

Two-phase Optimization Approach for Urban Express Vehicle Routing Problem with Customer Satisfaction Consideration

Mengyuan Zhang¹, Mengke Yang^{1,a,*} and Xiaoguang Zhou^{1,b}

¹*School of Automation, Beijing University of Posts and Telecommunications, Beijing, China*
a. yangmengke@139.com, b. zyg@bupt.edu.cn

**corresponding author: Mengke Yang*

Keywords: Vehicle routing problem, customer grouping, customer satisfaction.

Abstract: With the development of economy, customers have higher demands for express delivery services. How to plan a lower cost vehicle routing under the premise of meeting customer needs has become a focus of express delivery companies. Based on the traditional vehicle routing problem research, We design a two-phase optimization method from the perspective of customer needs. The first phase model use K-means clustering method to group customers reasonably after selecting grouping indicators. The second phase model is a vehicle route optimization model which can minimize the cost and improve customer satisfaction, and then solving with genetic algorithm. On this basis, it can not only improve the quality of distribution service, but also realize the effective control of distribution cost.

1. Introduction

In recent years, with the rapid development of national economy, e-commerce has gradually become a new highlight of economic growth, promoting the sustainable development of express industry. With the continuous development of commodity economy, customers put forward higher requirements for the service quality of express enterprises. In addition to time requirements, service quality has also become an important consideration for current customers. Therefore, how to improve the quality of distribution service has become the focus of express delivery enterprises. By reasonably arranging the vehicles routing, providing appropriate service for each customer, and delivering express to customers on time, it can not only improve customer satisfaction, win customer recognition, but also improve customer loyalty to the enterprise, so as to improve the market competitiveness of the enterprise. Therefore, fully considering the time, cost and service quality as the important indicators of vehicles routing is conducive to improving the efficiency of both express enterprises and customers.

Therefore, this paper proposes a two-phase optimization method for express vehicle routing problem, which considers customer demand and distribution cost. The first phase model use K-means clustering method to group customers reasonably after selecting grouping indicators. The second phase model is a vehicle route optimization model which can minimize the cost and improve

customer satisfaction, and then solving with genetic algorithm. On this basis, it can not only improve the quality of distribution service, but also realize the effective control of distribution cost.

2. Literature Review

VRP was proposed by Dantzig and Ramser at the end of the 20th century. Brandao [1] studied multi-type VRP, solved the problem with tabu search algorithm (TS), and analyzed the influence of vehicle type on the solution effect. Calvete [2] established a multi-objective optimization model (VRPSTW) of vehicle routing problem considering soft time window, and designed an objective planning method to solve the problem. In order to solve the VRPTW problem, Miguel [3] proposed an iterative path construction and improvement algorithm (RICI) based on path generation algorithm. Zhe liang proposed a two-phase method to solve the vehicle routing problems. In the first phase, a new heuristic algorithm is used to get the initial solution, and in the second stage, a tabu algorithm is used to improve the solution. Wee kit [9] focused on the hybrid heuristic algorithm of tabu algorithm and genetic algorithm, and then combined two advantages to design the optimal solution. Jorg and Hermann [4] used two evolutionary strategies to solve vehicle routing problems, and made full use of IOTA heuristic technology.

Bin Song [8] used genetic algorithm to solve VRPSTW. On the basis of the traditional genetic algorithm, Zhang Liangzhi et al.[5] added time constraint algorithm, aiming at the initialization population, according to the location of the customer point, let the mutation rate rise and reduced the inefficient calculation. Chen Min [6] combined with the research status, to explore the real-time vehicle routing problem with time windows, and combined with the various factors of the problem to carry out research, and build a new hybrid algorithm to optimize the solution. Wang Yongfeng et al.[7] analyzed the characteristics of VRPTW, combined with the use of hybrid genetic algorithm to optimize the solution, and according to the experimental results, the algorithm has a good effect, mainly for the high efficiency of the algorithm.

Based on the above analysis, although there are many researches on vehicle routing problems, there are few researches on customer demands. Therefore, in this paper, we select five indicators that affect customer grouping, and establish similarity matrix to cluster customers with similar product needs. On the basis of clustering results, we establish a vehicle routing optimization model, and design a genetic algorithm to solve the model. The purpose of this study is to make a theoretical contribution to VRP considering customer needs.

3. Model

3.1. Customer Grouping Model

we select five indicators that affect customer grouping, such as customer location, express service quality, express volume, express value and service security requirements, and establish similarity matrix to cluster customers with similar product needs.

3.1.1. Select Customer Grouping Indicators

In the city express delivery, the customers locations and express demands are the key that affect the quality of delivery. In the past, most of the researches on express vehicle routing only consider the location factor, and group customers according to the space distance between customers and distribution center. When there is a large gap in customer demand, it may lead to the imbalance of distribution volume in each distribution center. Therefore, in this paper, the geographical location of customers and the quality of express service are taken into account when clustering customers,

which can effectively improve the quality of express delivery service, and has a certain practical significance for urban express vehicle route planning. Table 1 defines the selected customer clustering indicators.

Table 1: Customer clustering indicator.

First-level indicator	Second-level indicator	Indicator description	Value definition
Customer properties	geographical position	It refers to the actual geographical location between two customers and between customer and distribution center.	Longitude and latitude of customer point
	Service quality requirements	Including customer requirements for delivery time, courier service attitude and service ability.	The value range is (0,1). The higher requirements for the quality of delivery service, the larger this value
Express properties	Express type	According to the size, it can be divided into four types: document, small express, parcel and bulk express.	Document, small express, parcel and bulk express, corresponding values (0.1,0.4,0.7,1) respectively
	Express value	It refers to the value of express delivery goods. For high-value express delivery, it can be delivered separately to provide better delivery services.	The value range is (0,1), the higher the value of express, the larger the value
	Express security requirements	It refers to the special requirements of customers for express safety, such as fragile products. The customers who have safety requirements are divided into a group for distribution.	If the express needs to be delivered separately, the value is 1, otherwise it is 0.

3.1.2. Data Standardization Processing

Most of the cluster statistics, especially the distance statistics, are usually influenced by the measurement units of the cluster variables. In this paper, the selected indicator data dimensions are different, and the order of magnitude units are also quite different. Therefore, the data needs to be transformed. Here, the commonly used standardized transformation method is used to preprocess the data of each type of indicator:

$$x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{S_j}, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, p \quad (1)$$

Where the mean value of samples is

$$\bar{x}_j = \frac{1}{n} \sum_{a=1}^n x_{aj} \quad (2)$$

The standard deviation of the sample is

$$S_j = \sqrt{\frac{1}{n-1} \sum_{a=1}^n (x_{aj} - \bar{x}_j)^2} \quad (3)$$

After transformation, the sample mean value is 0, and the sample standard deviation is 1. The data processed is unrelated with the dimension of variables.

3.1.3. Calculate Similarity Matrix

In clustering analysis, it is necessary to calculate the similarity between individuals, which is defined as similarity measurement. In this paper, we use cosine similarity method to calculate individual similarity. The matrix of standardized data is set as C, using cosine similarity method to generate customer similarity matrix W.

$$w_{ij} = \begin{cases} 1 & i = j \\ \frac{\sum_{k=1}^n (x_i^k \times x_j^k)}{\sqrt{\sum_{k=1}^n (x_i^k)^2} \times \sqrt{\sum_{k=1}^n (x_j^k)^2}} & i \neq j \end{cases} \quad (4)$$

Among them, w_{ij} represents the elements in the similarity matrix W, x_i^k is the element K of line I of matrix C, x_j^k is element K of row J of matrix C, W is a $n \times n$ matrix, n represents the number of customers.

3.2. Express Vehicle Routing Model based on Customer Satisfaction Function

In this paper, the problem of express vehicle routing based on customer satisfaction function is mainly based on the group vehicle routing after customer grouping. It introduces customer satisfaction model in the planning of express vehicle path, considering the requirements of customers for the timeliness and integrity of express delivery, and constructs a express vehicle routing model based on customer satisfaction.

3.2.1. Time Satisfaction Function

We design a soft time window, which includes the time window of customer preference and the maximum tolerable time window. Suppose the delivery time window of customer preference is $[s_i, e_i]$, the maximum tolerable time window is $[S_i, E_i]$. Within the time range $[s_i, e_i]$, the customer satisfaction is 100%. If the delivery time is out of range $[S_i, E_i]$, customer satisfaction is 0. If delivery time is within range $[S_i, s_i]$ or $[e_i, E_i]$, customer satisfaction decreases with the gap between delivery time and preference time window. The time satisfaction function is shown as figure1 below.

$$F(t_i) = \begin{cases} \left[\frac{t_i - S_i}{s_i - S_i} \right]^\alpha & S_i \leq t_i < s_i \\ 1 & s_i \leq t_i \leq e_i \\ \left[\frac{t_i - e_i}{E_i - e_i} \right]^\beta & e_i < t_i \leq E_i \end{cases} \quad (5)$$

t_i is delivery time for customer i , α 、 β is the time sensitive coefficient, which is determined by experience α , $\beta \in (1, 1.5]$.

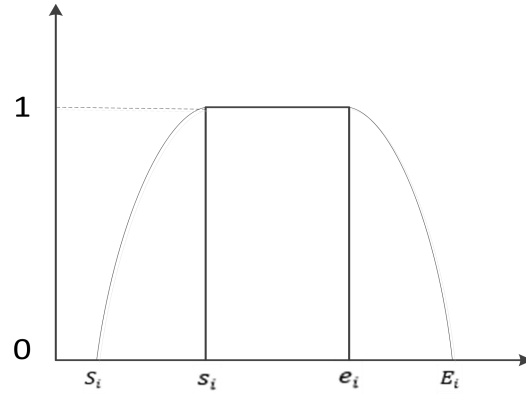


Figure 1: Time satisfaction function.

3.2.2. Goods Intact Satisfaction Function

In this paper, we introduce the concept of damage rate to measure customer satisfaction. The higher the damage rate, the lower the customer satisfaction. The formula for calculating the damage rate is as follows:

$$g_i = \theta_i d_{i0} \quad (6)$$

θ_i is good damage coefficient per kilometer during transportation, d_{i0} is the distance from the customer i to the distribution center.

Assume that the acceptable rate of damage is $[0, h]$. In this range, customer satisfaction will decrease with the increase of goods damage rate. When the damage rate exceeds this range, the customer does not accept this distribution service, and the customer satisfaction is 0. The satisfaction function of goods in good condition is shown in the figure below.

$$G(g_i) = \begin{cases} \left[\frac{h - g_i}{h} \right]^\gamma & 0 \leq g_i < h \\ 0 & g_i \geq h \end{cases} \quad (7)$$

γ is the sensitivity factor of the customer to the rate of goods damage.

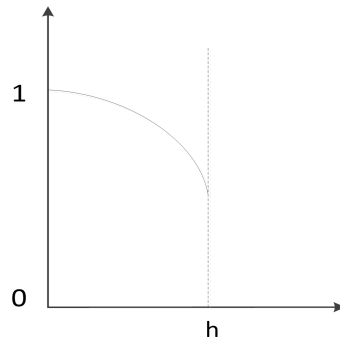


Figure 2: Goods intact satisfaction function.

3.2.3. Vehicle Routing Model

The symbols used in the model are described as Table 2.

Table 2: The symbols used in the model.

Symbol	Explain
V	A collection of customers and distribution center, $V = \{V_0, V_1, V_2, \dots, V_n\}$. V_0 is the distribution center, $\{V_1, V_2, \dots, V_n\}$ is customer demand points
K	Number of delivery vehicles
Q	Approved load capacity of vehicle
q_i	Customer point demand
C_1	Fixed cost per vehicle
C_2	Transportation cost per unit distance vehicle
d_{ij}	Distance between arcs (i,j)
t_{ij}	Travel time of vehicle in arc (i,j)
v	Average speed of vehicle transportation
y_i	Service time of vehicle at customer i
ω_1, ω_1	The weight of customer satisfaction function
x_{ijk}	$\begin{cases} 1 & \text{Vehicle } k \text{ drives from customer } i \text{ to customer } j \\ 0 & \text{otherwise} \end{cases}$
$F_i(t_i)$	Time satisfaction function of customer i
$G_i(g_i)$	Goods intact satisfaction function of customer i
dt_i	Departure time of vehicle from customer i
et_i	Time of vehicle arriving at customer i
S_i	Lower limit of the maximum tolerance time window of customer i
E_i	Upper limit of maximum tolerance time window for customer i

According to the express vehicle routing problem based on customer demand grouping, the following assumptions are made: (1) Distribution from one distribution center to multiple customers;

(2) Distributed by single model vehicles with fixed vehicle approved load; (3) Each customer is delivered only once by a single vehicle, and the sum of customer deliveries on the same route does not exceed the vehicle approved load; (4) All vehicles start from the distribution center and return to the distribution center after the completion of the distribution service, and all customers will be served; (5) The demand, required delivery time and geographical location of each customer point are known; (6) The driving speed of vehicles in the distribution process is affected by traffic congestion, and the speed is variable. So we establish the following vehicle routing model.

$$\min Z_1 = C_1 \sum_{j=1}^N \sum_{k=1}^K x_{0jk} + \sum_{i=0}^N \sum_{j=0}^N \sum_{\substack{k=1 \\ j \neq i}}^K C_2 d_{ij} x_{ijk} \quad (8)$$

$$\min Z_2 = \omega_1 \left(1 - \frac{1}{N} \sum_{i=1}^N F_i(t_i) \right) + \omega_2 \left(1 - \frac{1}{N} \sum_{i=1}^N G_i(g_i) \right) \quad (9)$$

$$\sum_{j=1}^N \sum_{k=1}^K x_{0jk} \leq K \quad (10)$$

$$\sum_{k=1}^K \sum_{\substack{i=0 \\ i \neq j}}^N x_{ijk} = 1, j = 1, 2, \dots, N \quad (11)$$

$$\sum_{k=1}^K \sum_{\substack{j=0 \\ i \neq j}}^N x_{ijk} = 1, i = 1, 2, \dots, N \quad (12)$$

$$\sum_{i=0}^N x_{iak} - \sum_{j=0}^N x_{ajk} = 0, a = 1, 2, \dots, N, k = 1, 2, \dots, K \quad (13)$$

$$\sum_{j=1}^N x_{0jk} = \sum_{j=1}^N x_{j0k} \leq 1, j = 1, 2, \dots, N \quad (14)$$

$$\sum_{i=0}^N \sum_{j=0}^N q_i x_{ijk} \leq Q, i \neq j, k = 1, 2, \dots, K \quad (15)$$

$$d_{ij} = t_{ij} \times v, i \neq j, i, j = 1, 2, \dots, N \quad (16)$$

$$S_j \leq et_j \leq E_j, j = 1, 2, \dots, N \quad (17)$$

$$dt_i + y_i + t_{ij} - H(1 - x_{ijk}) \leq dt_j, i, j = 1, 2, \dots, N, i \neq j \quad (18)$$

(8) (9) is the objective function, (8) represents the minimization of vehicle transportation cost and fixed cost, and (9) represents the maximization of customer satisfaction. (10) - (18) is the

constraint condition, (10) indicates that the total number of vehicles used does not exceed K ; (11) (12) indicates that each customer is only served by one vehicle; (13) indicates that the same vehicle arrives at the customer point and leaves the customer point; (14) indicates that only one vehicle on each route serves all users on this route; (15) indicates that the customer served by one vehicle. The total demand of the customer does not exceed the approved load capacity of the vehicle; (16) calculates the driving distance of the vehicle from customer point I to customer point J ; (17) indicates that the time when the vehicle arrives at the customer should be within the maximum tolerance time window; (18) indicates that the time when the customer arrives at the next customer point is later than the time when the customer arrives at the current customer point, where H is an infinite positive number.

4. Algorithmic

4.1. Clustering Algorithm

In this paper, we use k-means method to cluster, in which we use elbow method to determine the number of clusters. The key index of elbow method is SSE (sum of the squared errors).

$$SSE = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (19)$$

C_i is Cluster i , p is sample point of C_i , m_i is centroid of C_i , SSE is the clustering error of all samples, which represents the clustering effect. The shape of the elbow, and the corresponding K value of the elbow is the real clustering number of the data. If when $k = 4$, there is an obvious inflection point in the relationship, the best clustering number should be 4.

In this paper, we cluster the customers by similarity matrix, and finally divide the customers with high similarity into the same group. The specific algorithm steps are as follows:

Step 1: determine the number of clusters K , input the similarity matrix;

Step 2: randomly generate K objects as the initial clustering center;

Step 3: calculate the distance from each customer point to the cluster center, and divide the customers according to the minimum distance to complete the initial classification;

Step 4: recalculate the mean value of each cluster as the center of the next cluster;

Step 5: cycle step 3 and step 4 until the clustering result no longer changes, then the algorithm stops.

Step 6: output the final clustering results.

4.2. Genetic Algorithm

In this paper, natural number coding is used to construct chromosomes. The number 0 is used to represent distribution center. The number 1, 2, 3, n refers to n customer demand points provided by the distribution center. The vehicle routing problem studied in this paper requires distribution vehicles to start from the distribution center and return to the distribution center after completing the task, so the distribution path must start from 0 and end at 0. The example is 0-1-5-3-2-4-0.

The specific steps of the algorithm are as follows:

Step 1: determine the customer code corresponding to the gene code according to the result of customer demand grouping;

Step 2: construct chromosomes by natural coding, and set the coding of the first and last genes of chromosomes to 0;

Step 3: set the initial population number n , adaptive crossover probability p_c . Mutation probability p_m , maximum number of iterations N ;

Step 4: produce initial population, randomly generate N chromosomes, each chromosome represents a distribution path;

Step 5: calculate the individual fitness, population average fitness, cumulative fitness and other parameters of the new population;

Step 6: use the roulette to select individuals and generate the next generation of population;

Step 7: using partial matching crossover method and basic mutation method to generate a new generation of population;

Step 8: the number of iterations is accumulated continuously. When the algorithm termination condition is reached, the algorithm is terminated. Otherwise, return to step 5 and continue the cycle.

5. Case Study

In order to verify the validity of the model, we select 15 customer point data of Haidian District of Beijing as examples for analysis. The specific data is shown in Table 3.

Table 3: Data of 15 customers in Haidian District, Beijing.

Customer	Customer longitude	Customer dimension	Service quality	Express type	Express value	Demand (ton)	Time window
1	116.364	39.968	0.7	0.7	0.5	0.3	[8:00,10:00]
2	116.372	39.967	0.8	0.4	0.6	0.2	[10:00,12:00]
3	116.358	39.971	0.3	0.7	0.7	0.2	[9:00,10:30]
4	116.362	39.977	0.4	1	0.8	0.1	[14:30,17:00]
5	116.364	39.958	0.8	0.1	0.6	0.5	[9:00,12:00]
6	116.359	39.952	0.6	0.4	0.5	0.2	[14:00,15:00]
7	116.350	39.958	0.7	0.4	0.5	0.3	[15:00,16:00]
8	116.349	39.965	0.5	0.7	0.6	0.1	[8:00,11:00]
9	116.356	39.950	0.9	0.7	0.7	0.1	[13:00,17:00]
10	116.352	39.944	0.7	0.7	0.7	0.1	[8:30,12:00]
11	116.325	39.953	0.3	0.7	0.8	0.2	[10:00,11:00]
12	116.327	39.956	0.8	0.7	0.9	0.6	[14:00,15:00]
13	116.335	39.954	0.8	0.4	0.8	0.2	[11:00,11:30]
14	116.319	39.962	0.9	1	0.3	0.2	[15:00,16:00]
15	116.322	39.967	0.6	0.1	0.7	0.1	[15:00,17:00]

We cluster 15 customers and determine the cluster number is 3 by elbow method. Then we calculate a 15×15 similarity matrix, which is used as input data for k-means clustering. The final result is as Table 4.

Table 4: Clustering results of customer groups.

Group number	Group customer points	Total demand (ton)
Group 1	1、7、8、11、12	1.5
Group 2	2、3、4、10、13、14、15	1.1
Group 3	5、6、9	0.8

The parameters of genetic algorithm are as follows: the initial population number $inn = 100$, maximum number of iterations $gnmax = 100$, adaptive crossover probability $crossprob = 0.8$, mutation probability $muteprob = 0.08$. We run five times, and choose the minimum value of the objective function as the optimal solution, and finally get the following optimal vehicle routing in the group.

Table 5: Optimal route of each customer group.

Group number	Group vehicle routing	Transportation cost	Customer satisfaction
Group 1	0-1-8-11-12-7-0	177.3	92.79%
Group 2	0-3-2-10-13-14-15-4-0	192.8	88.64%
Group 3	0-5-6-9-0	158.6	96.07%

6. Conclusions

In this paper, according to the diversification of customer demand, the characteristics of customer demand are analyzed, and five characteristic indexes are selected to group customers. Then, in the construction of vehicle routing model, we introduce the customer satisfaction function, and use a two-phase heuristic algorithm to solve the problem. Finally, an actual case is used to verify the effectiveness of the model and method. Compared with the traditional vehicle routing problem, this paper considers the impact of customer demand on vehicle routing planning from the perspective of customers, which not only improves customer satisfaction, but also reduces distribution costs. This model can provide a theoretical basis for the optimization of the distribution routing of express enterprises, which is conducive to improving customer satisfaction and market competitiveness of enterprises.

Acknowledgments

This research was funded by the Fundamental Research Funds of Beijing University of Posts and Telecommunications, grant number 2018XKJC07 and National Development and Reform Commission, grant number JM-19-3-4.

References

- [1] B. Golden, A. Assad, L. Levy, and F. Gheysens, "The fleet size and mix vehicle routing problem," *Computers & Operations Research*, vol. 11, no. 1, pp. 49–66, 1984.
- [2] H. I. Calvete, C. Gale, M.-J. Oliveros, and B. Sanchez-Valverde, "A goal programming approach to vehicle routing problems with soft time windows," *European Journal of Operational Research*, vol. 177, no. 3, pp. 1720–1733, 2007.
- [3] F. Miguel, "Emissions minimization vehicle routing problem," *The 89th Annual Meeting of the Transportation Research Board*, pp. 224–231, 2010.
- [4] J. Cordeau, M. Gendreau, and G. Laporte, "A tabu search heuristic for periodic and multi-depot vehicle routing problems," *Networks*, vol. 30, no. 2, pp. 105–119, 1997.
- [5] Zhang Liangzhi, He Minai, Zhang Zhihai. *Optimizing Vehicle Routing Problem with Time Constrains Based on Genetic Algorithm[M]*. *Journal Transportation Systems Engineering and Information Technology*, 2006,6(5): 70–74.
- [6] Chen Min. *Study on Real-time Vehicle Routing Problem with Time Windows in Express Delivery Dispatching[D]*. Shanghai: Shanghai Jiao Tong University, 2009: 35–43.
- [7] Wang Yongfeng, Yang Yu, Gu Yongming, Wu Caiming. *Chaotic Genetic Algorithm for Solving Vehicle Routing Problems with Time Windows[J]*. *Application Research of Computers*, 2012, 29(7): 2423–2426.

- [8] Bin Song, Fu Zhuo. *An Improved Genetic Algorithm for Vehicle Routing Problem with Soft Time Windows*[J].*Systems Engineering*, 2003, 21,6: 79~85.
- [9] Wee-Kit. *A Hybrid Search Algorithm for the Vehicle Routing Problem with Time Windows* *International Journal on Artificial Intelligence Tools*[J]. 2001.Vol.10, No.3: 431-449.
- [10] Ting C.J., Huang C.H. *An improved genetic algorithm for vehicle routing problem with time windows*[J]. *International Journal of Industrial Engineering-theory Applications and Practice*, 2005, (12):218-228.
- [11] G. Y. Tu't'uncu", "An interactive GRAMPS algorithm for the heterogeneous fixed fleet vehicle routing problem with and without backhauls," *European Journal of Operational Research*, vol. 201, no. 2, pp. 593-600, 2010.
- [12] J. Qian and R. Eglese, "Fuel emissions optimization in vehicle routing problems with time-varying speeds," *European Journal of Operational Research*, vol. 248, no. 3, pp. 840-848, 2016.