Study on Logistics Distribution Route Optimization Based on Clustering Algorithm and Ant Colony Algorithm

Zhang Bin and Liu Xiao-Jun*

Department of Logistics and Information Management, Zhuhai College of Jilin University, Zhuhai, China

Abstract: Logistics distribution has become the key research aspect to improve the efficiency and reduce the cost of logistics. Based on the survey of current situation and optimization algorithms, a novel optimization scheme is presented in this paper. For a number of distribution sites in a city, firstly K-means clustering algorithm is adopted to examine local distribution centers and their scope. Following this, ant colony algorithm is also used to design the local optimal route inside each scope. The simulation results show that the presented scheme can improve the optimal distribution route compared to pure ant colony algorithm.

Keywords: Ant colony algorithm, Clustering algorithm, Logistics distribution, Optimization.

1. INTRODUCTION

With the rapid development of Internet and e-commerce, the logistics distribution scale of express industry has increased. Taking Nanjing as an example, the accumulative express delivery business volume from January 2013 to November 2013 in Nanjing was 172,000,000 packages and the daily average express delivery business volume was more than 500,000 packages. It was predicted that the annual express delivery business volume in Nanjing in 2015 will reach 450,000,000 packages and the business income will be up to 6,000,000,000 yuan. While the logistics distribution quantity increases rapidly; the consumers have new requirements on the distribution service quality [1]. The proposal of "door to door", "arrival on the same day in the same city", "arrival on the next day in the province" and other service slogans completely re-define the space and time concept of logistics distribution service. The starting point and destination of logistics transportation are scattered at various geographic corners and the logistics arrival time is greatly compressed due to the consumer experience and market competition. The key to improve the logistics efficiency and reduce the logistics cost is to design a reasonable distribution route and compress the distribution service time.

Informatization and intellectualization are the key features of modern logistics that are different from the traditional logistics. Facing the increased number of items, changing distribution destination and compressed distribution time and other requirements besides constrained by other various requirements, it is necessary to conduct in-depth study on the generating algorithm of logistics distribution route by relying on modern information technology to command and dispatch the logistics distribution and transportation operation efficiently and intelligently.

*Address correspondence to this author at the Department of Logistics and Information Management, Zhuhai College of Jilin University, Zhuhai, China; E-mail: lxj02041820@163.com

Based on the study of current situation of logistics and relevant distribution route algorithm, the clustering algorithm combined with the ant colony optimization is presented in this paper. First, K-means clustering algorithm is adopted to get local distribution centers and customer points within their scope, and following this, ant colony algorithm is used to find the optimal distribution route. The simulation results show that the presented scheme can improve the length of the optimal distribution route.

2. ANALYSIS OF LOGISTICS DISTRIBUTION PROBLEM

The transportation cost occupies a high proportion among various costs of logistics and the distribution is an important link in the logistics system. The rationality of distribution scheme imposes great influences on the logistics service level, cost and performance. Vehicle routing problem (VPR) is a core problem in the logistics distribution and one of the research hot topics all over the world and also occupies central position in modern logistics. VPR is generally considered as the distribution center in which the number and model of its vehicles as well as customers' demands are evaluate, to reasonably design a route for the distribution vehicles. Under the circumstance of meeting certain constraint, each vehicle starts from the distribution center and then returns to the distribution center after completing the distribution tasks of each demand point. By optimizing the vehicle scheduling, a certain goal can be achieved on the premise of satisfying the customer demand (such as the shortest driving distance and the fewest distribution vehicles).

VPR is limited by many conditions as follows:

Vehicle capacity limit: the demand of customer served by each vehicle should not be more than its maximum capacity;

② Time window limit: each customer must be served within a certain time scope;

1874-110X/15 2015 Bentham Open

③ The logistics enterprise may have many distribution centers to serve customers;

④ The customer may return some goods to the distribution center;

⑤ The customer can be served by different vehicle;

⑥ The customer number, demand and driving route are determined randomly;

⑦ Serving of customers may be limited.

Since the 1990s, scholars have introduced artificial intelligence for solving VRP, thus generating a large number of heuristic algorithms based on artificial intelligence. The heuristic search algorithm improves the solution of problem by constant local disturbance of current solution from the initial solution of problem until the optimal solution or satisfactory solution is obtained. The common heuristic algorithms are ant colony optimization, simulated annealing algorithm, particle swarm optimization and genetic algorithm [2].

2.1. Ant Colony Optimization

Ant colony optimization (ACO) is a novel bionic evolutionary algorithm for solving the combinational optimization problem inspired by the foraging behavior of ant colony in the nature. It, first put forward by an Italian scholar M.Dorig *et al*., is a new general heuristic algorithm for solving the combinational optimization problem, with strong optimization ability and good nature in solving many difficult combinational optimization problems. It is a random search algorithm [3].

2.2. Simulated Annealing Algorithm

Simulated annealing algorithm (SAA) was first put forward by Metropolis and was first applied to the combinational optimization problem successfully by Kark Patrick in 1983. SAA is derived from the solid annealing principle and its basic idea is to simulate the optimization process of optimization problem by solid annealing process. When the solid is in the minimum energy state, the objective function of optimization problem also reaches the global optimal value. SAA is different from the traditional random search method in terms of search strategy, which not only introduces proper random factor, but also introduces the natural mechanism of solid annealing process. Recently, SAA has been applied to solve the logistics distribution vehicle scheduling problem and some research achievements have also been made.

2.3. Particle Swarm Optimization

Particle swarm optimization (PSO) is an optimization algorithm inspired by the foraging behavior of birds in nature and was put forward by Kennedy and Ebethart *et al*. in 1995. Compared with other optimization algorithm, PSO retains the global search strategy based on population and evaluates individual fitness by cooperation and competition between individuals, thus avoiding complex individual operation. It also adopts the simple speed regulation model and can dynamically trace the current search status by unique memory. With simple principle and strong search ability, this algorithm cannot easily fall into local optimal solution. Now, it draws the attention of many scholars in China.

2.4. Genetic Algorithm

Genetic algorithm (GA) is a self-adaption and global optimization probability search method formed by simulating the evolutionary process of organisms in the nature. Based on the fitness function, GA is an iteration process that restructures the individual structure by genetic operation of individuals in the population, in which the individual with better fitness is realized after evolution to make the solution of problem optimized generation after generation and then gradually reach the optimal solution. Now, GA is not only an important method of solving the complex global optimization problem, but also has been widely used in the specific scientific practice of many research fields. Certain achievements have been made for its application to VRP.

However, current optimization computing methods only consider the shortest route in a certain region. Actually, for the first-tier or second-tier cities in China, if the city is regarded as distribution region then the scheme is applicable to the actual problem. In addition, in the actual distribution process, the weight of different express packages is greatly different to affect the specific design of distribution route. However, most of the existing schemes fail to consider the weight of express packages to be distributed [4].

Considering the clustering features of user, a novel solution combining weighted clustering algorithm and ant colony optimization is presented in this paper, which strikes an average in the shortest distribution route and the fewest distribution time. The basic idea is: regarding the destination as a point on the 2D map, taking the weight of express package as the weight, clustering by using clustering algorithm, calculating many regions and establishing a distribution center in each region. The vehicle starting from the distribution center distributes in the region according to the shortest route calculated by ant colony optimization.

3. CLUSTERING ALGORITHM AND ANT COLONY ALGORITHM

The principle of clustering algorithm and ant colony algorithm used below is introduced in this section.

3.1. Clustering Algorithm

The clustering analysis is based on similarity and the purpose of clustering is to maximize the similarity of objects of the same class and minimize the similarity of objects of different class.

K-means optimization algorithm in the clustering algorithm is adopted in this paper to classify the sample points. K-means optimization algorithm is a clustering algorithm classified according to the function criterion to minimize the clustering criterion function [5, 6]. The clustering criterion function used in this study is the quadratic sum of the distance from each sample point in each class to the clustering center. For K pattern classes, the criterion function is defined as follows:

$$J = \sum_{j=1}^{K} \sum_{i=1}^{N_j} \left\| X_{ij} - Z_j \right\|^2 ,\ X_i \in S_j \qquad (1)$$

Where, S_j represents j clustering set, in which the clustering center is Z_j ; N_j is the sample numbers in j clustering set S_j ; X_{ij} represents i sample in j clustering set. The clustering criterion of K-means optimization algorithm is that the clustering center Z_j should make the criterion function J minimum. To achieve this,

$$\frac{\partial J_j}{\partial Z_j} = 0 \quad (2)$$

It can be obtained combined with equation (1):

$$Z_j = \frac{1}{N_j}\sum_{i=1}^{N_j} X_{ij}, \; X_i \in S_j \quad (3)$$

This equation shows that the clustering center of S_j should choose the mean value of samples of this class.

The working process of K-means optimization process is as follows. Selecting K samples randomly from N pattern samples as the initial clustering centers; allocating the remaining samples to the most similar cluster (represented by the clustering center) respectively according to their similarity (distance) to the clustering centers; calculating the clustering center of each new cluster (mean of objects in this cluster) and constantly repeating this process until the convergence of standard measure function [7]. The calculation steps are as follows:

① Selecting K samples randomly as the initial clustering center $Z_1(1), Z_2(1), \ldots Z_K(1), K < N$. The number in the bracket represents the number of iterative operation times of searching the clustering center.

② Allocating other samples X to one of the K clustering centers according to the principle of minimum distance.

If $\min\{\|X - Z_i(k)\|, i = 1, 2, \ldots, K\} = \|X - Z_i(k)\|$ (4)

So $X \in S_j$

③ Re-calculating the new vector of each clustering center $Z_j(k+1), j = 1, 2, \ldots, K$.

$$Z_j(k+1) = \frac{1}{N_j}\sum_{X_i \in S_j(k)} X_{ij}, j = 1, 2, \ldots, K \quad (5)$$

The mean vector is taken as the new clustering center and the sample mean vector of K clustering center is calculated respectively.

④ If $Z_j(k+1) \neq Z_j(k)$ and $j = 1, 2, \ldots, K$, step (2) is followed by further reclassifying the pattern sample and repeating the iterative computation. If $Z_j(k+1) = Z_j(k)$ and $j = 1, 2, \ldots, K$, the clustering center remains unchanged, and the computation is completed.

3.2. Ant Colony Optimization

Ant colony optimization (ACO) is a probability algorithm for finding the optimal route in the map. ACO was first put forward by Marco Dorigo in his doctoral dissertation in 1992, inspired by the study of ants' behavior of finding the optimal route in the foraging process. According to the observation of Marco Dorigo, the ant chooses its initial starting direction and route randomly under the circumstance of the unknown place of food. The ant releases pheromone (which is a kind of volatile secretion and will disappear gradually with time) while walking [3, 8]. As the optimal road of finding the food is shorter than the other road, the walking times of ants along the road are greater than the other road and the pheromone on this road will be larger. Gradually, more ants are attracted to this road. After a long time, all the ants find their food along the shortest route.

Taking TSP (Travel Salesman Problem) as an example, the basic idea and principles of ant colony optimization are elaborated.

The variables and constants used in the steps are as follows:

m= Number of ants,

η_{ij} =Visibility of edge arc (i, j),

τ_{ij} =Locus strength of edge arc (i, j),

$\Delta\tau_{ij}^k$ =Number of locus pheromone per unit length left by ant k on edge arc (i, j),

Under different $\Delta\tau_{ij}^k$, the different type of ant colony optimization can be formed. Taking basic Ant-Cycle model as an example,

$\Delta\tau_{ij}^k = \begin{cases} Q / Z_k, \\ 0, \end{cases}$ if (i, j) is on the optimal route, Z_k is the objective function value, (6)

P_{ij}^k =Transition probability of ant k, which is directly proportional to $\tau_{ij}^a \cdot \eta_{ij}^\beta$. The renewal equation of locus strength is:

$$\tau_{ij}^{new} = \rho \cdot \tau_{ij}^{old} + \sum_k \tau_{ij}^k \quad (7)$$

Meaning of parameters:

α =Relative importance of locus ($\alpha \geq 0$),

β =Importance of visibility ($\beta \geq 0$),

ρ =Durability of locus ($0 \leq \rho \leq 1$),

Q=Constant that represents the number of locus left by the ants.

Main steps of ant colony optimization are as follows:

Step 1. Consider the iterations Nc as 0, record (Nc←0); initialize τ_{ij} and $\Delta\tau_{ij}$ and put m ants on n peaks;

Step 2. Place the initial starting point of each ant in the current set; for each ant k, move it to the next peak j according to the probability p_{ij}^k and put the peak j in the current set;

Step 3. Calculate the objective function value Z_k of each ant and record the current optimal solution;

Step 4. Modify the locus strength according to the renewal equation;

Step 5. For each edge arc (i, j), set $\Delta\tau_{ij} \leftarrow 0$; $\mathrm{Nc} \leftarrow \mathrm{Nc}+1$;

Step 6. If Nc is less than the preset iterations without degeneration (that is, all the solutions are the same), return to Step 2;

Step 7. Provide the best solution as output.

4. LOGISTICS DISTRIBUTION SCHEME BASED ON CLUSTERING ALGORITHM AND ANT COLONY OPTIMIZATION

Vehicle routing problem (VRP) is a core problem in the logistics distribution optimization. The VRP can be described as: limited by some conditions, and designing the optimal delivery or distribution route from one or more starting points to many different positions [9, 10]. Designing a route set with the minimum cost to:

① Make each vehicle visit each customer once;

② Make all the vehicles initiate from the starting point and return to the starting point;

③ Meet some constraints.

The general constraints are:

① Capacity limit. The freight of a vehicle should not be more than its capacity; this problem is called CVRP.

② Total length limit. The length of any route or time duration should not be more than a constant; this problem is called DVRP.

③ Time window. Customer i must be visited in a certain period of time; this problem is called VRPTW.

4.1. Overall Scheme Design

CVRP is taken as the research object in this paper. In the actual logistics distribution, the geographic features of customer distribution, customer demand, capacity of distribution vehicle and other factors should be considered [11]. The data size of logistics distribution customer is big generally, but the solution of VRP is associated to NP problem and the algorithm solving VCRP is based on exponential algorithm [12]. The presented logistics distribution route design scheme adopts K-means optimization algorithm according to the customer clustering problem to divide the customers into several regions and uses the ant colony optimization to obtain the optimal distribution route set in each region.

Assuming that several customers are known, along with the position coordinate and goods' demand of each customer, the vehicle capacity is fixed and each vehicle initiates from the starting point and returns to the starting point after completing the delivery task of several customer points; suppose each customer is visited only once and the freight distributed by each vehicle is not more than its capacity, it is required to satisfy all the customers' demands and minimize the gross transportation cost.

CVRPs can be described as the problem that __ vehicles starting from the parking lot serve __ customers. Setting d_i as the demand of customer i, c_{ij} is the transportation expense from customer i to customer j, and b_k is the capacity of vehicle k,

$$x_{ij}^k = \begin{cases} 1, \text{Vehicle k continuously serves} \\ \quad \text{customer j after serving customer i;} \\ 0, \text{Others.} \end{cases} \tag{8}$$

The objective of CVRPs is to find the optimal transportation route for the fleet to minimize the transportation cost, that is,

$$J = \min \sum_{k=1}^{m} \sum_{\substack{i,j=0 \\ i \neq j}}^{n} c_{ij} x_{ij}^k \Sigma\, \Sigma \tag{9}$$

And the following constrains should be satisfied:

$$\sum_{j=1}^{n} x_{0j}^k = \sum_{j=1}^{n} x_{j0}^k = 1, k \in (1, 2, \mathrm{K}\ , m)$$
(10)

$$\sum_{k=1}^{m} \sum_{\substack{j=0 \\ j \neq i}}^{n} x_{ij}^k = 1, i \in (1, 2, \mathrm{K}\ , n) \tag{11}$$

$$\sum_{k=1}^{m} \sum_{\substack{i=0 \\ i \neq j}}^{n} x_{ij}^k = 1, j \in (1, 2, \mathrm{K}\ , n) \tag{12}$$

$$\sum_{i=1}^{n} d_i \sum_{\substack{j=0 \\ j \neq i}}^{n} x_{ij}^k \le b_k, k \in (1, 2, \mathrm{K}\ , m) \tag{13}$$

Equation (10) implies that the vehicle must start from the parking lot and return to the parking lot after completing the task. Equation (11) and (12) represent that each customer will and only visited for once. Equation (13) represents that the freight of each vehicle for customers is not more than its capacity.

4.2. Specific Embodiment of Scheme

The processing process of the scheme can be expressed as Fig. (**1**).

① Classify all the samples by K-means optimization algorithm;

② For each pattern class, build the optimal distribution route by ant colony optimization;

(a) Initialize:

Iteration Nc←0 :

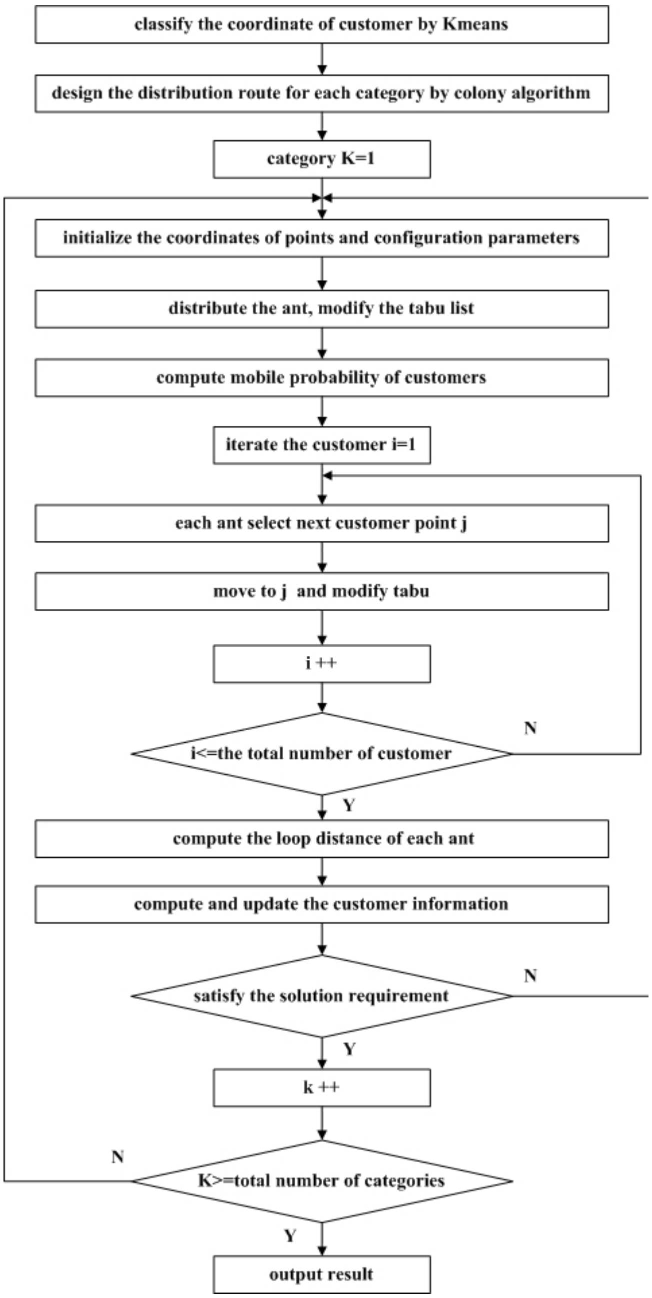

Fig. (1). Processing flow chart.

For each edge arc (i, j): $\tau_{ij} \leftarrow$ constant c; $\Delta\tau_{ij} \leftarrow 0$;

Remaining capacity of vehicle Ca←D;

(b) Considering the initial point into the current set;

(c) For each ant k;

peak J is selected according to the remaining capacity and transition probability p_{ij}^k. The transition probability p_{ij}^k is:

$$p_{ij}^k(t) = \begin{cases} \dfrac{[\tau_{ij}(t)]^a[\eta_{ij}(t)]^\beta}{\sum_{s \in al_k}[\tau_{is}(t)]^a[\eta_{is}(t)]^\beta}, \text{ÈĝÕ } j \in al_k \\ 0, \text{ · ñÚ} \end{cases} \quad (14)$$

Move ant k from the peak i to peak j;

Put peak j into the current set;

Repeat (c) until all the points are considered in the set;

(d) Record the number of ants m←k;

Optimize the route by local search mechanism;

Calculate the objective function value of each ant;

(e) for each ant k←1 to m;

Calculate each edge (i, j): $\Delta\tau_{ij} \leftarrow \Delta\tau_{ij} + \Delta\tau_{ij}^k$;

(f) Locus renewal:

Calculate each edge (i, j): $\tau_{ii} \leftarrow \rho \cdot \tau_{ii} + \Delta\tau_{ii}$;

(g) For each edge (i, j), set $\Delta\tau_{ii} \leftarrow 0$, Nc←Nc+1;

(h) If Nc is less than the preset iterations, return to (b);

(i) Provide the best solution as output.

5. SIMULATION EXPERIMENT

This paper performs simulation as an example for the logistics system optimization problem by using the above algorithm. Assuming that there are 33 customer points in the distribution operation plan of a day in a district of Nanjing, the coordinate points are shown in Fig. (**2**), and the vehicle capacity coefficient W is 9. To simplify the model, the package weight coefficient of each customer point t is 1.

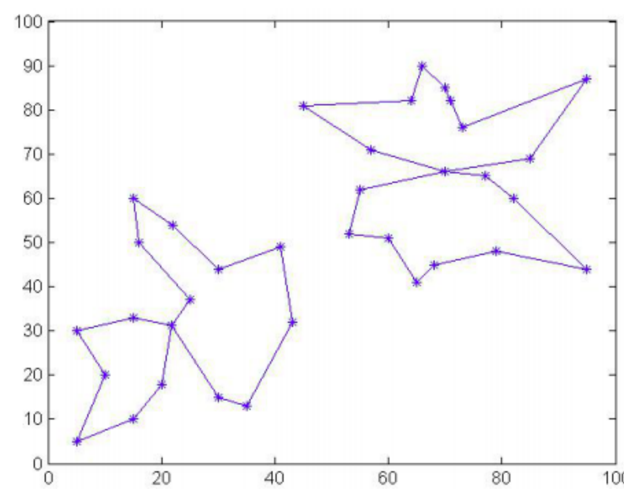

Fig. (2). Distribution route map obtained by the method presented in this paper.

First, the customer points are classified by K-means optimization algorithm, without losing generality. Under the circumstance that K is 2, the clustering center point is calculated, which is the local distribution center.

The points in Fig. (**2**) are the customer coordinate points in category 1 and 2, in which there are 18 customers points in category 1 and 15 customer points in category 2. The two local distribution centers are:

Distribution center 1= (70.0, 66.2);

Distribution center 2= (21.8, 31.3);

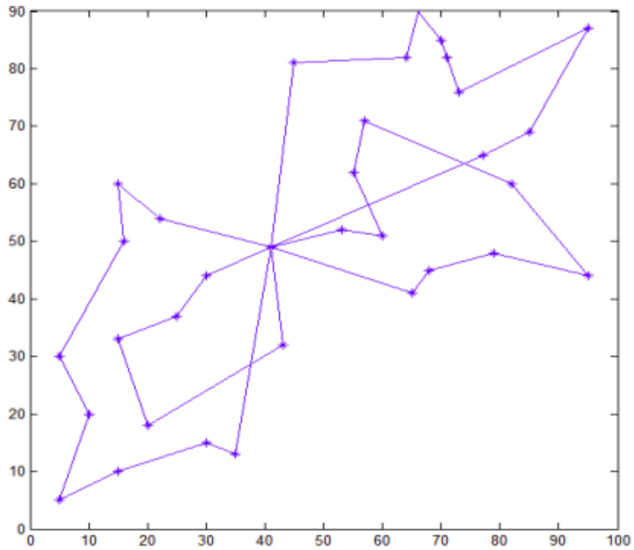

Fig. (3). Distribution route map obtained by ant colony optimization.

Then, the clustering center point is taken as the local distribution center, that is, the distribution vehicle starts from the clustering center point and then returns to it after completing the distribution task. The optimal vehicle distribution route is calculated by ant colony optimization in the two pattern class, in which m=60; Alpha=1; Beta=1; gama=2; Rho=0.15; NC_max=50; Q=15; W=9; qq=0.05.

After iteration, the optimal route map obtained is shown in Fig. (**2**).

The current optimal distribution length of category 1 and 2 is 244.8 and 206.4 respectively, 451.2 in total.

If the ant colony optimization algorithm is used, the optimal route map obtained by iteration is shown in Fig. (**3**), in which the optimal route length is 572.8 in total, which is 121.6 longer than the optimal route in Fig. (**2**). The method presented in this paper greatly improves the distribution length compared with the ant colony optimization.

The above algorithm is used in the logistics distribution planning of Nanjing. The data acquisition and analysis are used to guide the layout of local distribution center and the cloud computing technology is used to command and schedule the distribution operation plan.

CONCLUSION

The logistics distribution is an important link in the logistics system. The rationality of distribution scheme imposes great influence on the logistics service level, cost and performance and the logistics distribution route of vehicle routing problem and selection have become problems concerned by the logistics industry. First, K-means clustering algorithm is adopted to obtain local distribution centers and customer points within their scope, and following this, ant colony algorithm is used to find the optimal distribution route. The simulation experiment shows that the presented scheme can improve the length of the optimal distribution route.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Nanjing Daily, Nanjing Releasing, "*The First Development Plan of Express Service Industry*," [EB/OL]. http://www.nanjing.gov.cn/njgov/xxzx/mjxw/201401/t20140122_2454340.html.

[2] W. Chen, and S. Zheng, "Research on distribution route planning based on ant colony optimization," *Logistics Sci-Tech*, vol. 36, no. 7, pp. 45-47, 2013.

[3] B. Yang, "Influence of route optimization in express industry and customer service-based on ant colony optimization," *Logistics Engineering and Management*, vol. 35, no. 9, pp. 82-83, 2013.

[4] A. Ye, "Routing optimization problem of tobacco distribution vehicle based on genetic algorithm," *Computer Systems & Applications*, vol. 20, no. 4, pp. 241-244. 2011.

[5] Z. He, "Research on enterprise distribution route problem based on genetic algorithm," *Logistics Engineering and Management*, vol. 35, no. 11, pp. 96-97, 2013.

[6] J. Liu, Q. Zhou, G. Zou, "Research on logistics distribution route optimization of agricultural products based on partical swarm optimization," *Logistics Engineering and Management*, vol. 35, no. 1, pp. 72-75, 2013.

[7] Q. Gu, "Location selection algorithm of multiple logistics distribution center based on ant colony optimization," *Systems Engineering-Theory & Practice*, vol. 26, no. 4, pp. 120-124, 2006.

[8] K. Xie. "*Research and Application of Clustering Algorithm in Logistics Distribution System*," Zhejiang University: Hangzhou, 2006.

[9] M. Qin, D. Li, and C. Hao, "*Guide of Pattern Recognition*," Tsinghua University Press: Beijing: 2009, pp. 13-30.

[10] D. Marco, S. Thomas, Translated by Z. Jun "*Ant Colony Optimization*," Tsinghua University Press: Beijing, 2007, pp. 32-72.

[11] L. Ma, G. Zhu, and A. Ning, "*Ant Colony Optimization*," Science Press: Beijing, 2008, pp. 15-20.

Received: June 10, 2015 Revised: July 29, 2015 Accepted: August 15, 2015