to realize double mutation mechanism, so as to improve the optimal performance of algorithm. First, inertial weight is a very important parameter in PSO to control the development and detection capability of algorithm. The size of inertial weight decides the inheritance of particle on current speed. Larger inertial weight will make particle possess larger speed, so as to have stronger searching ability; smaller inertial weight will make particle possess strong development capability. There are generally constant and change (dynamic) in terms of choosing inertial weight. Shi and Eberhart put forward a new method for modifying inertial weight, i.e. inertial weight linearly decreasing with the increasing of algorithm iterations, as shown in formula (6). This thesis puts forward a non-linearly decreasing inertial weight function, mathematical expression as shown in formula (7).

$$\varpi = \varpi_{\max} - \frac{t(\varpi_{\max} - \varpi_{\min})}{T} \tag{6}$$

$$\varpi = e^{-20(t/T)^5} \tag{7}$$

In the above formulas, t is the current iterations of particle, and T is the maximum iterations of particle swarm. Second, the key of this algorithm is to guide particle to effectively mutate while particle falling into local extreme point. The conditions for particle mutation are set as: first, particle should be in the first half phase of iteration; second, particle to be mutated is bad in adaptive value; third, individual extreme value of particle fails to be update or is subtle in change in recent iterations. Randomly select particle meeting the above conditions to mutate; besides, not all of the iterations have to carry out judgment on conditions for particle mutation, spacing several iterations instead. In order to determine the conditions for particle mutation, this thesis brings in such two concepts as change rate of individual extreme value k and contribution rate of individual extreme value p. The mathematical expressions of k_i and p_{ip} are shown in formula (8) and formula (9).

$$k_i = \frac{|p_i(t) - p_i(t - t_{\mathrm{int}v})|}{p_i(t - t_{\mathrm{int}v})} \tag{8}$$

$$p_{ip} = \frac{1}{N}\sum_{i=1}^{i=N} p_i(t) / p_i(t) \tag{9}$$

In the above formulas, $p_i(t)$ means the individual extreme value of the i th particle in the t th iteration, and t is the current iterations,

$\mathrm{int}\,v$ is interval iterations, N is the number of particles. The calculation algorithm of judgment on conditions for particle mutation is as follows.

Inputs are information matrix of current particle swarm I, current iterations of particle t, and interval times of particle mutation $\mathrm{int}\,v$, controlling parameter of particle mutation ∂ and θ; output is the mutation information after t times of iterations of particle.

Input particle swarm information after $t-1$th time iteration;

If $\mathrm{mod}(t, t_{\mathrm{int}v}) = 0$, and $t \le T/2$, then calculate the change rate of individual extreme value of particle k_i and contribution rate of individual extreme value p_{ip};

If $k_i \le \partial$, $p_{ip} \prec 1$, $rand() \prec \theta$, then reset the location and speed of the particle.

If $x \ge x_{\max}$, then $x = x_{\max}(1 - 0.01 * rand())$;

If $x \le x_{\min}$, then $x = x_{\min}(1 - 0.01 * rand())$,

Return to the fourth step;

According to fitness function $fitness(x)$

If the current location of particle is better than the optimal location experienced, $pbest$=current location of particle

Return to the second step until $t = T$;

In the above optimization algorithm, parameter $\mathrm{int}\,v$ is able to indirectly control the mutation rate of particle; ∂ is able to assist judging whether particle is falling into local extreme value; θ is random mutation rate, able to directly control the mutation rate of particle; r is random mutation radius controlled variable, which is also set for directly controlling the mutation rate of particle. The mutation rate of particle can be controlled through adjusting such four parameters as $\mathrm{int}\,v$, t, ∂, θ and r.

4.5 Application of PSO and K_mans in Customer Classification

As there are two inherent defects of K-means clustering algorithm: one is that the random selecting of initial value may cause different clustering results, even no solution; the other is that the algorithm is the one based on gradient descent, thus constantly falling into local optimal solution inevitably. These two defects restrain its application scope. Thus, this thesis will make use of strong parallel computation capability of particle swarm, combing with K-means clustering algorithm, so as to conquer the defects of K-means algorithm.

In particle swarm algorithm, the location of each particle is comprised of K clustering centers, particles having speed and adaptive values besides location. As the dimension of sample vector is D, the location and speed of particle are K*D dimensional vector. Besides, particle has an adaptive value. The coding structure of particle is shown in formula 10.

$$\begin{array}{ll} Z_{11}, Z_{12}, Z_{13} \ldots\ldots Z_{1D} & V_{11}, V_{12}, V_{13} \ldots\ldots V_{1D} \\ Z_{21}, Z_{22}, Z_{23} \ldots\ldots Z_{2D} & V_{21}, V_{22}, V_{23} \ldots\ldots V_{2D} \\ Z_{31}, Z_{32}, Z_{33} \ldots\ldots Z_{3D} & V_{31}, V_{32}, V_{33} \ldots\ldots V_{3D} \\ Z_{41}, Z_{42}, Z_{43} \ldots\ldots Z_{4D} & V_{41}, V_{42}, V_{43} \ldots\ldots V_{4D} \end{array} \tag{10}$$

In PSO, fitness function is used to judge the strength and weakness of particle location in the evolution progress of swarm. Clustering results depend on the results of target function $J = \sum_{i=1}^{k} \sum_{p \in c_i} d(p - Z_i)$.

The smaller the J is, the better the clustering result is. Hence, fitness function of particle adopted in this thesis is shown in formula (11).

$$fitness(Z_i) = \frac{1}{1+J} = \frac{1}{!+\sum_{i=1}^{k} \sum_{p \in c_i} d(p - Z_i)} \tag{11}$$

In formula 11, J means intra-class distance sum; algorithm search the maximum value of fitness function $fitness(Z_i)$ through the iteration of particle swarm.

The steps of K-means clustering algorithm based on PSO are as follows. Inputs are Data Set, category number K and particle swarm algorithm parameter; output is optimal location coding $gbest$.

One-off dispatching each sample as certain category;

Calculate each clustering center Z_k;

Initialize particle swarm with clustering center Z_k;

Carry out clustering division on each sample according to Nearest Neighbor Principle;

If $x \le x_{\min}$, then $x = x_{\min}(1 - 0.01 * rand())$,

As for each particle, calculate new clustering center, update fitness value of particle, replace original coded value.

5. Experimental Verification

5.1 Object of Experimental Verification

The instance data of the experiment conduct empirical research on the customer data of the B2C transaction of certain enterprise website of the recent three years (totaling data of 41351 customers, 21 attributes in the data table are listed in the third part of the paper including customer characteristics type variables and customer behaviors type variables), making statistics on attribute values like annual transaction frequency, total amount, product cost, etc. of certain customer according to customer transaction records in information base, forming an information table (among which the decision attribute set D is null) [4].

5.2 Process of Experimental Verification

The process of the experimental verification can be listed as follows[5].

First, what is to be processed during the classification is the numeric data, so the numeric coding on character data should be conducted first;

Second, if the value number of certain attribute is equal to sample number, it means that it has little effect on classification, hence, remove such attribute first. Three attributes as Customer No., Post Code and Date of Birth are removed in this case.

Third, establish training sample set according to domain (prior) knowledge. Times of purchasing and total amount of purchasing of each customer are two major factors of customer classification (this is the prior knowledge of domain), so select 400 pieces of typical data among all the customers to form training sample set. And divide them into five types as Gold Customers, Silver Customers, Copper Customers, General Customers and Negligible Customers according to ABC management theory.

Fourth, use the customer classification algorithm above-mentioned, and the customer classification results can be expressed in Table 1. In the specific algorithm realization, this Paper simultaneously realizes ordinary K_means algorithm and customer classification algorithm based on BP neural network. The performance comparison of these three algorithms can be expressed in Table 1.

Table 1 Customer Classification Result of Some Website

Customer Type	Number of Customers	Percentage %	Profit Contribution Proportion
Gold Customers	2849	6.89	52.1
Silver Customers	5921	14.32	30.1
Copper Customers	10193	23.65	13.1
General Customers	13751	36.44	6.1
Negligible Customers	7319	17.70	-1.4
Total	41351	100.00	100.00

We can see from Table 2 that in the autonomous learning of algorithm of this Paper, such five factors as the educational background, income, occupation, times of purchasing, and total amount of purchasing of customers have a relatively great influence on customer classification. Through the classification result in Table 1, it can be seen that Gold Customers take up 6.89% of the total number of customers, while the profit takes up 52.1% of the total profit. These customers play a significant role in the existence and development of enterprises. However, the negligible customers account for 17.7%, who not only do not bring profit to enterprises, but also make enterprise lose money. These customers should be either further cultivated or eliminated according to the actual situation.

Table 2 Classification Performance Comparison of Each Algorithm

Algorithm	Algorithm in This Paper	Ordinary K-means Algorithm	BP Neural Network Algorithm
Accuracy Rate	99.7 %	88.47%	94%
E Value	104.33	159.81	119.96

We can see from Table 2 that the cluster accuracy rate of algorithm in this paper is the highest, reaching 99.7 %, obviously higher than ordinary K-means algorithm and BP Neural Network algorithm; the square errors and E values on customer classification of three algorithms are 104.33, 159.81 and 119.96 respectively. The smaller the E value is, the smaller the possibility of wrong classification is. Thus it can be seen that the square error and E value of the algorithm in this paper during the classification are far more less than ordinary K-means algorithm[4] and BP Neural Network algorithm[6]. Therefore, it shows that the improvement on K-means clustering algorithm in this paper turns out to be a success, with reasonable classification results.

6. Conclusion

Customer relations management of online trading is still developing. But to correctly and effectively classify online trading customers is the critical issue for reforming network marketing mode, improving customer management and service level and enhancing competitiveness of network enterprises[5]. On account of the shortcomings of the typical K_means and PSO classification algorithm in data mining, this paper puts forward several improvement measures, and applies them into the classification of online trading customers. Simulation results indicate that the improved online trading customer classification has higher accuracy rate on customer classification and more reasonable classification results.

References

[1] Liu Zhaohua:Study on Model of Customer Classification Based on the Customer Value ,A Dissertation of Huazhong University of Science and Technology,2018.
[2] Zhou Huan:Study of Classifying Customers Method in CRM,Computer Engineering and Design, Vol 29(2018), No.3, p.659-661.
[3] Deng Weibing:Wang Yan:B2C Customer Classification Algorithm Based on Based on 3DM,Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition),Vol 21(2017) No.4, p.568-572.
[4] Guan Yunhong:Application of Improved K-means Algorithm in Telecom Customer Segmentation, Computer Simulation, Vol 28,(2016)No.8, p.138-140.
[5] Qu Xiaoning:Application of K-means Based on Commercial Bank Customer Subdivision, Computer Simulation, 2016, Vol 28(2106) No.6, p.357-360.
[6] Yang Benzhao:Tian Gen, Research on Customer Value Classification based on BP Neural Network Algorithm, Science and Technology Management Research,Vol 23(2017) No.12, p.168-170.
[7] Bradley P S,Managasarian L.:*k-plane Clustering*.Journal of Global Optimization, (2009)16(1), p.23-32.
[8] Tang Yong:Rong Qiusheng. An Implementation of Clustering Algorithm Based on K-means.Journal of Hubei Institute for Nationalities,Vol.22 (2014)No.1, p.69-71.
[9] Zhang Y.F.,Mao J. L.:An improved K-means Algorithm, Computer Application, vol.23(2017) No.8, p. 31-33.
[10] Wang Zhigang:A modified particle swarm optimization, Journal of Harbin University of Commerce (Natural Sciences Edition), vol.25(2018)No.4, p. 454-456.
[11] Wang Hongli, Hou Qingjian:Improved Particle Swarm Optimization Algorithm and Its Simulation, Process Automation instrumentation, vol.30(2009)No.17, p. 28-30.
[12] Li Jing, Wu Cheng:Particle Swarm Optimization Algorithm Based on Grouping and Cooperating, Science Technology and Engineering, vol.9(2009)No.16, p. 4806-4809.