Journal of Engineering Science and Technology Review 11 (2) (2018) 111 - 120

Research Article

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Two-stage Algorithm for Capacitated Vehicle Routing Problem

Xueyuan Wang[1,*], and Hongyu Zhu[2]

[1] *College of Economy & Management, Neijiang Normal University, Neijiang 641112, China*
[2] *Department of Business and Economics , Linnaeus University, Växjö SE-35195, Sweden*

Received 9 January 2018, Accepted 24 March 2018

Abstract

Distribution is the key link of logistics, and its cost accounts for 30%-70% of logistics expense. The key to control distribution transportation cost is to reasonably plan the vehicle route. With the rapid distribution expansion, the complexity of vehicle routing has increased, and the calculation amount increases by an order of magnitude. The quality and efficiency of the solution can hardly satisfy the practical demand. A three-class and three-tier distribution model was established, and a new two-stage algorithm combined with clustering and ant colony algorithms was proposed to solve the capacitated large-scale vehicle routing problem. In the first stage, the solved data set clustering was divided into multiple clusters using K-means algorithm, which reduced the problem scale. In the second stage, the improved ant colony algorithm was employed to enhance the local optimization and improve the quality of vehicle routing. Finally, the proposed two-stage algorithm was simulated by taking the distribution vehicle routing in the road freight transport planning of Longquan Logistics Center in Sichuan Province as an example and compared with the basic ant colony algorithm. The results reveal the following: (1) the two-stage algorithm can effectively reduce size of the problem the problem and improve the solution quality with low complexity in time and space. (2) The average quality optimization rate of two-stage algorithm is 54.13%. As the solution efficiency increase, the advancement is notable. (3) The correlation between the solution quality and the scale increase of the two-stage algorithm is 0.99. The solution quality is relatively stable and has good robustness. (4) The proposed two-stage algorithm can solve the capacitated vehicle routing problem (CVRP) with large capacity constraint. The study provides a novel method to solve the actual large-scale vehicle routing problem.

Keywords: Capacitated vehicle routing problem, Two-stage algorithm, Ant colony algorithm

1. Introduction

With the rapid development of e-commerce, the application demand of vehicle distribution is continuously increasing. However, vehicle distribution has several restrictions, such as vehicle capacity, route, customer, distribution center, and road network [1]. In addition, many serious problems have been emerging, including increasingly complex road traffic, high transportation costs, and changing customer demands. Therefore, researchers are notably challenged to efficiently solve the capacitated vehicle routing problem (here in after referred to as CVRP) and consequently carried out several studies to solve this problem. Heuristic algorithms, such as Tabu Search, Sweep, genetic algorithm, and ant colony algorithm, have been used to solve the problem, which have so far achieved a significant progress [2–4].

Most laboratory studies focus on small-scale problems, because CVRP belongs to NP-hard problem, which indicates its computational complexity [5]. The number of nodes in the CVRP example is generally less than 1,200 [6]. However, the scale of continuously increases as the logistics industry rapidly develops, with actual application of over tens of thousands and even tens of thousands of nodes. Meanwhile, several scholars have discovered 1,200 to 20,000 application nodes in large-scale CVRP [7]. This study result can hardly meet the demands of large-scale vehicle distribution application. Therefore, designing an algorithm that cannot only meet the actual demand of large-scale vehicle distribution but also increase the solution efficiency and improve the solution quality is a top priority.

Consequently, the two-stage algorithm is studied, which reduces the problem scale and improves solution quality through clustering and is expected to prove its applicability and advantage for solving large-scale CVRP and provide a reference for relevant study.

2. State of the art

Plenty of studies of CVRP have been carried out by scholars. Baldacci R. et al. designed a double-rise heuristic algorithm [8], which significantly improved the solution efficiency of CVRP. However, the algorithm was less suitable for solving large-scale CVRP, due to the high complexity of the algorithm based on set partitioning. Contardo C. et al. explored the multi-depot vehicle routing problem with path length and capacity constraints [9], and the lower limit was calculated using the vehicle flow formula. Although the complexity of solving the set partition was reduced, high requirements for lower limit, which led to the difficulty of solving several CVRP examples, are increasing. Dell'Amico dealt with the route problem with mixed capacity constraint

*E-mail address: lengxuer_505@163.com

doi:10.25103/jestr.112.16

using iterated local search meta heuristic algorithm [10] and adopted the neighborhood search operator to enhance the local search efficiency. However, the solution efficiency is high. The larger the generation matrix of the edge (arc) is, the longer the time and the higher the time complexity will be. Considering the influence of the distribution route length on drivers' income, Jorge Oyola et al. sought a CVRP solution for path equilibrium [11], which minimized the differences in the route and total lengths. Although good approximate Pareto set was obtained, the distribution law of customers in distribution area and customers' demand for distribution time were not fully considered. On the basis of the two environments , mathematical and constraint logic programming, Pawel Sitek solved the two-echelon CVR [12]. The algorithm has high complexity due to the parameters such as multiple constraints and multiple targets. Mahvash B. et al. overcame the vehicle routing problem with three-dimensional loading capacity constraints using a heuristic algorithm based on column generation technique [13]. The three-dimensional loading constraints were considered, but a few limitations in practical application were observed due to the complexity of minimum supporting surface and the transportation demand. In view of the absolute minimum vehicle number preference and the path optimization preference, Wang Chao et al. explored the multi-objective solution method for CVRP with different target preferences [14], but they failed to respond to the changes of customer demands in time.

Scholars proposed a variety of improved strategies for the existing heuristic algorithms of solving CVRP. For example, Wan Bo et al. optimized the hybrid leapfrog algorithm by integrating adaptive differential perturbation mechanism and chaotic local search strategy [15]. Fu Zhengtang et al. improved the simulated annealing algorithm through inter-path adjustment and intra-path optimization [16]. Ma Xiaolu et al. enhanced the population diversity of the basic genetic algorithm by designing crossover operators of paternal and individual summation [17]. Pang Qi et al. introduced the operator operation to solve the discrete problem of artificial bee colony algorithm [18]. Li Mingjv et al. solved the disadvantages of the traditional Tabu Search algorithm on the basis of the principle of hierarchical cost structure by adding reservation table [19].

The above methods optimize the solution quality and efficiency of CVRP to some extent. However, these methods are more suitable for medium and small-scale CVRP and have certain limitations in solving the actual problem. A few study results of large-scale CVRP are available, particularly relatively large-scale CVRP.

A three-class and three-tier distribution model was established for large-scale CVRP, and a two-stage algorithm was proposed based on this model. Large-scale CVRP was transformed into several small- and medium-scale CVRPs through clustering, which reduced the problem scale and the complexity of the two-stage algorithm. The 2-OPtimization (2-OPT) strategy was used to improve the solution quality of the two-stage algorithm. Finally, Pearson correlation coefficient and quality optimization rate index are introduced to analyze the simulation. Furthermore, the two-stage algorithm is proven to be suitable and efficient in solving large-scale CVRP.

The rest of this study is organized as follows. Section 3, the three-class and three-tier model of CVRP is constructed. Section 4 carries out the simulation by using the two-stage algorithm CVRP of Longquan Logistics Center in Sichuan Province as an example. Finally, Section 5 provides the conclusions.

3. Methodology

3.1 Establishment of CVRP three-class and three-tier model

Large-scale CVRP has a complicated distribution structure and a wide distribution area in practical application. Consequently, the three-class and three-tier model was established, as shown in Fig. 1. The model divided the distribution structure of large-scale CVRP into three classes, namely, fist-and second-class distribution centers and third-class distribution site, which resulted in the reduction of the complexity of distribution structure. Meanwhile, the three-tier distribution improved the distribution efficiency. The first tier is the delivery of goods from logistics center to the first-class distribution center in bulk. The second tier is the delivery of goods from the first- to the second-class center. The third tier is the transportation of goods from the second-class distribution center to the third-class distribution site. The logistics center delivers the goods of the carload, which are mainly from factories, suppliers, and consignors. Lastly, the third-class distribution site is the destination for vehicle delivery, including sales stores, consignees (customers), and distribution service stations.

Therefore, the solution of large-scale CVRP is actually the solution of three-class and three-tier model.

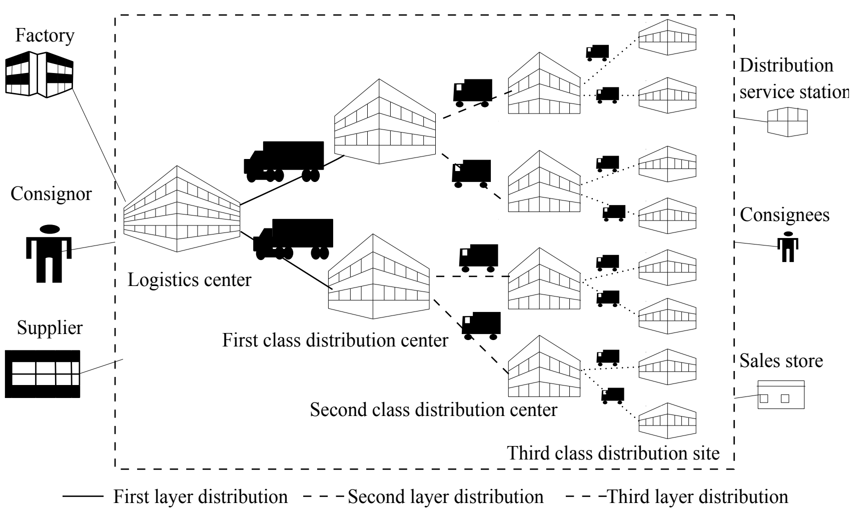

Fig. 1. CVRP three-class and three-tier simulation model

3.2 Solution of CVRP three-class and three-tier model by two-stage algorithm

The two-stage algorithm integrating the K-means algorithm and 2-OPT improved ant colony algorithm were used to solve the large-scale three-class and three-tier CVRP model.

In the first stage, a three-class distribution structure is established. K-means algorithm was used to partition clustering. The distribution structure including the first-, second-, and third-class distribution sites based on the clusters obtained through clustering partitioning was established.

In the second stage, the vehicle distribution route is solved. 2-OPT improved ant colony algorithm based on the clustering results was used for three-tier distribution routes.

3.2.1 First stage: establishment of a three-class distribution structure

In transportation distribution, multiple vehicles distribute goods to customers, and the distribution volume is large. Thus, an efficient algorithm is needed to satisfy all kinds of constraints, such as non-overlapping regions and similar distribution volume. K-means is a classical algorithm based on partition and is suitable for the data type. However, K-means is insensitive to the order of data input.

The set is denoted with $X=\{x_1, x_2, ..., x_n\}$, the number of clusters with k, the clusters with C_j $(j=1,2,...,k)$. Each cluster includes $\{n_1^k, n_2^k, ..., n_k^k\}$ nodes. The Euclidean distance between two data objects is calculated using Formula (1). The criterion function E is used as the objective function to measure the clustering quality, as shown in Formula (2) :

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{n}(x_{ik} - x_{jk})^2} \qquad (i \neq j) \tag{1}$$

$$E = \sum_{i=1}^{k}\sum_{x \in C_i} d^2(C_i, x) \tag{2}$$

The specific steps for establishing a three-class distribution structure are as follows:

Step-l: Let $I=1$. Select randomly the k data points as the initial cluster center of k clusters, and obtain the initial value of the cluster center set $m_j(I)$ $(j=1,2,...,k)$.

Step-2: Compare the distance $d(x_i, m_j(I))$ $(i=1,2, ...,n; j=1,2,...,k)$ between each data point and k cluster center according to Formula (1). If $d(x_i, m_j(I))=\min\{d(x_i, m_j(I)), j=1,2,...,k\}$ is satisfied, then $x_i \in C_j$,and x_i point is placed in cluster C_j .

Step-3 : Recalculate the k new clustering centers according to $m_j(I+1) = \sum_{i=1}^{N_j} x_i / N_j$ $(j=1,2,...,k;\ x_i \in C_j)$.

Step-4: Calculate the criterion function E using Formula (2).

Step-5: Determine whether it is convergence. If convergence, then return to clustering. Otherwise, execute Step-2.

Step-6：Establish a three-class structure based on the obtained cluster C_j $(j=1,2,...,k)$. Establish the first-class distribution center by referring to cluster center set $m_j(I)$ .Let $\{n_1^k, n_2^k, ..., n_k^k\}$ nodes in each C_j $(j=1,2,...,k)$ except for the first-class distribution center be the second-class distribution centers. The third-class distribution site refers to all destination nodes included in each node in the second-class distribution centers $\{n_1^k, n_2^k, ..., n_k^k\}$.

3.2.2 Second stage: solving the three-tier distribution route

The first tier distribution is the round trip of the vehicle. The 2-OPT was used to improve the ant colony algorithm to solve the route. The route of the second and third tier distributions can be solved using 2-OPT improved ant colony algorithm based on the clustering results.

(1) Problem description

The CVRP requires the scientific planning of the distribution route so that the mileage is at the minimum and the vehicle capacity is satisfied under the condition of completing the distribution tasks.

The specific description is as below. Weight graph $G=(V,E)$ is used to describe traffic network. Set $V=\{0, 1, 2,...,n\}$ represents one distribution center and n customers. Edge set $E=\{(i,j)|\ 0 \le i \ne j \le n\}$ represents the edge formed by any two points. d_{ij} is the edge length weight. Several m vehicles have the maximum load of Q. Customer demand is denoted with q_i $(i=1,2,...n)$.

The objective function is as follows:

$$\min z = \sum_i \sum_j \sum_k d_{ij} x_{ijk} \tag{3}$$

The constraints are as follows:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} q_i x_{ijk} <= Q \qquad (1 \le k \le m) \tag{4}$$

$$\sum_{i=0}^{n}\sum_{j=0}^{n} d_{ij} x_{ijk} \le L \qquad (1 \le k \le m) \tag{5}$$

$$\sum_{j=1}^{n}\sum_{k=1}^{m} x_{ijk} = 1 \qquad (1 \le i \le n) \tag{6}$$

$$\sum_{i=1}^{n}\sum_{k=1}^{m} x_{ijk} = 1 \qquad (1 \le j \le n) \tag{7}$$

$$\sum_{j=1}^{n}\sum_{k=1}^{m} x_{ojk} = \sum_{i=1}^{n}\sum_{k=1}^{m} x_{iok} = m \tag{8}$$

$$x_{ijk} = \begin{cases} 1 & \text{Vehicle i passes arc } (i,j) \text{ and moves from i to j} \\ 0 & \text{Otherwise} \end{cases} \qquad (0 \le i,\ j \le n, 1 \le k \le m) \tag{9}$$

Formula (3) is the objective function. That is, the objective of the minimum route length of all m vehicles should be achieved. Formula (4) is the load constraint of each vehicle. Formula (5) is the constraint on the operating distance of each vehicle, Formulae (6) and (7) guarantee that one vehicle serves a particular customer. Formula (8) is the distribution center which is the starting and ending points of all vehicles. Finally, Formula (9) indicates that the route of each vehicle is a simple loop.

(2) Solution method

The basic ant colony algorithm is one of the commonly used heuristic algorithms for solving CVRP, which is a population-based evolutionary algorithm proposed by Italian scholars M. Dorigo et al. by simulating the ant colony foraging behavior, a probabilistic algorithm used to find the optimal path. Substantially, this is a distributed parallel algorithm. Using the positive feedback principle can accelerate the evolutionary process to a certain extent and also has strong robustness and the ability to find better solutions. However, a few defects such as slow convergence rate and local optimal solution are recognized.

Furthermore, the 2-OPT strategy was introduced for local search and optimization. The solution for nodes that do not meet the vehicle load condition is obtained by exchanging its access order, which accelerates the convergence speed of the ant colony algorithm and optimizes the quality of solution.

The processes of solving vehicle distribution route based on 2-OPT improved ant colony algorithm are as follows:

Step-1: Parameter initialization. Let time $t=0$ and cycle time $N_c=0$ and the maximum cycle time be $N_{c\max}$, put m ants on n elements (cities), let the initialization information amount $\tau_{ij}(t)=const$ on each edge(i,j) of a directed graph be $\tau_{ij}(t)=const$, where const represents a constant and the initial time is $\Delta\tau_{ij}(0)=0$.

Step-2: Let $N_c = N_c +1$ the cycle time.

Step-3: In *tabu* list of ant *k*, $k=1$. Let $tabu_k$ $(k=1,2, ...,m)$ records the city which the ant *k* is currently passing through and adjust the set dynamically with the $tabu_k$ evolution process.

Step-4: Let $k=k+1$ be the number of ants.

Step-5: Calculate the transition probability $p_{ij}^k(t)$ based on Formula (10) and choose city *j* by roulette to calculate whether the load constraints are satisfied or not. If the constraints are satisfied, then the ant individual *k* is transferred from city *i* to city *j*, as shown in the following formula:

$$p_{ij}^k(t)=\begin{cases}\dfrac{\left[\tau_{ij}(t)\right]^{\alpha}\cdot\left[\eta_{ij}(t)\right]^{\beta}}{\sum\limits_{s\in J_k(i)}\left[\tau_{is}(t)\right]^{\alpha}\cdot\left[\eta_{is}(t)\right]^{\beta}}, if\ \ j\in J_k(i)\\ 0,\quad else\end{cases} \tag{10}$$

Where α is information heuristic factor and represents the relative importance of the trajectories. The greater the value, the more the ants tend to choose the path of the majority of other ants. Thus, the collaboration is stronger, and the study is less random. The smaller the value, the earlier the ant colony search is prematurely trapped in the local optimum. β is the expected heuristic factor and represents the relative importance of visibility. The greater the value, the more likely the ant chooses the local shortest path. Although the convergence speed of the algorithm is accelerated, the randomness of the path search is weakened and causes an easy fall into the local optimal. $J_k(i)$ is the city that the ant *k* can choose in the next step, τ_{ij} is information amount on edge(i,j), and η_{ij} is a heuristic function, which reflects the heuristic degree of ants moving from city *i* to city *j*. $\eta_{ij}(t)$ is obtained using Formula (11):

$$\eta_{ij}(t)=\frac{1}{d_{ij}} \tag{11}$$

Where d_{ij} represents the distance between two adjacent cities. For ant k, the smaller the d_{ij}, the greater the $\eta_{ij}(t)$ and $p_{ij}^k(t)$, which indicates that the expectation that the ant moves from city *i* to city *j* is greater.

Step-6: Modify the *Tabu* list pointers. That is, the ant moves to a new city, and the city is added to the *Tabu* list.

Step-7: If the cities in set *C* are not traversed, then let $k<m$ and jump to Step-4.

Step-8: Check for the optimal solution. If an optimal solution is found, then implement the 2-OPT local optimization. Otherwise, jump to Step- 9.

Step-9: To prevent the excessive residual information from submerging the heuristic information, update the residual information after each ant passes through or traverses all the *n* cities and update the information amount on the route (i,j) at $t+n$ according to Formulae (12) and (13) :

$$\tau_{ij}(t+n)=(1-\rho)\cdot\tau_{ij}(t)+\Delta\tau_{ij} \tag{12}$$

$$\Delta\tau_{ij}=\sum_{k=1}^{m}\Delta\tau_{ij}^k \tag{13}$$

$$\Delta\tau_{ij}^k=\begin{cases}\dfrac{Q}{L_k}, & \text{If ant k passes the edge } (i,\ j)\\ 0, & \text{Otherwise}\end{cases} \tag{14}$$

where $\Delta\tau_{ij}$ is the information increment on edge(i,j) in this iteration, $\Delta\tau_{ij}^k$ is the information amount left by the ant on edge(i,j) in this iteration, ρ is information evaporation (or volatilization) coefficient, and $1-\rho$ is the persistence coefficient $(0<\rho<1)$.

Formula (14) denotes the information updating strategy. *Q* indicates the information intensity, which affects the convergence rate of the algorithm to some extent. L_k represents the total length of the route passed by *k* ants in the cycle.

Step-10: If the end condition is satisfied, that is, if the number of cycles $N_c \ge N_{c\max}$, then end the cycle and output the route solution. Otherwise, clear the *Tabu* list and jump to Step2.

4 Result Analyses and Discussion

The simulation was carried out based on MATLAB 2016 by using the distribution vehicle routing problem in the highway freight transportation planning of Longquan Logistics Center in Sichuan Province as an example.

4.1 Simulation data

Sichuan Province has 21 cities (prefectures), 183 counties, and an area of 486,000 square kilometers, as shown in Fig. 2. Solving the route using the straight line distance is difficult, because the three ethnic autonomous prefectures of Ganzi, Liangshan, and Aba in Sichuan Province belong to the alpine plateau area and the terrain is complex. Moreover, the population density is small. Therefore, only three prefecture capitals are considered in the simulation: Kangding, Xichang, and Malkang. However, a total of 138 counties (cities and districts) are involved in the simulation.

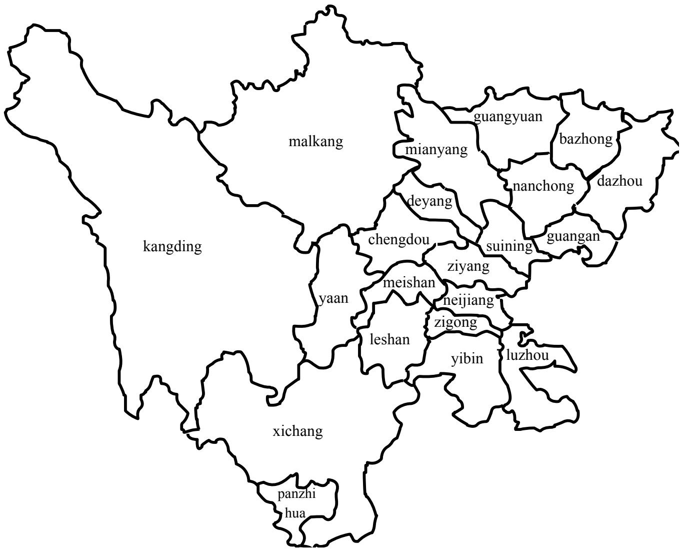

Fig. 2. Administrative map of Sichuan Province

The longitude, latitude, and cargo demand of 138 counties (cities and districts) in Sichuan Province are shown in Table 1. The data of each region are represented by vectors (city number, longitude, latitude, and cargo demand). The codes of administrative division of 138 counties (cities and districts) in Sichuan Province are arranged from small to large. Longitude and latitude data are collected using the Baidu map coordinate pickup system point by point according to city number. The demand for goods is estimated on the basis of the 2016 GDP index of the counties (cities and districts), which is three times the difference between the GDP index of the counties (cities and districts) and 100.

Longitude and latitude data source: Baidu Map coordinate pickup system:

http://api.map.baidu.com/lbsapi/getpoint/index.html.

Administrative division code source: Statistics Bureau of the People's Republic of China:

http://www.stats.gov.cn/tjsj/tjbz/tjyqhdmhcxhfdm/2016/51.html.

The source of GDP index of counties (cities and districts) in 2016: "Table 12-4 GDP index of Counties (cities, districts) (2016) "in 2017 Sichuan Statistical Yearbook:

http://www.sc.stats.gov.cn/tjcbw/tjnj/

Table. 1. Longitude, latitude, and cargo demand of 138 counties (cities and districts) in Sichuan

City number, longitude, latitude, cargo demand	**City number, longitude, latitude, cargo demand**	**City number, longitude, latitude, cargo demand**	**City number, longitude, latitude, cargo demand**	**City number, longitude, latitude, cargo demand**	**City number, longitude, latitude, cargo demand**
1,104.08,30.67,21	2,104.05,30.68,22	3,104.05,30.7,20	4,104.05,30.65,22	5,104.1,30.67,23	6,104.27,30.57,12
7,104.23,30.88,26	8,104.15,30.83,24	9,103.83,30.7,24	10,104.43,30.85,32	11,103.92,30.58,24	12,103.88,30.82,37
13,103.52,30.58,32	14,103.5,30.2,32	15,103.82,30.42,36	16,103.62,31,29	17,103.93,30.98,23	18,103.47,30.42,32
19,103.67,30.63,34	20,104.55,30.4,23	21,104.77,29.35,18	22,104.72,29.35,27	23,104.77,29.37,21	24,104.87,29.27,31
25,104.42,29.47,24	26,104.98,29.18,26	27,101.7,26.55,22	28,101.6,26.6,21	29,101.73,26.5,27	30,102.12,26.88,30
31,101.85,26.7,26	32,105.45,28.88,30	33,105.37,28.77,28	34,105.43,28.9,30	35,105.38,29.15,29	36,105.83,28.82,28
37,105.43,28.17,26	38,105.82,28.05,28	39,104.38,31.13,25	40,104.68,31.03,26	41,104.5,31.32,32	42,104.28,30.98,28
43,104.17,31.13,21	44,104,31.6,27	45,104.73,31.47,27	46,104.75,31.47,25	47,105.08,31.1,25	48,105.38,31.22,22
49,104.57,31.53,22	50,105.17,31.63,23	51,104.45,31.82,25	52,104.53,32.42,17	53,104.75,31.78,25	54,105.05,29.58,23
55,105.97,32.32,25	56,105.88,32.65,25	57,106.28,32.23,23	58,105.23,32.58,23	59,105.52,32.28,24	60,105.93,31.73,26
61,105.57,30.52,29	62,105.45,30.35,29	63,105.72,30.78,25	64,105.38,30.87,28	65,105.25,30.58,25	66,105.05,29.58,22
67,105.07,29.6,23	68,104.67,29.53,23	69,104.85,29.78,24	70,105.28,29.35,25	71,103.77,29.57,26	72,103.55,29.42,28
73,103.82,29.4,24	74,103.08,29.25,13	75,103.95,29.22,27	76,104.07,29.65,22	77,103.57,29.73,27	78,103.9,28.97,18
79,103.27,29.23,21	80,103.55,28.83,21	81,106.08,30.78,26	82,106.1,30.77,25	83,106.05,30.77,25	84,106.07,31.35,24
85,106.57,31.08,23	86,106.42,31.03,25	87,106.28,31.27,24	88,105.88,31,24	89,106,31.55,24	90,103.83,30.05,23
91,104.15,30,26	92,103.87,30.2,26	93,103.37,29.92,24	94,103.52,30.02,25	95,103.85,29.83,26	96,103.29,29.36,25
97,104.62,28.77,24	98,104.55,28.7,26	99,104.98,28.85,26	100,105.07,28.73,27	101,104.92,28.58,24	102,104.52,28.43,24
103,104.72,28.45,25	104,104.52,28.17,25	105,105.23,28.3,25	106,104.33,28.83,23	107,106.43,30.55,23	108,106.28,30.35,23
109,106.93,30.33,25	110,106.32,30.18,24	111,106.38,30.26,25	112,106.77,30.38,24	113,107.48,31.22,15	114,107.5,31.2,22
115,107.72,31.35,25	116,107.87,31.08,22	117,107.2,30.73,25	118,106.97,30.83,25	119,108.03,32.07,20	120,103,29.98,20
121,103.12,30.08,25	122,102.85,29.8,25	123,102.65,29.35,23	124,102.37,29.23,24	125,102.75,30.07,27	126,102.92,30.15,31
127,102.82,30.37,31	128,106.77,31.85,26	129,106.63,31.8,31	130,107.23,31.92,19	131,106.83,32.35,21	132,107.1,31.57,24
133,104.65,30.12,25	134,105.33,30.1,22	135,105.02,30.28,24	136,103.58,31.48,26	137,101.97,30.05,24	138,102.27, 27.9,18.9

4.2 Solution of the distribution route in 138 counties (cities and districts)

Considering the policy, site, population density, economic development situation, and convenient transportation, the scholars selected as an example Longquan Logistics Center of Sichuan Province to carry out the simulation experiment. The center longitude of Longquan Logistics Center is 104.27,30.57, which is one of regional integrated logistics centers in Sichuan Province, located in Longquanyi District, Chengdu City, the capital of Sichuan Province.

4.2.1 Solving the distribution route using the ant colony algorithm

The parameter setting of the ant colony algorithm directly affects the algorithm performance. The optimal parameter setting is obtained after multiple tests: information heuristic factor α =1.0, expected heuristic factor β =5.0, information evaporation coefficient β =0.05, iterations $N_{c\,\max}$ =200, and vehicle capacity Q=200. The overall number of ants is 138.

Using the basic ant colony algorithm, the optimal route length from Longquan Logistics Center to 138 counties (cities and districts) is 177.88, which takes 106.617 seconds to complete. The optimal route is shown in Fig. 3.

As shown in Fig. 3, the region span of several routes is large, with many cross regions, resulting in the increase of the transportation cost and affecting the transportation efficiency to a certain extent.

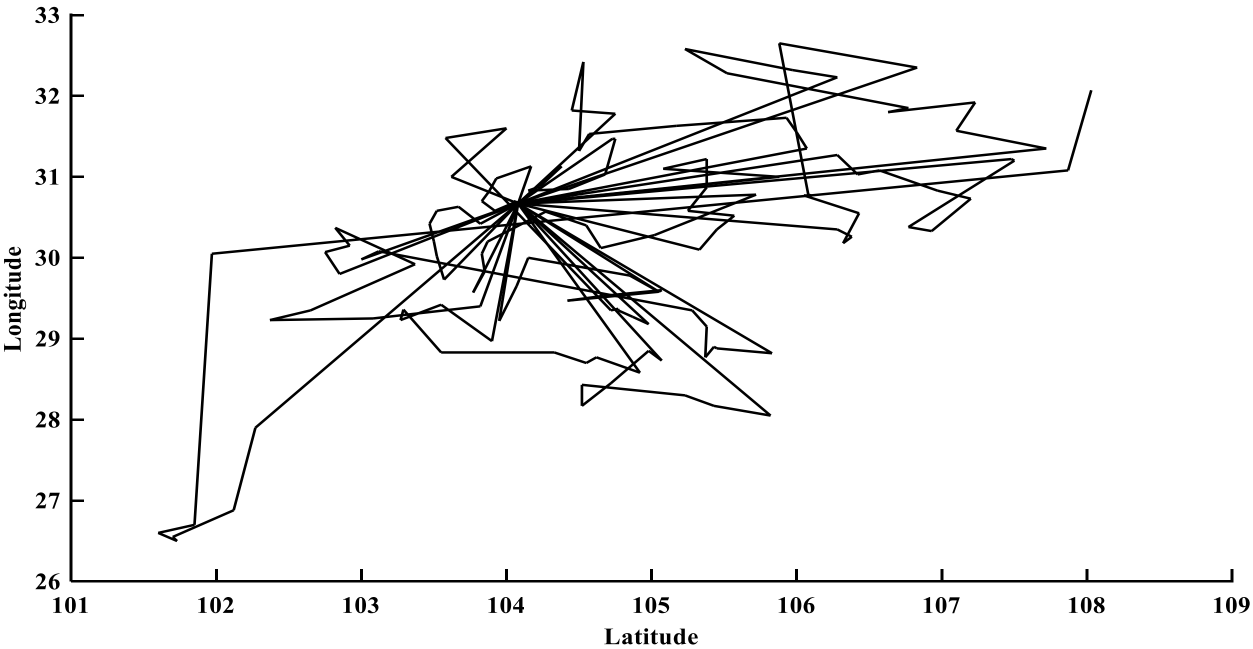

Fig. 3. First class optimal route solved using the basic ant colony

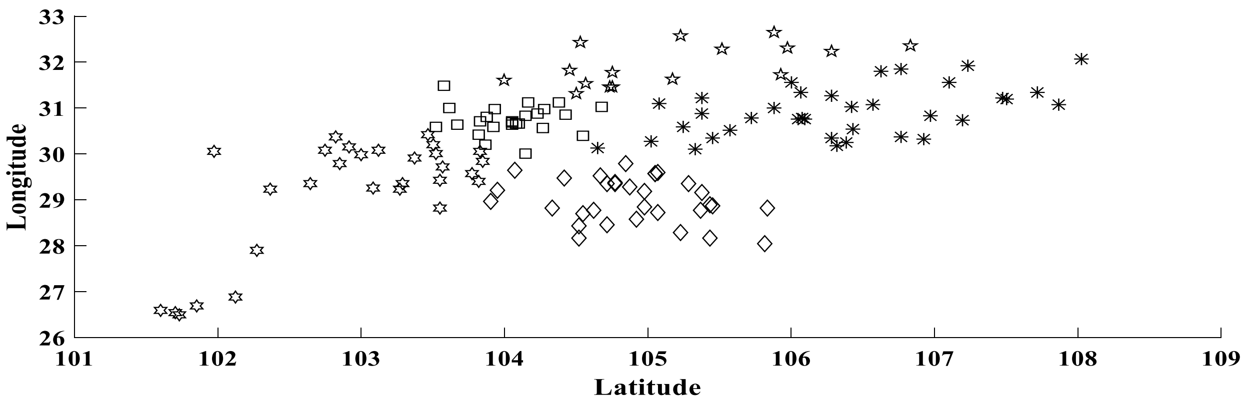

Fig. 4. Cluster map of 138 counties (cities and districts) in Sichuan Province

4.2.2 Solving the distribution route using two-stage algorithm

(1) Establishment of three-class distribution structure

Using the solution method proposed in Section 3.2.1, the two-stage algorithm was used to cluster 138 nodes in Sichuan Province on the basis of its latitude and longitude. Five optimal clusters were obtained, as shown in Fig. 4. The central coordinates of the clusters are the following: 104.05,30.68; 104.72,29.35; 104.75,31.78; 105.25,30.58 and 103.37, 29.92.

From Fig. 2 and 4, the cluster center obtained by cluster analysis is consistent with the developed logistics area in Sichuan Province, which indicates that the area division using clustering algorithm is effective.

The five-cluster centers are the first-class distribution centers. The cluster nodes except for cluster centers are the second-class distribution centers, that is, 133 counties. The third-class distribution sites refer to the destination nodes corresponding to each second-class distribution center. Thus, the three-class distribution structure of Longquan Logistics Center in Sichuan Province was established.

(2) Solution of the first tier distribution route

The two-stage algorithm proposed in Section 3.2.2 was used to solve the first tier distribution route from Longquan Logistics Center to the first-class distribution center based on the clustering results. The node where the cargo demand is 0 is the logistics center. The cargo demand of the first-class distribution center is the sum of the second-class distribution centers in the same cluster, as shown in Table 2. Finally, the optimal path length is 50.6, and the time to complete the route is 32.22 seconds. The optimal route is the round trip of the vehicle, as shown in Fig. 5.

Table.2. Latitude, longitude, and cargo demand of the logistics and first-class distribution centers

City number	Latitude	Longitude	Cargo demand
11	104.27	30.57	0
2	104.05	30.68	642
22	104.72	29.35	802
53	104.75	31.78	390
65	105.25	30.58	872
93	103.37	29.92	722

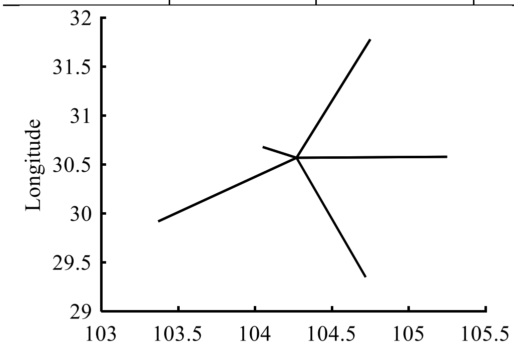

Fig. 5. First-class distribution route

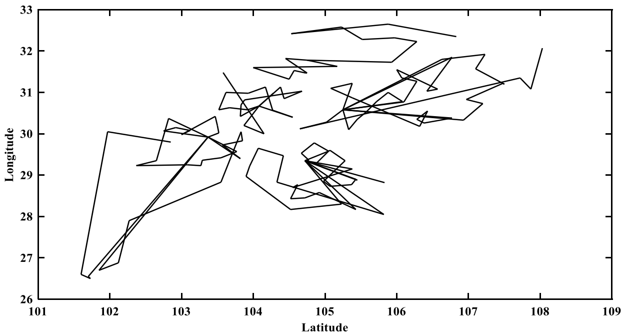

Fig. 6. Optimal routes of five clusters

(3) Solution of the second tier distribution route

On the basis of the solution method in Section 3.2.2, using the cluster divided by the clustering, the vehicle load is balanced using the same parameter setting as the basic ant colony algorithm under the precondition of satisfying the load constraints. The optimal route length of the second-class distribution is 108.91, which takes 56.41 seconds. During the 200 iterations, 2-OPT is operated for seven times, and the optimization rate is 3.5%. Finally, the optimal vehicle routes of five clusters are solved, as shown in Fig. 6. The average full load rates are 81.4%, 86.4%, 79.25%, 90.92%, and 80%.

In Fig. 6, each route passes through a small area span, which is conductive to solving the second tier distribution path satisfying the vehicle load in a short period of time. Combined with the load of the five clusters, the average vehicle load rate is 83.59. The vehicle load is relatively balanced, and the vehicle arrangement is reasonable.

4.2.3 Analysis of simulation

To evaluate the performance of the two-stage algorithm, the quality optimization rate index is introduced. Let the optimal path length obtained by two-stage algorithm and basic ant colony algorithm be x and y. The calculation method is: 1-x/y.

The results of distribution routes in 138 counties (cities and districts) solved using the two-stage and basic ant colony algorithms are the following: The optimal route length is 159.51 and 177.88, and the consumed time is 86.63 seconds and 106.617 seconds, respectively. Meanwhile, the optimization rate of 2-OPT is 3.5%. Compared with the operation results, the time consumed by the two-stage algorithm is decreased by 18.74%, and the route length is decreased by 18.37%. The quality optimization rate reaches 10.33%. Hence, the two-stage algorithm not only improves the solution efficiency but also improves the solution quality to a certain extent.

4.3 Solution of the third tier distribution route

The second-class distribution center with high population density was selected, and ten examples with 600–12,000 third-class distribution points were randomly generated, in the range of longitude and latitude of the cluster. The two-stage and basic ant colony algorithms were used to solve the optimal route of the third tier distribution. The optimal route length and quality optimization rate of the two algorithms are shown in Table 3. Fig. 7 shows the route length solved by the two algorithms. Fig. 8 shows the quality optimization rate of the two-stage algorithm.

Table. 3. Optimal route length and quality optimization rate of two-stage and basic ant colony algorithms

Solution efficiency	x	y	Quality optimization rate (%)
CVRP-1000	3876.52	6693.46	42.09%
CVRP-2000	7659.47	13707.00	44.12%
CVRP-3000	16050.05	30490.22	47.36%
CVRP-3500	19245.35	37110.20	48.14%
CVRP-4000	24440.64	48979.24	50.10%
CVRP-4800	29691.55	61318.31	51.58%
CVRP-5200	32317.00	69290.31	53.36%
CVRP-6000	34567.90	79907.30	56.74%
CVRP-7200	37444.26	91887.75	59.25%
CVRP-8000	50695.17	135476.14	62.58%
CVRP-9100	59359.45	171906.89	65.47%
CVRP-10000	62812.04	202619.48	69.00%

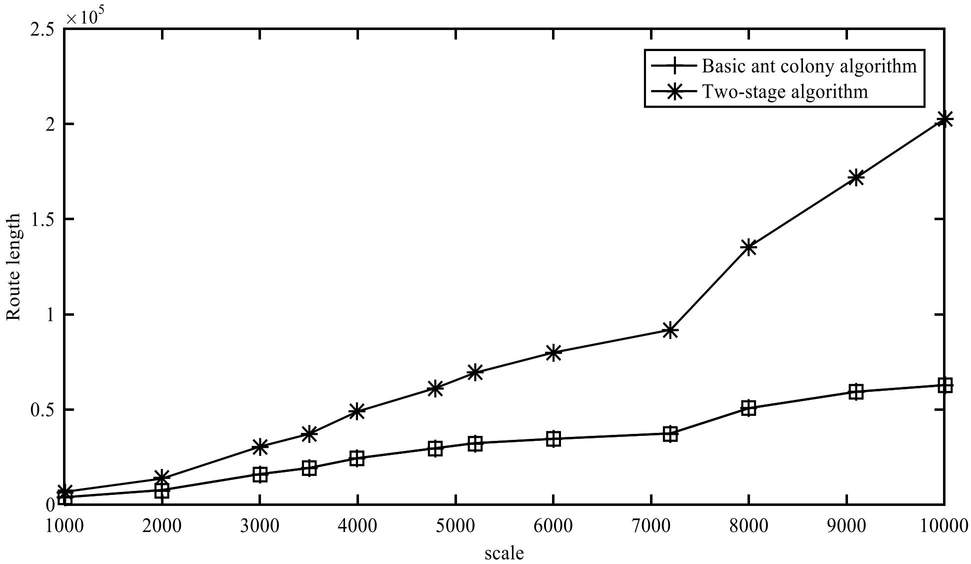

Fig. 7. Route length solved by the two algorithms

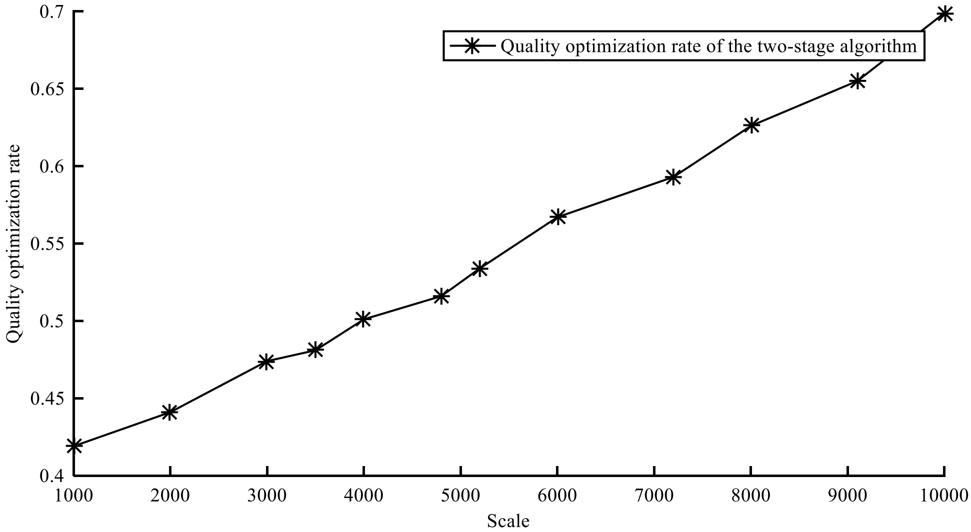

Fig. 8. Quality optimization rate of two-stage algorithm

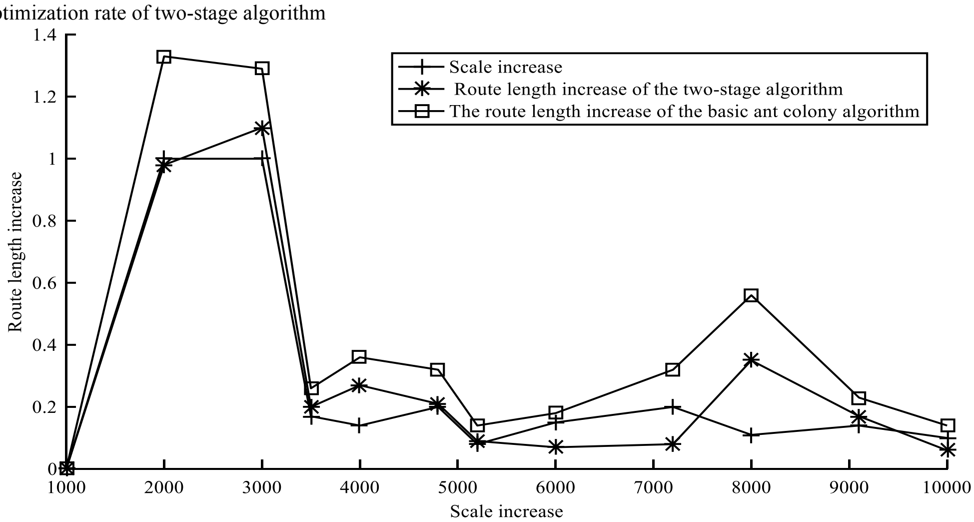

Fig. 9. Broken-line graph of the scale increase and the route length increase

As shown in Table 3 and Fig. 7, the route length obtained using the basic ant colony algorithm is higher in comparison with the two-stage algorithm. With the expansion of the solution efficiency, the difference among the route length values is greater. As shown in Fig. 8, the larger the problem scale is, the more obvious the upward trend of the broken line of the quality optimization rate of the two-stage algorithm will be, and the greater the increase extent, which indicates that the algorithm quality is better. Based on the calculation, the average quality optimization rate is 54.13%. In general, the larger the scale is, the more significant the advantages of the two-stage algorithm over the basic ant colony algorithm will be.

(1) Analysis on the correlation between the performance of two-stage algorithm and the increase of problem scale

To study the relationship between the solution quality and the solution efficiency of the two-stage algorithm, a broken-line graph of the scale increase and the route length increase of the two algorithms is as shown in Fig. 9.

As shown in Fig. 9, as the scale increases, the route length of the basic ant colony algorithm is significantly longer than the scale increase. The solution quality of the two-stage algorithm is less affected by the scale, and the negative correlation even occurs in several regions.

Pearson correlation coefficient is introduced to quantify the relationship between quality and scale. Pearson correlation coefficient is used to measure the linear correlation between two variables X and Y, and its value is between –1 and 1. The formula is as follows:

$$Corr(X,Y)=\frac{n\sum xy-\sum x\sum y}{\sqrt{n\sum x^2-(\sum x)^2}\sqrt{n\sum y^2-(\sum y)^2}} \tag{15}$$

In Formula (15), the correlation between the solution quality and the problem scale of the basic ant colony algorithm and the two-stage algorithm is 0.9823 and 0.9915, respectively. The results show that the solution quality of the basic ant colony algorithm has a lower correlation with the scale. That is, the larger the scale is, the more discrete the solution results will be. On the contrary, the solution quality of two-stage algorithm is correlated with its scale. The larger the scale is, the more stable the solution quality will be. This result proves that the two-stage algorithm has good robustness and convergence for solving large-scale CVRP.

(2) Analysis of time and space performance of two-stage algorithm

To further prove the advantage of the two-stage algorithm, the time and space complexities are introduced for comparison and analysis. n is the number of the third-class distribution sites, k is the number of clusters, m is the number of ants, and T is the number of iterations. The number of iterations is set to the multiple of the third-class distribution sites, and the number of ants is equal to the number of third-class distribution sites.

Therefore, the space complexity of the basic ant colony algorithm is $O(3*n*n+n*m)$, and the time complexity is $o(n*(n-1)*m*T\div 2)$. When n tends to infinity, the space complexity of the basic ant colony algorithm is $O(n^2)$, and the time complexity is $O(n^4)$. The space complexity of the two-stage algorithm is expressed as $O(n)+O(\sum_{i=1}^{k} n_i^2)$, where $\sum_{i=1}^{k} n_i=n$. The time complexity is expressed as $O(T*n*k)+O(\sum_{i=1}^{k} n_i^4)$, where $\sum_{i=1}^{k} n_i=n$.

When $n=10000$, $k=5$, $m=n$, and $T=10000$, the space and time complexities of the basic ant colony algorithm are 4.0E+08 and 1.0E+17, respectively, where as space and time of the two-stage algorithm are 2.0E+07 and 8.1E+13, respectively, which are much smaller than those of the basic ant colony algorithm. The result proves that the two-stage algorithm not only reduces the problem scale and improves the solution efficiency but also has certain advantage.

5. Conclusions

To effectively solve the large-scale CVRP, a three-class and three-tier distribution model was established, and the optimal route was solved using the two-stage algorithm which first clustered and then searched the route. The results of the distribution vehicle routing problem in the highway freight transportation planning of Longquan Logistics Center in Sichuan Province were analyzed on the basis of the quality optimization rate. Finally, the following conclusions could be drawn:

(1) Based on the CVRP three-class and three-tier distribution model, the space division clustering and 2-OPT strategies proposed by the two-stage algorithm notably reduce the time and space complexities of the algorithm.

(2) The larger the CVRP scale, the higher the quality optimization rate of the two-stage algorithm, and algorithm performance has notable advantage.

(3) On the basis of the relationship between the solution quality and the scale increase, the solution quality of the two-stage algorithm is less affected by the scale difference, and the two-stage algorithm has good robustness.

(4) On the basis of the speed and satisfactory quality, the proposed two-stage algorithm is not only suitable for solving large-scale CVRP but also has certain advantage.

The two-stage algorithm has local and global optimization performance, which could be the solution of the actual relatively large-scale or large-scale vehicle routing problems. Moreover, the constraints will be updated in future studies to make the vehicle route more feasible, due to the constantly changing constraints required in practical application, thus reducing the logistics costs.

Acknowledgements

This work was supported by the Sichuan Province Social Science Key Research Base (Grant Nos. WHCY2016b10, SDJJ1627 and DSWL16-5).

References

1. Psaraftis, Harilaos N, M. Wen, and C. A. Kontovas.,"Dynamic vehicle routing problems: Three decades and counting ". *Networks* ,67(1),2016,pp.3-31.
2. Sarasola, Briseida, et al.,"Variable neighborhood search for the stochastic and dynamic vehicle routing problem". *Annals of Operations Research* ,236(2),2016,pp.425-461.
3. Gounaris C E, Wiesemann W, Floudas C A.,"The Robust Capacitated Vehicle Routing Problem Under Demand Uncertainty". *Operations Research*, 61(3),2013,pp. 677-693.
4. Li Jun, Hu Xiangpei, Wang Zheng.,"Reviews on Vehicle Routing Problem and Its Solution Methods". *Systems Engineering*, 24(11),2013,pp. 31-37.

5. Rao Weizhen, Jin Chun.," An Efficient Heuristic for Solving Large-scale Capacitated Vehicle Routing Problem". *Journal of Industrial Engineering and Engineering Management*, 28(2),2014, pp. 45-54.
6. Kytojokia J,Nuortio T,Braysy O,et a1.,"An eficient variable neighborhood search heuristic for very large scale vehicle muting problems". *Computers & Operations Research*,34(9),2007,pp.2743-2757.
7. Li FY,Golden B ,Wasil Edward.,"Very large—scale vehicle routing ：new test problems , algorithms, and results". *Computers & Operations Research*,32(5),2005,pp. 1165 -1179.
8. Baldacci R, Mingozzi A, Roberti R.,"New Route Relaxation and Pricing Strategies for the Vehicle Routing Problem". *Operations Research*, 59(5),2014, pp. 1269-1283.
9. Contardo C, Martinelli R.,"A new exact algorithm for the multi-depot vehicle routing problem under capacity and route length constraints". *Discrete Optimization,* 12(1),2014, pp. 129-146.
10. Dell’Amico M, Díaz J C D, Hasle G, et al.,"An Adaptive Iterated Local Search for the Mixed Capacitated General Routing Problem". *Transportation Science*, 50(4),2016, pp. 1-29.
11. Oyola J, Løkketangen A.,"GRASP-ASP: An algorithm for the CVRP with route balancing". *Journal of Heuristics*, 20(4),2014, pp. 361-382.
12. Pawel Sitek.A.,"Hybrid Approach to the Two-Echelon Capacitated Vehicle Routing Problem (2E-CVRP) ". *Recent Advances in Automation, Robotics and Measuring Tech*niques, 38 (3),2014,pp.251-263.
13. Mahvash B, Awasthi A, Chauhan S.,"A column generation based heuristic for the capacitated vehicle routing problem with three-dimensional loading constraints". *International Journal of Production Research*, 55(6),2015, pp. 1730-1747.
14. Wang Chao et al.,"Multi-objective Model and Methods for Capacitated Vehicle Routing Problem Facing Different Target Preference". *Application Research of Computers*, 33(8),2016,pp. 2270-2274.
15. Wang Bo, Lu Yu, Chen Liyun et al., " Study of Modified Shuffled Frog Leaping Algorithm for Solving CVRP". *Application Research of Computers,* 28(12),2011, pp. 4503-4506.
16. Fu Zhengtang, Hu Zhihua, Fei Haiping.,"Hazardous Container Yard Patrol Optimization Research with UAVS". *Journal of Railway Science and Engineering,* 14(11),2017, pp. 2467-2472.
17. Ma Xiaolu, Li Hecheng.,"A New Genetic Algorithm for the Capacity Constraints Vehicle Routing Problem". *Advances in Applied Mathematics*, 03(4), 2014,pp. 222-230.
18. Pang Qi, Ruan Pingnan, Guan Zhiqiang., " Hybrid Artificial Bee Colony Algorithm for Solving CVRP Problem in Coal Logistics". *Modern Management Science*,14(1),2014,pp. 72-74.
19. Li Mingjv, Liang Liping, Lu Yanxia., "A Model of Vehicle Routing Problem Based on Improved Tabu Search Algorithm". *Journal of Highway and Transportation Research and Development,* 34(10),2017, pp. 108-114.