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 to improve 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 lot of distribution sites in a city, firstly K-means clustering algorithm is adopted to get local distribution centers and their scope, and then ant colony algorithm is 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 increases. 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 is 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. 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-defines the space and time concept of logistics distribution service. The starting point and destination of logistics transportation are scattered at the 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 the 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 increasing items, changing distribution destination and compressed distribution time and other requirements and constrained by other various requirements, it is necessary to conduct in-depth study of the generating algorithm of logistics distribution route by relying on modern information technology so as to command and dispatch the logistics distribution and transportation operation efficiently and intelligently.

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 then 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 the various costs of logistics and the distribution is an important link in the logistics system. The reasonableness 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 focuses all over the world and also occupies the central position in the modern logistics. VPR is generally described as: the distribution center and the number and model of its vehicles as well as position of a series of customer demand points and customer demand are known, a route for the distribution vehicles is required to be designed reasonably. 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:

① 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;
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 occur randomly;

⑦ Serving consequence of customers may be limited.

Since the 1990s, scholars have introduced artificial intelligence into the solving of VRP, thus generating a large number of heuristic algorithms based on artificial intelligence. The heuristic search algorithm improves the solution of problem by the 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, partial swarm optimization and genetic algorithm.

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 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 the solving of many difficult combinational optimization problems. It is a random search algorithm.

2.2. Simulated Annealing Algorithm

Simulated annealing algorithm (SAA) was first put forward by Metropolis and then 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 does not only introduce the proper random factor, but also introduce 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 been made.

2.3. Partial Swarm Optimization

Partial swarm optimization (PSO) is an optimization algorithm inspired by the foraging behavior of birds in the 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 the individual fitness by the cooperation and competition between individuals, thus avoiding the complex individual operation. It also adopts the simple speed regulation model and can dynamically trace the current search status by the 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 and better fitness occurs after evolution to make the solution of problem optimized generation after generation and then gradually reach the optimal solution. Now, GA does not only become an important method of solving the complex global optimization problem, but also have been widely used in the specific scientific practice of many research fields. Certain achievements have been made for its application to VRP.

However, the current optimization computing methods just consider the shortest route in a certain region. Actually, for the first-tier or second-tier cities in China, it is very big if the city is regarded as a distribution region and 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 so as 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.

Considering the clustering features of user, a novel solution combined 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: regard the destination as a point on the 2D map, take the weight of express package as the weight, cluster by using clustering algorithm, calculate many regions and establish 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 the 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 classifying according to the function criterion to minimize the clustering criterion function. The clustering criterion function used here 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:
\[ J = \sum_{j=1}^{K} \sum_{i=1}^{N_j} \| x_i - z_j \|^2 \quad x_i \in S_j \]

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 \( S_j \); \( x_i \) represents i sample in j clustering set. The clustering criterion of K-means optimization algorithm is: the clustering center \( z_j \) should make the criterion function \( J \) minimum. To achieve this,

\[
\frac{\partial J}{\partial z_j} = 0
\]

It can be obtained combined with equation (1):

\[
Z_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_i, \quad x_i \in S_j
\]

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:

1. Choose K samples randomly as the initial clustering centers; allocate the remaining samples to the most similar cluster (represented by the clustering center) respectively according to their similarity (distance) to the clustering centers; then calculate the clustering center of each new cluster (mean of objects in this cluster); constantly repeat this process until the convergence of standard measure function. The calculation steps are as follows:

   1. Chose 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 looking for the clustering center.

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

   \[
   \text{If } \min \| x - z_i(k) \|, i = 1, 2, \ldots, K = \| x - z_i(k) \|
   \]

   So \( x \in S_j \)

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

   \[
   Z_j(k+1) = \frac{1}{N_j} \sum_{x_i \in S_j(k)} x_i, j = 1, 2, \ldots, K
   \]

   Take the mean vector as the new clustering center and then calculate the sample mean vector of K clustering center respectively.

   4. If \( Z_j(k+1) \neq Z_j(k) \) and \( j = 1, 2, \ldots, K \), return to step (2), reclassify the pattern sample and repeat the iterative computation; if \( Z_j(k+1) = Z_j(k) \) and \( j = 1, 2, \ldots, K \), the clustering center remains unchanged, 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 their initial starting direction and route randomly under the circumstance of not knowing the place of food. The ant releases pheromone (which is a kind of volatile secretion and will disappear gradually with the time) during walking. As the optimal road of finding the food is shorter than other road, the walking times of ant along the road are more than 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,
- \( \eta_{ij} \) = Visibility of edge arc \((i, j)\),
- \( \tau_{ij} \) = Locus strength of edge arc \((i, j)\),
- \( \Delta \tau_{ij} \) = Number of locus pheromone per unit length left by ant \( k \) on edge arc \((i, j)\).

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

\[
\Delta \tau_{ij} = \begin{cases} \frac{Q}{Z_{ij}}, & \text{if } (i, j) \text{ is on the optimal route}, \\ 0, & \text{otherwise} \end{cases}
\]

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

\[
\tau_{ij} ^{new} = \rho \cdot \tau_{ij} ^{old} + \sum_k \tau_{ij} ^k
\]

Meaning of parameters:
- \( \alpha \) = Relative importance of locus \( (\alpha \geq 0) \),
- \( \beta \) = Importance of visibility \( (\beta \geq 0) \),
- \( \rho \) = 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. Put the iterations \( Nc \) as 0, record \((Nc \leftarrow 0)\); initialize \( \tau_{ij} \) and \( \Delta \tau_{ij} \) and put \( m \) ants on \( n \) peaks;
Step 2. Put 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 \( Z_i \) 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 \( \Delta \tau_{ij} \leftarrow 0 \); \( \text{Nc} \leftarrow \text{Nc} + 1 \);

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

Step 7. Output the best solution.

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, design the optimal delivery or distribution route from one or more starting points to many different positions. Design a route set with the minimum cost to:

1. Make each vehicle visit each customer once;
2. Make all the vehicles start 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 length 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. The data size of logistics distribution customer is big generally, but the solving of VRP belongs to NP problem and the algorithm solving VCRP belongs to exponential algorithm. The presented logistics distribution route design scheme adopt K-means optimization algorithm according to the customer clustering problem to divide the customers into several regions and then use the ant colony optimization to get the optimal distribution route set in each region.

Suppose that several customers are known, the position coordinate and goods demand of each customer are known, the vehicle capacity is fixed and each vehicle starts from the starting point and then returns to the starting point after completing the delivery task of several customer points; suppose each customer is just visited once and the freight distributed by each vehicle cannot be more than its capacity, it is required to satisfy all the customer demand and minimize the gross transportation cost.

CVRPs can be described as the problem that \( n \) vehicles starting from the parking lot serve \( m \) customers. Set \( d_{ij} \) is 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 \( k \),

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

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

\[
J = \min \sum_{k=1}^{m} \sum_{i=1}^{n} c_{ij} x_{ij}^k \sum \sum \tag{9}
\]

And the following constrains should be satisfied:

\[
\sum_{j=1}^{n} x_{ij}^k = \sum_{j=0}^{n} x_{j0}^k = 1, k \in (1, 2, \ldots, m) \tag{10}
\]

\[
\sum_{k=1}^{m} \sum_{j=0}^{n} x_{ij}^k = 1, i \in (1, 2, \ldots, n) \tag{11}
\]

\[
\sum_{k=1}^{m} \sum_{j=0}^{n} x_{ij}^k = 1, j \in (1, 2, \ldots, n) \tag{12}
\]

\[
\sum_{i=1}^{n} \sum_{j=0}^{n} d_{ij} x_{ij}^k \leq b_k, k \in (1, 2, \ldots, m) \tag{13}
\]

Equation (10) represents that the vehicle must start from the parking lot and then return to the parking lot after completing task; equation (11) and (12) represent that each customer will and can only be visited once; equation (13) represents that the freight of each vehicle for customers should not be 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 \( \text{Nc} \leftarrow 0 \).
For each edge arc (i, j): \( \tau_{ij} \leftarrow \text{constant} \); \( \Delta \tau_{ij} \leftarrow 0 \);

Remaining capacity of vehicle \( C_a \leftarrow D \);

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

(c) For each ant \( k \):

Choose peak \( J \) according to the remaining capacity and transition probability \( p^k_{ij} \); the transition probability \( p^k_{ij} \) is:

\[
p^k_{ij}(t) = \begin{cases} 
\frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}(t)]^\beta}{\sum_{i,j} [\tau_{ij}(t)]^\alpha [\eta_{ij}(t)]^\beta} & \text{if } \langle i,j \rangle \in al_k \\
0 & \text{otherwise} 
\end{cases}
\]  

(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 put in the set;

(d) Record number of ants \( m \leftarrow k \);

Optimize the route by local search mechanism;

Calculate the objective function value of each ant;

(e) For each ant \( k \leftarrow 1 \) to \( m \);

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

(f) Locus renewal:

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

(g) For each edge \( (i, j) \), set \( \Delta \tau_{ij} \leftarrow 0 \), \( N_c \leftarrow N_c + 1 \);

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

(i) Output the best solution.

5. SIMULATION EXPERIMENT

Solve a simulation example in the logistics system optimization problem by using the above algorithm. Suppose 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), the vehicle capacity coefficient \( W \) is 9. To simplify the model, the package weight coefficient of each customer point \( t \) is 1.

First, classify the customer points by K-means optimization algorithm, without loss of generality. Under the circumstance that \( K \) is 2, calculate the clustering center point, that is, 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 classes, in which $m=60; \text{Alpha}=1; \text{Beta}=1; \text{gama}=2; \text{Rho}=0.15; \text{NC}_{\text{max}}=50; Q=15; W=9; \text{qq}=0.05$.

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 just used, the optimal route map obtained by iteration is shown in Fig. (3), in which the optimal route length is 572.8 in total, 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 gradually used in the logistics distribution planning of Nanjing. The data acquisition and analysis are used to guide the layout of the 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 reasonableness of distribution scheme imposes great influences on the logistics service level, cost and performance and the logistics distribution route of vehicle routing problem and selection has become one of the problems concerned by the logistics industry. First, K-means clustering algorithm is adopted to get local distribution centers and customer points within their scope, and then 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