

# Risk Analysis and Fairness Evaluation of Insurance Consumption Behavior Based on Big Data

Mei Du<sup>1</sup>, Yanliang Yu<sup>2</sup>, Xu Chen<sup>3,\*</sup>

<sup>1</sup>Library, Jilin Engineering Normal University, Changchun, China

<sup>2</sup>City College of Dongguan University of Technology, Dongguan, China

<sup>3</sup>Business and Management Department, Jilin Engineering Normal University, Changchun, China

\*.Correspondence Author

**Keywords:** Risk Analysis; Evaluation; Insurance; Consumption

**Abstract:** Introducing cluster analysis into customer segmentation can effectively process a large amount of customer information. This paper analyzes various methods of cluster analysis and selects ant colony clustering algorithm for customer segmentation. Before the implementation of the algorithm, the method of principal component analysis is selected to preprocess the original data. Finally, the customer segmentation model of real estate enterprises based on clustering analysis is constructed. The feasibility and practicability of the real customer information inspection model of Baoyu Real Estate are verified. Based on the actual customer data obtained from the questionnaires of Baoyu Real Estate customers, the customer segmentation model constructed in this paper is adopted to cluster the collected customer information, and the clustering results are compared with the original customer data for analysis, and the company's resource allocation strategy and customer management strategy are formulated based on the basis of the results.

## 1. Introduction

Compared with traditional customer segmentation, the customer segmentation of this technology can discover the potential favorable information in the customer relationship. However, among the numerous data mining and analysis technologies, clustering analysis technology based on partition is widely used in the field of customer segmentation, and its advantages are as follows: its processing time is fast; its operation is simple; and it is easy to understand. Therefore, it is one of the widely used data mining algorithms, and k-means clustering algorithm is the most typical representative. At present, clustering analysis technology has been used in the field of dealing with remote sensing imaging, pattern recognition, statistics, biology, financial business, geography and so on, but it is rarely used in the field of real estate. Based on the real estate enterprises which adopt clustering algorithm to segment customers, different characteristics of customers are obtained in this paper, so as to assist the real estate enterprises to formulate a production and service strategy of corresponding buyers, carry out a targeted management of house purchasers, excavating the purchasers' potential, improve their contribution, and rationally utilize the resources of the real estate enterprises, thus maximizing the profits of enterprises.

## 2. Ant colony clustering algorithm

The host node distributes ANT - NUM ants equally to each slave node, and broadcasts the distance matrix to each slave node. In each iteration, the host node broadcasts the current global pheromone matrix, and each slave node uses the pheromone matrix to construct the local optimal solution, and then the host node collects the local optimal solution of each slave node, based on which the global pheromone matrix is modified. This iterates many times until the satisfactory result is obtained. The parallel strategy of ant colony algorithm is shown in Figure 1.

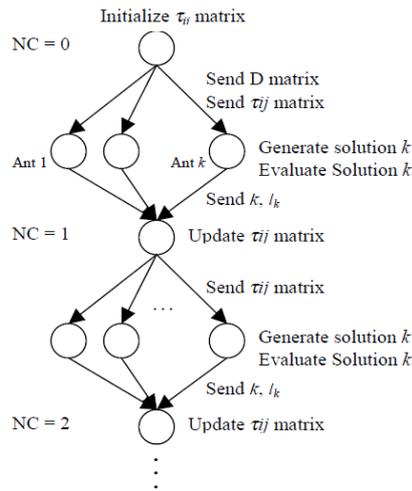


Figure 1 Parallel strategy of ant colony algorithm

Description of parallel algorithm: (1) According to the number of nodes involved in the calculation, the host node distributes ANT - NUM ants to each slave node equally; (2) the host node calculates the distance matrix, visibility matrix and initial pheromone matrix between cities, and broadcasts the distance matrix and visibility matrix to each slave node; (3) the distance matrix and visibility matrix received from the slave node are stored locally; (4) the host node broadcasts pheromone matrix; (5) the pheromone matrix received from the broadcast nodes and then it is stored locally; (6) According to the local distance matrix and pheromone matrix, each slave node constructs a local optimal solution by circulating N ants once; (7) the host node collects the local optimal solutions calculated by each slave node, compares and obtains the global optimal solution of this cycle, and updates the pheromone matrix according to the pheromone updating rules; (8) if the end condition is met, it will exit; Otherwise, it will turn to (4) for the next iteration. The algorithm flow chart is as follows:

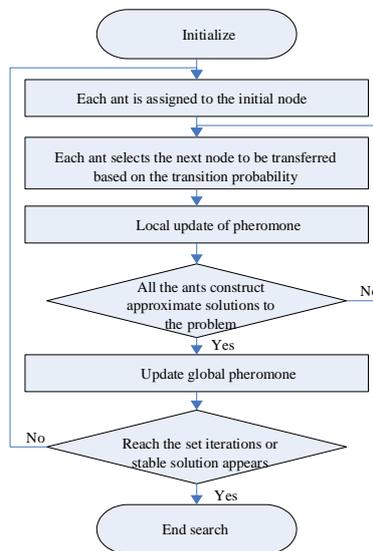


Figure 2 Flow chart of ant colony parallel algorithm

The algorithm execution diagram is as follows:

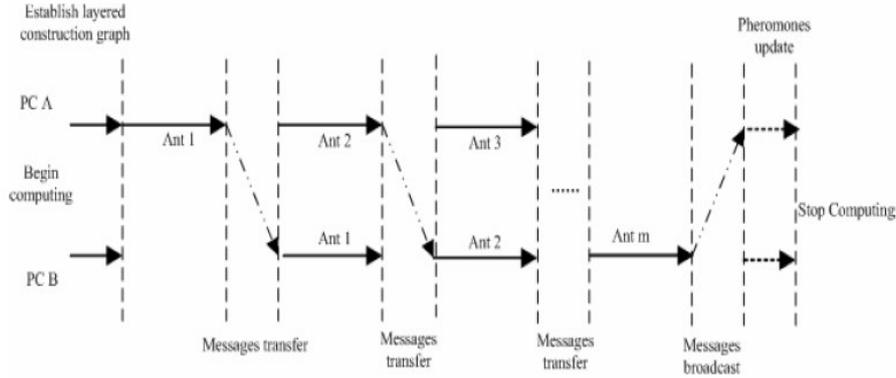


Figure 3 Execution diagram of ant colony parallel algorithm

### 3. Empirical analysis

In this paper, the questionnaire star is used among the customers of Baoyu Real Estate to provide technical support. The the data in the collected effective questionnaires is carried out by a descriptive statistical analysis, and the information of 40 customers is extracted as the original data. Meanwhile, combining with the selected segment index, a 40 x 14 data matrix is generated, and the property fields in the matrix are shown in Table 1.

Table 1 Description of property fields

Fields	Description	Fields	Description
C11	Purchase amount (ten thousand Yuan)	C42	Repayment timeliness
C12	Purchase amount (ten thousand Yuan)	C34	Times of recommending customers
C21	Gross profit (ten thousand Yuan)	D11	Customer elasticity
C22	Relationship cost	D12	Purchase frequency
C31	Customer development potential	D13	Purchase shares (%)
C33	Transaction purchase	D21	Customer satisfaction
C41	Bank credit	D23	Conversion cost

The results are the iterative history table of ant colony clustering algorithm analysis. It can be seen that a total of four iterations are carried out. The distance between the class center formed after the first iteration and the initial class center is 0.00, 1.642, 1.685, 2.522, 1.555, 1.750, 1.522, 1.780, respectively. After the second iteration, the distance is 0.00, 0.00, 0.223, 0.00, 0.271, 0.00, 0.00, 0.333, respectively. After the third iteration, the distance is 0.00, 0.00, 0.00, 0.00, 0.193, 0.00, 0.00, 0.188, respectively. After the fourth iteration, the class center point doesn't change or changes very little, with a distance of 0.00, as shown in Table 2.

Table 2 Iteration history

	1	2	3	4	5	6	7	8
1	0.000	1.642	1.685	2.522	1.555	1.750	1.522	1.780
2	0.000	0.000	0.223	0.000	0.271	0.000	0.000	0.333
3	0.000	0.000	0.000	0.000	0.193	0.000	0.000	0.188
4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

### 5. Conclusion

this paper takes the actual customer data of Baoyu Tianyi of Harbin as an example, and extracts

customer samples for segmentation through customer segmentation theory and cluster analysis, firstly, the collected sample data is preprocessed, the principal component is extracted, and then cluster analysis is carried out on the variables, the customer segmentation model is built and suggestions are put forward for Baoyu Real Estate Company on the basis of the cluster analysis, hoping to use the segmentation results to help deal with practical problems, retain customers and reasonably allocate the resources of enterprise. As a result, this paper classifies the customers of real estate enterprises with cluster analysis technology, and establishes a customer segmentation index system from two perspectives of customer value and customer loyalty, and adheres to the concept of 'customer first' throughout the whole development, sales and later services process, so as to improve the efficiency of real estate enterprises with the customer demand as the starting point and foothold of management decisions and plans.

### **Acknowledgement**

Foundation Project for Key Disciplines of Insurance Characteristic in City College of Dongguan University of Technology.

### **References**

- [1] Hernandez P R G, Carvalho L F, Proenca M L. Digital Signature of Network Segment Using Flow Analysis through Genetic Algorithm and ACO Metaheuristics[C]// Chilean Computer Science Society. 2016.
- [2] Boubaker J, Mouna G, Sahnoun M. Metaheuristic Techniques to Optimize the Steaming Process of Elastic Denim Yarns[J]. Journal of Natural Fibers, 2017, 14(6):1-9.
- [3] Singh S, Chana I. Cloud resource provisioning: survey, status and future research directions[J]. Knowledge & Information Systems, 2016, 49(3):1005-1069.
- [4] Sundareswaran K, Vigneshkumar V, Sankar P, et al. Development of an Improved P&O Algorithm Assisted Through a Colony of Foraging Ants for MPPT in PV System[J]. IEEE Transactions on Industrial Informatics, 2016, 12(1):187-200.
- [5] Armaghani D J, Mohamad E T, Hajihassani M, et al. Evaluation and prediction of flyrock resulting from blasting operations using empirical and computational methods[J]. Engineering with Computers, 2016, 32(1):109-121.