

Received 2 September 2022, accepted 12 September 2022, date of publication 15 September 2022,
date of current version 26 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3206947

RESEARCH ARTICLE

Two-Stage Vehicle Routing Optimization for Logistics Distribution Based on HSA-HGBS Algorithm

QI SUN¹, HAIFEI ZHANG^{2,3}, (Graduate Student Member, IEEE), AND JIANWU DANG⁴

¹Key Laboratory of Railway Industry of BIM Engineering and Intelligent for Electric Power, Traction Power Supply, Communication and Signaling, Lanzhou Jiaotong University, Lanzhou 730070, China

²School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, Jiangsu 214122, China

³Key Laboratory of Advanced Process Control for Light Industry, Ministry of Education, Jiangnan University, Wuxi, Jiangsu 214122, China

⁴School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

Corresponding author: Haifei Zhang (1256121964@qq.com)


This work was supported in part by the Opening Foundation of Key Laboratory of Railway Industry of Building Information Modeling (BIM) Engineering and Intelligent for Electric Power, Traction Power Supply, Communication and Signaling under Grant BIMKF-2021-01; and in part by the Opening Foundation of State Key Laboratory of Rail Transit Engineering Informatization (FSDI) under Grant SKLKZ22-06.

ABSTRACT Aiming at the problems of complex urban road network, low efficiency of logistics distribution, and the difficulty of large-scale logistics distribution area division and routing planning, this paper proposes a two-stage logistics distribution vehicle routing optimization (VRP) method based on the establishment of a multi-factor complex road network constrained logistics distribution mathematical model. Considering the complex traffic elements and road network topological structure in logistics and distribution, in the first stage, a heuristic simulated annealing (HSA) distribution region partitioning algorithm is proposed with the objective of balancing vehicle task load to divide the urban logistics distribution network under complex road networks, so as to reduce the region scale and path search cost. In the second stage of route decision making, aiming at minimizing the total cost of logistics distribution, combining the VRP problem with complex road network conditions, a heuristic path search method combined with complex road network model constraints is proposed. In this stage, a hybrid genetic beam search(HGBS) algorithm is used to plan the path nodes, reduce the randomness of the model in the initial search for paths by heuristic genetic algorithms, then combine with Beam Search methods to reduce the space and time used for the search, and use optimization algorithms to improve the accuracy of independent sub-region routing optimization and the rationality of overall physical distribution route selection. Finally, the proposed method is validated in this paper with two practical cases. The experimental results show that the two-stage decision-making algorithm proposed in this paper has certain advantages in partitioning schemes, minimizing total cost and iteration times. Through comparison, the optimization ability of this method for logistics distribution networks is proved.

INDEX TERMS Vehicle routing optimization, complex road network, two-stage algorithm, heuristic simulated annealing, hybrid genetic beam search.

I. INTRODUCTION

With the rapid development of online shopping and the modern logistics industry, logistics distribution has become an important link connecting producers and consumers, and it plays an increasingly important role in the whole supply chain. Optimizing the vehicle routing of urban logistics

The associate editor coordinating the review of this manuscript and approving it for publication was Vivek Kumar Sehgal .

distribution is of great significance to reduce logistics operation costs and improve customer service satisfaction. Therefore, the research of vehicle routing optimization (VRP) for multi-factor complex road network constraints has been paid more and more attention.

Since Dantzig and Ramser [1] proposed the truck scheduling problem, researchers have been studying the relationship between vehicle routing planning and delivery planning. It is considered a typical case of VRP, involving the distribution

of goods from central warehouses to geographically dispersed customers. By planning a reasonable driving route, the cost (such as the shortest distance and the least time) can be minimized. At this period, some scholars have studied many external factors that affect vehicle routing planning in logistics distribution. In the study of constraints related to vehicle routing planning, Wagner [2] studied the relationship between logistics region partitioning and transportation through empirical analysis in the case of a whole logistics area in Hamburg, Germany, and concluded that traffic flow would affect the distribution of the logistics region. This kind of direct path planning for the entire logistics network can be called level-1-based VRP.

Logistics distribution can also divide the entire logistics distribution network into several smaller sub regions, and further study the vehicle routing problem in each sub region. This method can be called two-level VRP, where, Christofides [3] meshed the distribution region into square areas of unequal sizes and aggregated each square area into distribution units according to the historical path connection frequency of different square areas, so as to divide the distribution area and optimize the vehicle path in each area. Besides, Wang *et al.* [4] proposed a hybrid algorithm based on extended particle swarm optimization and genetic algorithm (EPSO-GA). By establishing a two-level logistics distribution system, the efficient urban logistics distribution can be realized under the minimization of logistics distribution cost. Based on the above research, it can be seen that traffic flow, regional scale, distribution cost, and other factors have become the key factors affecting the quality of vehicle routing planning, which directly affect the efficiency, cost, capacity, and service level of logistics distribution. In addition, the traffic elements and topological structure of urban roads also affect the operation of the entire logistics distribution system. Because it is very complex to establish a VRP mathematical model considering the complex road network, there are fewer studies considering the impact of the complexity of urban roads on urban logistics distribution.

The partitioning of logistics distribution region is a combinatorial optimization problem with multiple constraints and multi-objective decision. In recent years, research on solving the optimal vehicle routing method is emerging in endlessly. Most of the methods need to establish mathematical models to complete the vehicle path optimization by defining different types of variables, constraint functions and objective functions. The commonly used methods mainly include exact algorithms and heuristics [5], [6], [7], [8], [9]. Among them, exact search algorithms mainly focus on the branch and bound method, the branch cutting method, sequence generation method and dynamic programming method etc [10], [11]. Because VRP is a NP-hard problem, it will consume too much computing power and storage space when using exact algorithms to optimize, which limits the accuracy of the optimal vehicle path. This method can only be applied to small-scale VRP solution.

With the gradual increase of the problem scale, some scholars propose to use heuristic search algorithm to solve the vehicle path planning problem. General computational intelligence heuristic search algorithms are divided into ten types: biological based, social based, chemical based, physical based, music based, mathematics based, sports based, population-based, plant-based and water based [8], [9]. The heuristic algorithm is based on the optimization algorithm. Its basic idea is to give a feasible solution to the combinatorial optimization problem within an acceptable range. In the VRP Problem, the heuristic search algorithms mainly used in VRP problems include evolutionary algorithm [12], particle swarm optimization algorithm [13], ant colony algorithm [14], genetic algorithm [15], intelligent water drop, tabu search [7] and their improvement types [16], [17], [18]. Compared with the exact search algorithm, the heuristic search algorithm has better robustness and feasibility when dealing with large-scale VRP problems.

Based on the existing research, both exact search algorithm and intelligent heuristic search algorithm can be used to solve VRP and related problems. Exact algorithms can find the optimal solution for the problem. However, it is highly dependent on the solution space, the number of constraints and the number of decision variables in the problem model, and cannot provide a general solution strategy for different types of variables, objectives and constraints [7], [9]. However, when the scale of the problem becomes larger, there will be a “combination explosion” phenomenon in exact algorithm that will consume too much computing power and storage space. By designing the heuristic function, heuristic algorithms can get the optimal solution to a search problem in a very short time. For the NP problem, it can also get a better solution in polynomial time. Heuristic algorithms can further improve the accuracy of vehicle routing. The classification and advantages and disadvantages of vehicle routing optimization methods for logistics distribution are shown in Table 1.

At present, there are three kinds of data used in VRP research. The first is the standard Solomon data-set [19], which is limited to dozens of points. However, in the actual distribution process, especially in industries closely related to daily life, such as garbage collection, milk collection and distribution, cigarette distribution, and so on, the customer group consists of residents or retailers distributed in all corners of the city, and the scale of the problem to be solved is often in the order of 100 or 1000. Whether exact algorithms or heuristic algorithms are directly used to solve large-scale problems, they have limitations. The second is the data set based on experimental simulation. Compared with the standard data set, the scale of this kind of data will expand. For example, Li *et al.* [21], Mester and Bräysy [22], and Accorsi and Vigo [23] obtained a large-scale CVRP example with 1200 customer points according to the example generator designed by themselves, and designed effective solution algorithms such as simulated annealing and variable neighborhood search. Duhamel *et al.* [24] presented a VRP

example using a two-stage algorithm to deal with large-scale and multi-vehicle models. Vonolfen *et al.* [25] gave an example of VRP with a time window processing of 1000 points. However, the scale of these computer-generated experimental simulation examples was still less than that of the actual distribution cases, and the constraints caused by the complex geographical environment and special industrial requirements that the actual distribution needs to face cannot be reflected in the standard examples. In this regard, Wasil [26] and Cheong *et al.* [27] studied the distribution problem of the beverage industry and combined it with the two-stage algorithm to realize the distribution path optimization. Beasley and Christofides [28] studied the example of a large mail order company in the UK that accepts orders by telephone or mail and then delivers them to customers' homes. They all adopted the customer region partition method based on postal code, and realized the distribution route selection through the customer region partition based on postal code. These studies put forward many effective strategies for standard examples, simulation examples, and practical cases, which have important guiding significance for the study of large-scale VRP in actual distribution.

The urban road network has complexity, which is mainly reflected in detailed and complex traffic elements and complex topology (see Section 2.3). Complex traffic elements mainly include traffic lights, accidents, peak hours, etc., as well as the restrictions of urban road intersections and road sections. The complex urban road network topology is mainly reflected in the connection path between two points, including multiple choices. This road network structure affects the operation of the whole logistics distribution system. At present, in a logistics distribution system, the location and quantity of logistics centers are usually determined. Therefore, the two most important links in large-scale urban logistics distribution under the condition of a complex road network are the reasonable partitioning of distribution regions and the optimization of vehicle routing. The partitioning of logistics distribution regions can reduce the burden of large-scale logistics distribution and provide convenience for urban logistics planning. Vehicle routing optimization can alleviate urban congestion, improve mobility, and reduce pollution. Based on previous studies, it is of great practical significance to solve the vehicle routing optimization problem of large-scale logistics distribution facing the constraints of multi-factors and complex road networks. The main contributions to this paper are as follows:

- (1) Aiming at the problem of complex road network composed of detailed and complex traffic elements and the topological structure of road network in urban logistics distribution, the mathematical model of urban complex road network is established.
- (2) Aiming at the vehicle routing optimization problem considering complex road networks, a two-stage decision-making algorithm is established to improve the accuracy and timeliness of vehicle routing in

logistics distribution. In the first stage, aiming at balancing the vehicle task load, the overall distribution region is divided into independent sub-regions by simulated annealing algorithms, so as to reduce the region scale and path optimization range, reduce the path search cost, and improve the path search effect.

- (3) In the second stage of route decision-making, aiming at minimizing the total cost of logistics distribution and considering the complex road network constraints in each sub-regions after the partition of the regional road network, a hybrid genetic beam search algorithm is proposed to realize vehicle routing optimization and enhance the accuracy of independent sub-region routing optimization and the rationality of overall physical allocation routing selection.

II. VRP MATHEMATICAL MODEL CONSIDERING COMPLEX ROAD NETWORK

A. THE DIFFERENCE BETWEEN VRP CONSIDERING COMPLEX ROAD NETWORK AND TRADITIONAL PROBLEMS

Traditional VRP can be defined as [1]: (1) Multiple customers need transportation services at the same time, and multiple vehicles are required to solve customer demand problems. (2) Each customer can only be visited once by one vehicle. (3) All vehicles start from the depots and finally return to the depots. (4) All vehicles must meet the loading capacity constraints. Under the above constraints, reasonably arrange the distribution lines to minimize the total distribution distance and shorten the distribution time.

The precondition for solving the traditional VRP is that the location of the logistics distribution center, the location of the customer point, and the shortest path between any two customer points are known. On this basis, the customer is assigned to different vehicles, and the customer access sequence is arranged for each vehicle, so as to determine the problem solution.

The search traversal network graph of traditional VRP is relatively simple, which is an undirected graph with customers as nodes and the shortest known path as the edge. The search network graph of the VRP model considering complex road networks is a directed graph composed of path nodes (refers to the intersection of two roads) [27], customer nodes, and paths, with additional consideration of actual road traffic restrictions (such as one-way driving and motor vehicles). Therefore, the VRP model considering complex road networks is more practical.

B. ROAD NETWORK CONSTRAINTS AND CUSTOMER DEMAND CONSTRUCTION

In logistics distribution, urban road conditions (mainly reflected in the complexity of urban road networks) affect the operation of the whole distribution system. The complex road network basically reflects the road network structure of the real city. Its complexity is mainly reflected in the

TABLE 1. Classification of exact search algorithm and heuristic search algorithm.

Classification	VRP solution algorithms	Advantages	Disadvantages	Applicable scale
Exact search algorithm	branch and bound method	The optimal solution can be found without evaluating all possible feasible solutions.	The solution speed is slow and the applicability is poor.	Small scale VRP
	sequence generation method	It can solve large-scale linear programming problems with high speed.	Only linear relaxation problems can be solved, and the solving speed of subproblems is slow.	Relatively large-scale
	dynamic programming method	Practical knowledge and experience can be used to improve the solution speed.	There is no unified standard model and numerical method, and there is a dimension disaster when solving the problem.	Small scale
Heuristic search algorithm	PSO	It is easy to realize and the convergence speed is fast.	As the number of iterations increases, the diversity of particles is gradually lost, resulting in the algorithm easily falling into local extremum.	Large scale
	ACO	It has strong robustness and inherent parallelism.	Slow convergence speed, easy to fall into local optimization, and lack of initial pheromones	Large scale
	GA	Taking the objective function value as the search information directly has the characteristics of group search.	The crossover and mutation have randomness, and the algorithm is prone to slow convergence.	Large scale
	Tabu Search	Fast search speed and strong local search ability.	The global searching ability is weak.	Large scale VRP
	water based	It has good robustness and strong global optimization ability.	The movement of water droplets is prone to stagnation and blockage, and it is easy to fall into local extremum and converge slowly.	Large scale

detailed and complex traffic elements and complex topology. Specifically, there may be many connection paths between two points. The traffic network data in the process of logistics distribution and transportation belongs to spatial information. Before modeling, it is necessary to grid the urban logistics distribution region according to the road network.

The model of the urban complex road network can be expressed as:

$$G = (V(G), A(G), P(G)) \tag{1}$$

$$V(G) = \{v_1, v_2, \dots, v_n\} \tag{2}$$

$$A(G) = \{R(v_i, v_j) | R(v_i, v_j) = 1, i, j = 1, 2, 3, \dots, n\} \tag{3}$$

$$\forall i = j, R(v_i, v_j) = 0 \tag{4}$$

$$v_i = v_i(x, y) \tag{5}$$

$$P(G) = \{P(v_i, v_j) | P(v_i, v_j) = (P_1 \times U_2) / (P_2 \times U_1) \text{ and } (v_i, v_j) = 1, i, j = 1, 2, 3, \dots, n\} \tag{6}$$

Formula (1) represents the road network G consists of point set $V(G)$, directed path set $A(G)$ and road congestion information $P(G)$. In formula (2), $V(G)$ is the set of all path nodes, v_i is the i^{th} path node, and n is the number of path nodes. Formula (3) represents that the elements of $A(G)$ are

composed of path matrix $R(v_i, v_j)$. Formula (4) represents that for the path node itself $R(v_i, v_j) = 0$. Formula (5) represents the coordinate of path node of v_i . Formula (6) is the description of road congestion: The impact of natural conditions and road grade on road capacity is evaluated, import formula $P = (P_1 \times U_2) / (P_2 \times U_1)$, (P_1 and U_1 are the average fuel consumption and speed of national roads in plain regions. P_2 and U_2 are the corresponding parameters of a certain class of highway in a certain place), and then, in formula (6), $P(G)$ indicates that when path nodes v_i to v_j are passable, road congestion information $P(v_i, v_j)$ including all path nodes exists.

The demand information of the customer distribution center can be described as:

$$C = \{c_1, c_2, c_3, \dots, c_d, \dots, c_s\} \tag{7}$$

$$c_d = c_d(\bar{x}, \bar{y}, R(v_u, v_p), q_d) \tag{8}$$

$$Q = \sum_{d=1}^S q_d \tag{9}$$

$$J = \{J_1, J_2, J_3, \dots, J_s\} \tag{10}$$

In formula (7), the customer distribution center demand is represented by set C , where, c_d is the demand information of the d^{th} customer distribution node J_d , and S is the total number of customer distribution points, $d = 1, 2, 3, \dots, S$.

In formula (8), \bar{x} and \bar{y} represents the coordinate of the d^{th} customer distribution node J_d , where $R(v_u, v_p) = 1$ also indicates that the customer distribution node J_d is on the traffic side where v_u points to v_p (that is, there is a path between any two path nodes). q_d represents the distribution and delivery demand of all customers at the d customer distribution node J_d . v_u represents the average speed of all path nodes. Formula (9) represents the sum of the demand of all customers for distribution and delivery. Formula (10) represents the set of customer distribution nodes.

This paper uses a two-stage decision-making method to solve the problems of distribution region partitioning and vehicle distribution path planning. In the first stage, the problem of vehicle distribution region partitioning is solved first. Customers are scattered throughout the city, and the location of the distribution center is generally fixed. Customers are mainly clustered around the distribution center as the central point in logistics distribution. In addition, due to the insufficient utilization of vehicle load capacity caused by uneven vehicle task load and the increase in maintenance cost caused by vehicle selection differences caused by uneven task load, the goal of balancing vehicle task load is adopted in the first stage decision-making:

$$\min G = \sum_{m=1}^N (Q_m - \bar{Q})^2 \quad (11)$$

where, $\bar{Q} = Q/N$ represents the average of vehicle capacity, Q_m is the freight volume of vehicle m . Equation (10) represents minimizing the vehicle operating variance during region partitioning, where Q represents the sum of the demand of all customers for distribution and delivery, N represents the total number of vehicles.

In the second stage of decision-making, a better heuristic algorithm is used for path planning. Through actual examples to prove the proposed algorithm has the ability of path optimization and can reduce the total distribution cost. On the whole, the two-stage decision-making takes into account the optimization of distribution costs and benefits on the premise of ensuring that customers are served.

C. MAPPING DESCRIPTION OF COMPLEX ROAD NETWORK

(1) When the customer distribution point is mapped to the road network node: When the customer distribution node c_d coincides with the road network node v_i , then the path search to the customer distribution node c_d is equivalent to the path search to the road network node v_i .

(2) When the customer distribution center is a certain distance from the main road (e.g. in the inner center of the community), it is processed according to the actual path from the trunk road to the customer distribution point. A path node is added at the point where the trunk road enters the customer point. The distance from the customer to the trunk road is the distance from the customer to the path node, which is closer to the actual situation.

(3) When the customer distribution center is just beside the street, it is necessary to consider that the two customer distribution centers are distributed on different sides of the two-way traffic road and there is an isolation belt in the center of the road. At this time, the ideal shortest distance between the two customer distribution centers should be the driving distance of the vehicle entering the node of the road network first and then turning back to the customer distribution center on the other side.

(4) Vehicle transportation path description: the vehicle transportation path is described by the arrangement of path nodes and customer distribution points, indicating the path nodes and customer distribution points that the vehicle passes through in turn.

III. DISTRIBUTION REGION PARTITIONING

The partitioning of the logistics distribution region is a complex and comprehensive problem that needs to consider many factors. For many years, it has been a research hotspot for scholars, involving logistics regional segmentation, vehicle scheduling, vehicle optimal combination, facility scale, distribution time, and distribution cost. These problems will directly affect the efficiency, cost, capacity, and service level of logistics distribution regional planning.

The problem of logistics distribution region partitioning is to determine a set of large and small car distribution schemes and then determine the selection of transfer stations and the attribution division of distribution units, so as to minimize the total cost of the objective function. The schematic diagram before and after region partitioning is shown in Figure 1.

The logistics distribution partitioning considering of complex road network can more easily deal with all links of logistics distribution, and carry out effective management and decision-making analysis on the problems involved, so as to meet the requirements of modern logistics and help logistics distribution enterprises make effective use of existing resources, reduce consumption and improve efficiency.

In the first stage, the heuristic simulated annealing algorithm is used to divide the vehicle distribution region. Firstly, the heuristic simulated annealing algorithm is used to generate the initial partition decomposition: In the first stage, heuristic simulated annealing algorithm is used to divide the vehicle distribution region. First, the heuristic SA algorithm is used to generate the initial partition decomposition:

(1) Set initial value: $m = 0$, $kn = 0$, set safe radius $sr = \sqrt{(\max(x) - \min(x))^2 + (\max(y) - \min(y))^2}$, where, kn represents the number of iterations, $\max(x)$ represents the maximum value of abscissa of all customer distribution points c , and other meanings are similar, turn to (2).

(2) Whether m is less than or equal to the limited number of vehicles N , if so, proceed to (3); Otherwise, it indicates that the division is completed, the algorithm ends, and turn to (6) to evaluate the solution.

(3) Randomly generate the center coordinate $AC_m(x_m, y_m)$ of the distribution region of vehicle m . $x_m = rand \times (\max(x) - \min(x))$, $rand$ represents a random number between (0, 1),

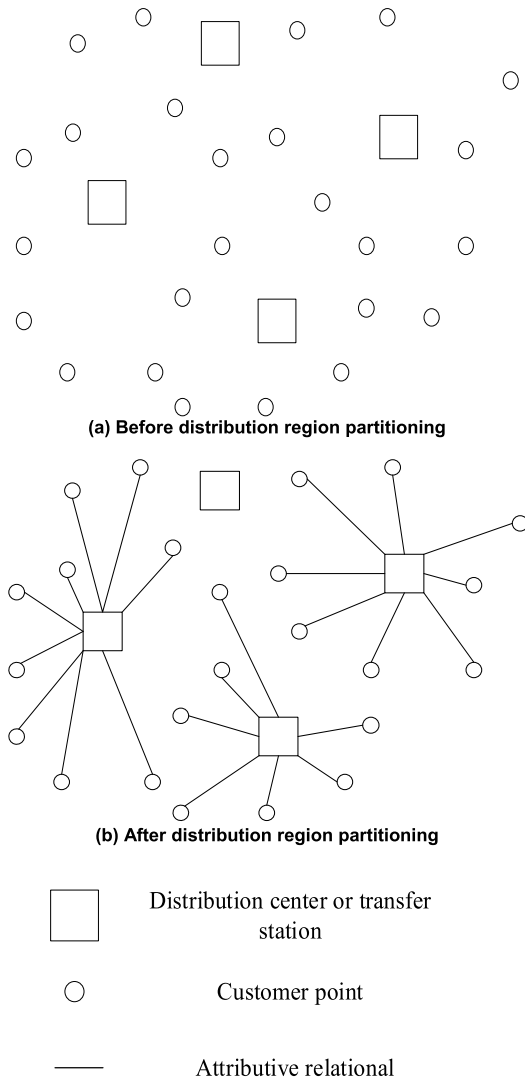


FIGURE 1. Distribution region partitioning.

$y_m = rand \times (\max(y) - \min(y))$, if $m = 1$, turn to (4), if $m > 1$, then turn to (5).

(4) From $m' = 1$ to $m - 1$, verifying whether the center distance from the distribution region center of the m^{th} vehicle to the first $m - 1$ vehicle is greater than or equal to the safety radius sr . if so, $m = m + 1$, turn to step (2). If not, turn to step (5).

(5) Record the number of repeated calculations $kn = kn + 1$, if kn is less than the iteration limit value KN , return to step (3) to regenerate the random center. If $kn = KN$, it indicates that the algorithm may stagnate, and return to step (1) for recalculation.

(6) Cluster the coordinates of all customer points to each AC point according to the principle of minimum Euler distance, and calculate the G value from Equation (11).

Simulated annealing algorithm has been reported in VRP for the generation of specific distribution paths. In this paper, the simulated annealing algorithm is used to divide the distribution region of vehicles [28], [29], [30]. The flow of simulated annealing algorithm is shown in Figure 2.

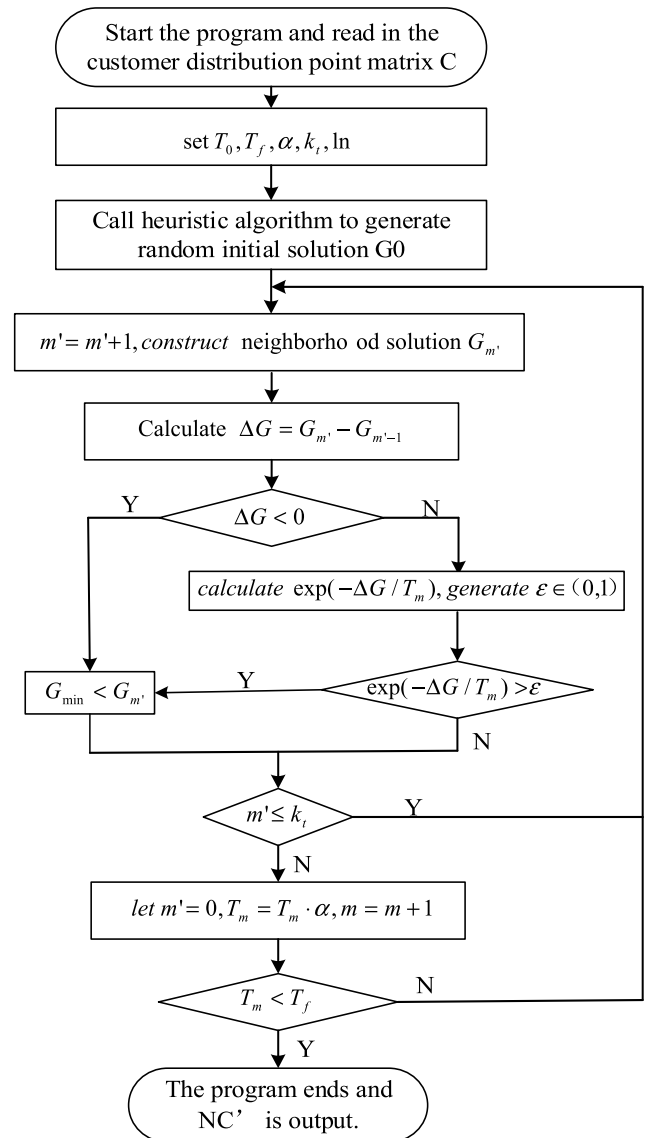


FIGURE 2. Flow chart of simulated annealing algorithm.

The logistics distribution model considering the complex road network constructed in the previous stage is taken as the input of the first stage decision-making, the distribution region NC_m of divided N vehicles is taken as the output, and transmitted to the second stage decision-making as the input.

IV. HEURISTIC PATH SEARCH ALGORITHM

After the first stage decision is completed, the distribution task set NC' of N vehicles will be output. For the second stage of decision-making, a heuristic hybrid genetic-beam search algorithm is proposed in this paper.

A. LOGISTICS DISTRIBUTION CONSTRAINTS AND OBJECTIVE FUNCTION CONSTRUCTION CONSTERING COMPLEX ROAD NETWORK

Considering the vehicle routing problems of complex road networks, in the actual distribution line, railways, rivers, and other traffic obstacles cannot be crossed directly. If the

distribution line passes through the above obstacles, vehicles need to bypass them, resulting in detour costs. The obstacles that the bypass cost exceeds the scope of distribution must be avoided in the route planning.

The total cost of vehicle distribution considering the complex road networks can be expressed as:

$$GD = \min(\sum_{k' \in M} \sum_{v_i, v_j \in D} f_{y_{v_i, v_j}} x_{J_d, J_{d'}}^{k'} + \sum_{k' \in M} F \sum_{J_d \in V'} x_{v_0, J_{d'}}^{k'} + l) \quad (12)$$

$$l = \omega \left(\sum_{k' \in M} \sum_{J_d \in S} \sqrt{\left(\bar{x} - \frac{\sum_{J_d \in S} \bar{x}}{S} \right)^2 + \left(\bar{y} - \frac{\sum_{J_d \in S} \bar{y}}{S} \right)^2} \right) \quad (13)$$

$$\sum_{k' \in M} \sum_{J_d \in V'} x_{J_d, J_{d'}}^{k'} = 1, \quad \forall J_d \in V' \quad (14)$$

$$\sum_{J_d \in V} x_{J_d, v_p}^{k'} - \sum_{J_{d'} \in V} x_{J_{d'}, v_j}^k = 0, \quad \forall v_p \in V', \forall k' \in M \quad (15)$$

$$\sum_{J_{d'} \in V'} \sum_{J_d \in D} x_{J_d, J_{d'}}^{k'} q_d \leq Q, \quad \forall k' \in M \quad (16)$$

$$x_{J_d, J_{d'}}^{k'} \in (0, 1), \quad \forall J_d, J_{d'} \in D, \forall k' \in M \quad (17)$$

$$\sum_{k' \in M} x_{J_d, J_{d'}}^{k'} = 0, \quad a_{v_i, v_j} \in NC \quad (18)$$

Equation (12) is the total vehicle operation cost, which is composed of the driving cost, vehicle fixed cost, and time delay cost. Equation (13) is the time delay cost, where, ω represents the time delay cost corresponding to the unit distance. $\sum_{d=1}^S |J_d|$ represents the number of all customer points in all urban logistics distribution regions. The time delay cost l is directly proportional to the sum of the distances from all customers in the line to the geographical center. The time delay cost here is only used to compare the advantages and disadvantages of schemes, and the value of a single scheme has no practical operational significance. Equations (14) and (15) ensure that a customer distribution center is visited only once, and the incoming vehicles must drive out. Equation (16) is to ensure that the loading capacity of each vehicle meets the limit of rated loading capacity. In equation (17), $x_{J_d, J_{d'}}^{k'}$ represents the decision variable, when $x_{J_d, J_{d'}}^{k'} = 1$, it means that the k' -th vehicle accesses the customer distribution node $J_{d'}$ after visiting the customer distribution node J_d , otherwise, it takes 0. Equation (18) ensures that the traffic obstacles that must be avoided do not appear in the distribution line.

Input parameters:

- J : A set of customer distribution nodes, $J = \{J_1, J_2, J_3, \dots, J_s\}$
- C : set of customer distribution point requirements
- S : Total number of customer distribution points, $d = 1, 2, 3, \dots, S$
- c_d : Demand information of the d^{th} customer distribution node J_d
- (\bar{x}, \bar{y}) : Coordinate of the d^{th} customer distribution node J_d
- F : Vehicle fixed cost
- l : Time delay cost
- v_0 : Distribution Centre
- M : set of vehicles, $M = (1, 2, \dots, k', \dots, m, \dots, N)$
- ω : Time delay cost corresponding to unit distance
- D : set of distribution centers and customers
- $V' = D \setminus \{v_0\}$: s customer distribution nodes
- FY : Cost matrix, cost $f_{y_{v_i, v_j}} \in FY$ corresponding to each path $a_{v_i, v_j} \in A$
- $v_i = v_i(x, y)$: Coordinates of path node v_i
- Q : Loading capacity of the vehicles
- A : The set of paths is composed of the shortest paths between any two points in D . The traffic fault paths that must be avoided in A are put into the set $NC, NC \in A$
- S' : The number of distribution regions, $S' \leq S$

B. VEHICLE ROUTING OPTIMIZATION

According to the distribution task set NC' of N vehicles output after the completion of the first stage decision. When using genetic algorithm to plan the vehicle path, because the algorithm has a certain dependence on the selection of the initial population [31], [32], [33], and has a certain randomness when searching the path, it can be improved in combination with some heuristic algorithms.

Beam search algorithm [34], [35], [36], [37] is a heuristic graph search algorithm, which is usually used when the solution space of the graph is relatively large. In order to reduce the space and time occupied by the search, some nodes with poor quality are cut off and some nodes with high quality are retained during each step of depth expansion. This reduces space consumption and improves time efficiency. In the second stage, the hybrid genetic-beam search (HGBS) algorithm is used to plan the path nodes.

The algorithm flow is shown in Figure 3. The algorithm takes the customer node as the starting point of initialization and the path search minimization cost as the output. The specific steps are as follows:

- (1) Chromosome coding [34]

The route information and distribution information are chromosome coded, and the two-dimensional chromosome coding method is adopted.

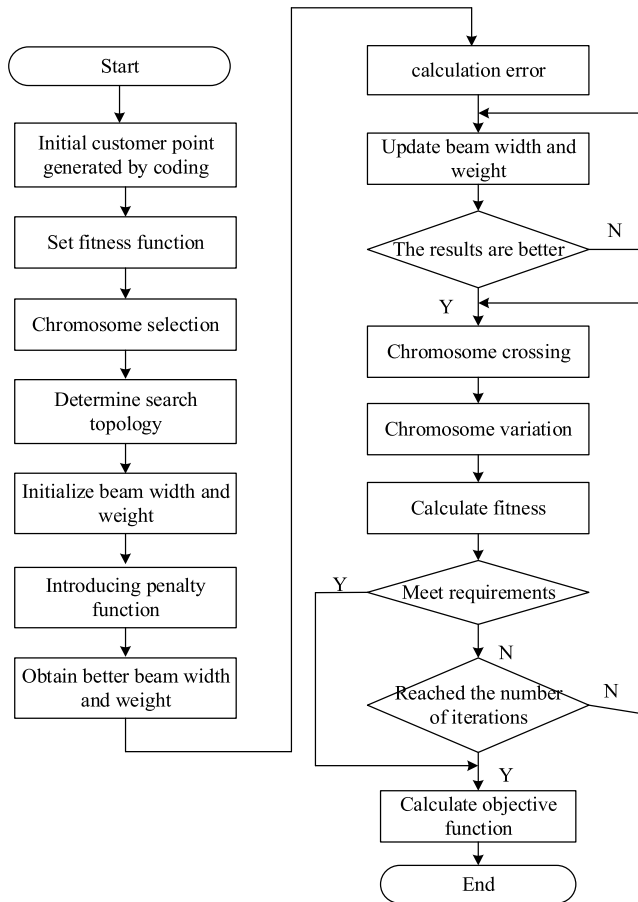


FIGURE 3. Flow chart of HGBS algorithm.

The first dimension is the natural number sequence: $1, 2, 3, \dots, S', S'$ the number of distribution regions.

The second dimension is the position of the chromosome. The chromosomes encoded are illustrated in Table 2. For example, $[1, (3, 1)]$ represents in the sub-region 1, the NO. 3 customer point is served by No. 1 distribution center.

The chromosome position is expressed as the serial number of the customer point assigned to each distribution center, y_{fd} represents the distribution center assigned to the d customer distribution node in chromosome f .

(2) Fitness function

$$Z = \frac{1}{GD + QZ \times S'_d} \quad (19)$$

where, QZ is the penalty weight. The penalty weights can facilitate the genetic algorithm to search the global optimal solution from both feasible and infeasible domains. In our work of logistics distribution of complex road networks, the value of QZ cannot be set to 0 in order to obtain the optimal solution set. S'_d is the chromosome number violating the maximum transportation distance, and the number of unqualified chromosomes is added by 1, that is $S'_d = S'_d + 1$, the initial $S'_d = 0$.

(3) Chromosome selection

Using Monte Carlo method to select operators, the greater the individual fitness, the higher the probability of being selected. If the population number is m' and the appropriate value of individual n' is F , the probability of selecting individual a is:

$$P = \frac{F_{n'}}{\sum_{n'}^{m'} F_{n'}} \quad (20)$$

(4) Determine the topology of beam search, that is, generate a set of divisible customer distribution points.

(5) Initialize search beam width and weight

Judge whether the number of initial nodes N is greater than bw , if $N > bw$, then continue. If $N < bw$, branch the initial node, where bw is the beam width of beam search.

(6) Introducing penalty function

$$\max GD \geq QZ \quad (21)$$

(7) Crossover and variation

Two individuals w_r and w_l are randomly selected from the parent generation, and the connection values are randomly and independently selected for exchange. The cross operation at bit b is as follows:

$$w_{rb} = w_{rb}(1 - \beta) + w_{lb}\beta \quad (22)$$

$$w_{lb} = w_{lb}(1 - \beta) + w_{rb}\beta \quad (23)$$

where β is the random number between $[0, 1]$.

Complete uniform mutation with a predetermined probability to improve individual fitness and approach the optimal solution from a local point of view. The new gene value after mutation is:

$$w' = \gamma(w_{\max} - w_{\min}) + w_{\min} \quad (24)$$

where w_{\max} , w_{\min} are the maximum and minimum values of the initial individual, γ is the random number between $[0, 1]$.

To sum up, after many cross variations, the optimal path of urban logistics distribution region is obtained.

V. CASE ANALYSIS

In this section, we use Matlab 2020a to carry out simulation experiments. The experimental equipment is a computer with i7-7500 2.90GHz CPU. The computer system is windows10 64 bit Professional Edition with 4G RAM.

In the first stage of decision-making, we use two examples for analysis. Case1: Based on the road network of a city in southern Jiangsu, a logistics distribution model considering complex road networks is established according to the road network. There are 286 road network nodes in the city, 10 logistics distribution centers and 95 customer distribution centers. The 95 customer distribution centers are expressed with C1, C2, C3, ..., C95, and the 10 distribution centers are represented by D1, D2, D3, ..., D10. The geographical distribution of each point is roughly shown in Figure 4. Case 2: Based on the basic data of the urban logistics distribution region partitioning in Literature [4], the

TABLE 2. Schematic of chromosome codes.

The first dimension	1	1	2	...	S'
The second dimension	(3,1)	(5,2)	(3,4)	...	(d, y_{fd})

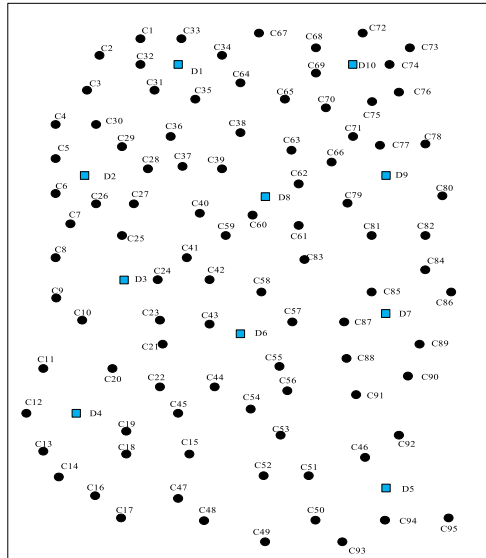


FIGURE 4. Spatial distribution of distribution center and customer point.

data in [4] includes 279 road network nodes, 105 distribution units, five candidate distribution centers, and 1 logistics center. According to the basic data of urban logistics distribution, we regard the logistics center as a distribution center and jointly undertake the basic logistics distribution tasks.

The purpose of the distribution region partitioning is to determine the number of distribution centers and their corresponding locations. In addition, each distribution unit or customer point should be reasonably allocated to a certain logistics distribution center. Table 3 measures the distance between 10 logistics distribution centers and the top 15 customer distribution centers in Case 1.

Table 4 lists the coordinates of some customer distribution nodes and road network nodes and the demand information for customer distribution nodes.

In the second stage of decision-making, we first discussed the hybrid idea of the algorithm in this study. In order to solve the problem of a large search range of genetic algorithms, we mixed genetic algorithm and beam search algorithm to realize the pruning operation of beam search in three cases: before chromosome crossover, after chromosome mutation, both before chromosome crossover and after chromosome mutation. Then, the efficiency of the proposed heuristic hybrid algorithm is evaluated on the Solomon data set. Finally, the vehicle route optimization is realized by the region partitioning result of the example in the first stage of decision-making.

A. REGION PARTITIONING CALCULATION

For the simulated annealing algorithm, the initial temperature is 600° , the cooling coefficient is 0.99, the cooling times are 1000, the number of internal cycles before each cooling is 100, and the number of distribution vehicles is set according to the number of distribution centers.

Fig. 5 is an iterative convergence diagram of the first stage region division objective function G . It can be seen from Fig.5 that when the simulated annealing algorithm performs region division, there is a large fluctuation in the early stage, and the g value converges to Case 1:418 and Case 2:512 in the later stage. Figure 6 and Figure 7 are the regional division results of Case 1 and Case 2 respectively. It can be seen from the figure that the division results have good results in terms of customer quantity and load (calculated load variance: Case 1 412, Case 2: 509, the variance of customer nodes: Case 1:0.627, Case 2:0.718).

B. VEHICLE ROUTING PLANNING CALCULATION

Since each customer distribution node usually contains more than 20 customers, the requirements of the Solomon data-set or extended Solomon data-set are collected for each customer. In order to better reflect the actual distribution demand of each customer point, in this embodiment, the customer delivery demand of RC1_2_1 in Solomon data set [19] is multiplied by 20 as the delivery demand, that is, the data set is expanded to meet the actual distribution demand.

For the HGBS algorithm, setting vehicle carrying capacity $Q = 2000$, the fixed cost of vehicles $F = 500$, the fixed cost of vehicles $F = 500$, the maximum transportation distance of vehicle $LD = 35$ mile, the beam Search width of bundle search $bw = 4$, the maximum number of generations is $S_{max} = 500$, the crossover probability $p_c = 0.8$, the mutation probability $p_m = 0.02$.

In the first step, the fusion methods of the three algorithms are compared and verified on the Solomon data set. The minimum cost, number of iterations, and algorithm running time obtained by random experiments 5 times are shown in Table 5.

It can be seen from Table 5 that before the chromosome crossing, the beam search algorithm is integrated, and the pruning function is introduced to cut the chromosome sequence that does not meet the transportation distance limit and vehicle capacity constraint, and pruning is performed once before the chromosome crossover, after the chromosome mutation, both before and after the chromosome crossover. The experimental results show that after five simulation experiments, the experimental results show that the average total cost of pruning before chromosome crossing is

TABLE 3. The distances (miles) between 10 logistics distribution centers and top 15 customer distribution centers.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
D1	12.7	48.7	54.9	48.6	33.8	41	22.2	55.9	14.2	51.5	13.8	27.6	46.9	58.5	66.1
D2	69.7	58.2	36.5	32.7	33.5	23	38.1	38.3	61.3	59.6	48.3	28.4	60.9	40	26.3
D3	70.4	65.7	41.3	46.7	44	15.6	22.6	35.3	17.3	63.3	14.6	16.9	12.8	17.1	23.3
D4	69.3	32.3	13.9	67.9	73.5	32.9	8.3	19.4	30.7	44.6	19.7	45.2	53.3	16.6	8.8
D5	22.3	23.9	31.8	6.4	60.1	2.2	54.8	36.6	43.4	17.8	34.4	72.2	39.1	17.4	36.7
D6	46.8	50.9	29.7	74.1	2.8	66.4	68.5	59.7	7.4	19.6	25.2	51	10.2	54.1	8
D7	49	37.1	58.4	53.6	67.8	66.8	25.1	52.4	14.8	2.3	55.8	37.5	36	45.7	46.3
D8	64.5	60.4	43.3	13.7	18	66.5	12.6	73.4	53.5	4.5	51.1	3.2	5.4	7.3	61.4
D9	54.2	11.2	49.5	38.9	73	60	34	32.4	61.9	6.3	10	13	29.3	62.4	60.3
D10	29.9	39.5	31.3	49.3	47.1	21.9	1.2	73.8	12.5	27.9	14.9	25.5	71.4	69	4

Note: Di = the i-th distribution center, Ci = the i-th customer point.

TABLE 4. Customer distribution node coordinates and requirements.

Number	Coordinates / load (kg)	Number	Coordinates / load (kg)	Number	Coordinates / load (kg)
1	(5.0,35.0)/200	70	(28.0,35.0)/700	123	(40.0,50.0)/320
2	(0,40.0)/500	71	(18.0,75.0)/400	124	(20.0,80.0)/640
3	(18.0,75.0)/100	72	(45.0,68.0)/410	125	(87.0,30.0)/0
4	(8.0,40.0)/0	73	(30.0,50.0)/0	126	(65.0,60.0)/0
5	(2.0,40.0)/160	74	(58.0,75.0)/190	127	(90.0,35.0)/230
6	(22.0,75.0)/210	75	(40.0,69.0)/230	128	(8.0,70.0)/160
...

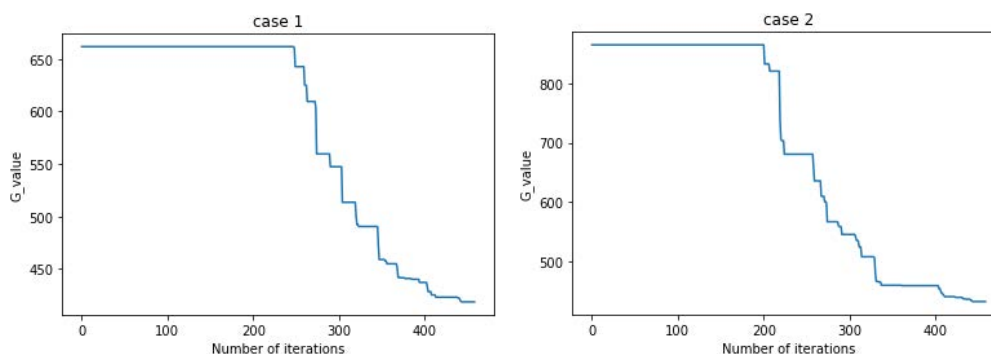


FIGURE 5. Iterative convergence curve of simulated annealing algorithm.

the lowest, which is \$28815.8, which is about 0.72% lower than the average total cost of pruning after mutation, and about 0.07% lower than the average total cost of pruning before chromosome crossing and after mutation. Under the three different operations, it is relatively better to prune both before crossing and after mutation at the same time in terms of the number of iterations and the execution time of the algorithm to obtain the optimal results, the reason is that the pruning operation is carried out before crossing and after mutation at the same time, so that the nodes that do not meet the current optimum at each stage are pruned. This is a greedy strategy. Therefore, the number of iterations and the execution time of the algorithm are relatively small, but the result will fall into the local optimum and the global optimum

solution cannot be accurately obtained. Therefore, in this study, we chose to prune before crossing. In addition, the number of iterations of the algorithm is significantly better than that of the other two cases in both minimum cost and optimal results. Pruning the unqualified chromosome before the chromosome crossing ensures the best of the father generation, and the offspring obtained after the cross mutation basically inherits the best of the parent, so as to ensure the relative optimal result. However, if the chromosome is pruned after the cross mutation, it is pruning in the offspring, which cannot guarantee global optimization, so the result is optimal.

In the second step, the efficiency of the proposed heuristic hybrid algorithm is verified on Solomon data set. We use the same Solomon dataset to implement and test different

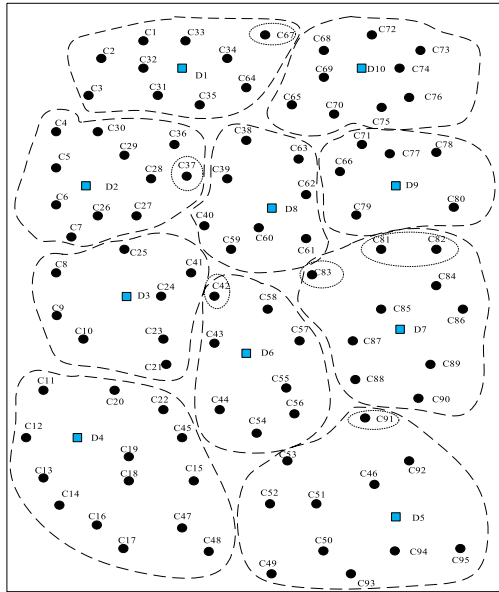


FIGURE 6. Case 1: distribution region partitioning result.

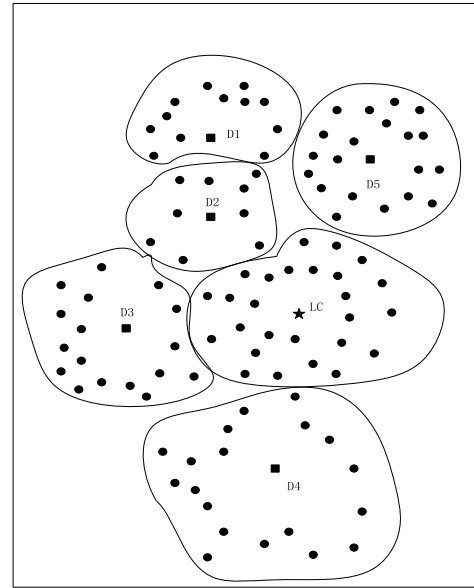


FIGURE 7. Case 2: distribution region partitioning result.

algorithms related to this paper such as HPSO [38], GA [39], ACO [40], EPSO-GA [4], 2MPSO [41], IPSO [42], TS-MOEA [12], HPSO-HGA [18].

Most of the routing based meta heuristic search and optimization algorithms which have been searched and collected for the first time in this paper have been developed relatively more recently. All algorithms are carried out under equal conditions. Equal conditions mean using the same starting and termination criterion, equal number of starting search points, the same data set, same hardware running the algorithms.

Each algorithm is executed for 10 times, and we selected the optimal solution as the distribution region partitioning results for each method. For the optimal cost, Yin *et al.*'s optimal solution is $GD=\$31577$, Shima *et al.*'s optimal solution is $GD=\$32693$, and Chen *et al.*'s optimal solution is $GD=\$34672$, and Wang *et al.*'s optimal solution is $GD=\$29993$, and Okulewicz *et al.*'s optimal solution is $GD=\$31106$, and Hannan *et al.*'s optimal solution is $GD=\$29653$, and Wang *et al.*'s optimal solution is $GD=\$28906$, and Lu *et al.*'s optimal solution is $GD=\$30017$.

In addition, HGBS algorithm and other algorithms are executed 10 times with the optimal cost, the number of iterations, and algorithm running time for convergence. The results are shown in Table 6, Table 7, and Table 8. When HGBS algorithm is used for vehicle route planning, the average total cost is the lowest, about $\$28912.6$. Compared with TS-MOEA ($\$29025.2$), which is the best performing algorithm in other algorithms, the average total cost is reduced by about 0.39%.

For statistical analysis, t-test method has been performed, as shown in the last column of Table 6 and Table 7.

In order to compare the performance of the proposed algorithm, the same iteration parameters are used. In the t-test:

H_0 : It is argued that there is no difference between the means.

H_a : It is argued that there is a meaningful difference between the means.

P: Probability value

t Stat: t statistic value

Pearson Correlation: The correlation coefficient between LCA and OIO samples

t Critical one-tail: Single-sided t critical value

t Critical two-tail: Double-sided t critical value

alpha: Significant level

In the t-test, there are two hypotheses, H_0 and H_a . When the p value is less than 0.01, the H_0 hypothesis is rejected and H_a is accepted. When p value is greater than or equal to 0.01, H_0 hypothesis is accepted and H_a is rejected. If the test result value of P is less than 0.01, there is a significant difference between the two groups of data. The smaller the P value, the more obvious the difference is, and the greater the difference is from the benchmark data.

It can be seen from Table 6 and Table 7 that compared with HGBS algorithm, the p-values of the other eight algorithms are significantly less than 0.01 and all are negative. The t-test of ts-moea is relatively optimal, which is -1.18 . The t-test results show that the results obtained by this algorithm are relatively good compared with other algorithms, and the t-test results of ACO are relatively poor, which is -28.65 . The t-test results show that the results obtained by this algorithm are relatively poor compared with other algorithms.

Based on the t-test results, both the distribution cost per day and the number of iterations are significantly different than each of the other methods. Compared with the Genetic Algorithm, the total cost of the algorithm is reduced by 14.1%, and the number of iterations required to reach the optimal solution is reduced by 152 times. The reason why HGBS algorithm has achieved this experimental effect is that the pruning operation is carried out before the chromosome crossing of GA, which

TABLE 5. Comparison of experimental results under three fusion methods.

Operation	Total cost			Number of iterations			Execution time		
	Before crossover	After mutation	Before crossover and after mutation	Before crossover	After mutation	Before crossover and after mutation	Before crossover	After mutation	Before crossover and after mutation
1	28623	28932	28902	312	314	308	2.87	2.88	2.77
2	28941	28996	28703	314	315	310	2.76	2.92	2.71
3	28623	29008	29001	320	311	307	2.89	2.81	2.80
4	28835	29104	28731	313	312	312	2.90	2.83	2.69
5	29062	29083	28843	318	317	311	2.84	2.85	2.82
Average	28816.8	29024.6	28836	315.4	313.8	309.6	2.852	2.858	2.758

TABLE 6. Different optimization algorithms comparison (total cost).

Method	1	2	3	4	5	6	7	8	9	10	Average	t-test
HGBS(this paper)	28837	28951	28996	28623	29017	28623	29618	28623	28836	29002	28912.6	—
HPSO(Yin et al., 2006)	33829	31577	31996	32098	34006	31910	33217	32926	32152	31609	32532	-12.11
GA(Shima et al., 2006)	33598	34609	32901	34017	33969	33497	35109	32693	33281	32886	33656	-17.93
ACO(Chen et al., 2011)	36011	34905	35217	34672	36218	34810	35521	36254	34516	35171	35329	-28.65
EPSO-GA (Wang et al., 2015)	30117	32096	31249	30843	32091	33095	30076	32174	31548	29993	31328	-6.94
2MPSO(Okulewicz et al 2017)	31652	32001	31408	31301	31189	31285	31106	31260	31456	31309	31396.7	-19.93
IPSO(Hannan et al 2018)	29653	29773	30017	30827	29831	30188	30711	30227	30142	30219	30159	-8.20
TS-MOEA(Wang et al 2018)	28991	29102	28906	29001	29053	29034	29068	29083	28996	29018	29025.2	-1.18
HPSO-HGA(Lu et al 2020)	30298	30139	31031	30082	30017	31105	30182	30421	30118	30234	30362.7	-9.38

ensures the optimal selection of the parent generation, so that the offspring obtained by the individuals after the pruning after the mutation crossing operation basically inherits the optimal selection of the parent generation, thus ensuring the relatively optimal results.

In addition, as can be seen from Table 8, in terms of the speed of obtaining the optimal solution, through 10 experiments and simulations, the average algorithm execution time of HGBS algorithm is 68.5s, which is 18.85% higher than that of HPSO-HGA (84.41s), and 55.19% higher than that of 2mpso (106.31s). Compared with other algorithms, HGBS algorithm can obtain the optimal solution faster. This means that HGBS algorithm is more likely to find the optimal solution than other methods. That is, the method proposed in this paper can better capture the partitioning scheme, optimal cost and iteration times, and show its ability to solve the complex distribution region partitioning problem and logistics distribution vehicle routing optimization in urban logistics distribution networks. The reason why HGBS algorithm can obtain the optimal solution in a short time is that the pruning operation before the individual crossing avoids the search for

the relative inferior solution and reduces the execution time of the algorithm on the premise that the individual of the parent generation is guaranteed to be optimal.

Compared with the other four algorithms, HGBS algorithm has the following merits:

- (1) HGBS algorithm is a hybrid algorithm combining genetic algorithm and beam search algorithm, which has global and local search capabilities.
- (2) BS algorithm is a graph search algorithm, it is similar to the directed graph of logistics distribution constructed by the logistics distribution model considering complex road networks.
- (3) The unique pruning function of BS algorithm in the search process can set the search width for the next node search at any time according to the quality of the customer points obtained from the expansion.
- (4) HGBS algorithm has inherent advantages and can be applied to large-scale logistics distribution network with thousands of customers. As shown in Table 6 and Table 7, the number of iterations of HGBS algorithm is significantly less than that of other methods. As the number

TABLE 7. Different optimization algorithms comparison (number of iterations).

Method	1	2	3	4	5	6	7	8	9	10	Average	t-test
HGBS(this paper)	312	319	321	311	317	308	310	314	307	313	313.2	—
HPSO(Yin et al., 2006)	388	412	396	408	421	417	409	399	405	414	407	-26.60
GA(Shima et al., 2006)	453	469	457	461	477	480	454	466	471	459	465	-45.59
ACO(Chen et al., 2011)	510	496	487	501	512	483	492	505	488	517	499	-46.73
EPSO-GA (Wang et al., 2015)	365	389	371	402	393	378	369	392	387	411	386	-14.71
2MPSO(Okulewicz et al 2017)	344	341	346	348	350	339	342	345	351	357	346.3	-14.77
IPSO(Hannan et al 2018)	352	357	360	351	364	347	358	362	366	359	358	-18.51
TS-MOEA(Wang et al 2018)	320	324	328	323	319	342	337	329	322	330	327.4	-5.12
HPSO-HGA(Lu et al 2020)	332	337	329	348	341	340	336	338	331	342	337.4	-10.39

TABLE 8. Different optimization algorithms comparison (algorithm running time/second).

Method	1	2	3	4	5	6	7	8	9	10	Average
HGBS(this paper)	67.8	69.2	68.3	68.4	69.1	67.7	68.9	69.3	67.8	68.5	68.5
HPSO(Yin et al., 2006)	98.3	101.1	99.6	98.6	100.3	99.4	95.9	97.6	98.3	99.4	98.85
GA(Shima et al., 2006)	102.3	100.2	103.2	99.8	102.2	100.7	101.6	101.4	101.6	101.8	101.48
ACO(Chen et al., 2011)	103.1	102.9	101.9	104.8	106.2	107.2	104.5	105.7	103.8	103.6	104.37
EPSO-GA (Wang et al., 2015)	100.2	101.5	102.1	103.5	104.5	102.6	106.7	103.7	106.2	107.3	103.83
2MPSO(Okulewicz et al 2017)	108.3	107.2	107.6	108.3	109.2	106.2	105.2	106.2	103.2	101.7	106.31
IPSO(Hannan et al 2018)	102.2	100.2	103.3	104.2	103.9	102.8	103.2	101.3	103.4	101.4	102.59
TS-MOEA(Wang et al 2018)	98.2	96.2	97.2	96.3	96.2	98.3	97.5	97.4	96.4	96.8	97.05
HPSO-HGA(Lu et al 2020)	88.8	89.2	89.5	87.3	78.2	80.2	80.4	82.4	84.5	83.6	84.41

TABLE 9. Results of the two cases.

Method	Case 1			Case 2		
	cost	number of iterations	Algorithm running time	cost	number of iterations	Algorithm running time
This paper	28912	313	68.5	31174	407	78.3
GA	37416	467	101.48	39505	515	104.31
EPSO-GA	32173	356	103.83	33038	396	105.62
IPSO	30017	334	102.59	33419	466	104.88
HPSO-HGA	29434	330	84.41	33409	426	98.37

of customers increases, this advantage will become more obvious.

In summary, the HGBS algorithm proposed in this paper is a hybrid algorithm combining genetic algorithm and beam search algorithm. It is designed and implemented by updating penalty weight and pruning the parent chromosome.

In addition, in the two-stage algorithm, we first improved the HSA algorithm with the goal of balancing the vehicle load to achieve the division of logistics distribution regions. Through the HSA algorithm, the entire distribution region is divided into independent sub regions, so as to reduce the region scale and the path optimization range, reduce the path search cost, and improve the global search effect of the path. Secondly, in order to minimize the total cost of logistics distribution, considering the complex road network constraints of each sub region after the division of regional road network, an improved path search strategy based on genetic algorithm is proposed. Before the chromosome crossing of genetic algorithm, the beam search algorithm is used to prune the unqualified chromatids to ensure the optimization of the parent nodes, so as to improve the search ability of GA for the local optimal solution of the path. However, the algorithm still has some disadvantages, such

as high computational complexity, imperfect decision space performance, and difficult experimental parameter setting. In the follow-up study, the above factors need to be comprehensively considered to improve the performance of the algorithm.

The proposed hybrid algorithm can obtain a better optimal solution for most of the randomly generated initial populations, but for those poor initial populations, the pruning function rarely appears in the optimal solution. Compared with other methods, the optimization success rate of the method is very high (3 times in 10 times).

Finally, according to the regional division results of two examples in the first stage decision-making, the logistics distribution vehicle route is optimized, and the classical genetic algorithms GA, EPSO-GA [4], IPSO [42], and HPSO-HGA [18] are compared to verify the effectiveness of the method proposed in this paper on two examples. The operation comparison results are shown in Table 9.

The experimental results of the two practical cases are shown in Table 9. The average total cost of the two-stage algorithm proposed in this paper on Case1 is about \$28912, which is about 1.8% lower than the optimal HPSO-HGA (\$29434) among other algorithms, the number of iterations

of the algorithm is reduced by about 5.4%, and the execution time of the algorithm is increased by about 23.2%. The average total cost on Case2 is \$31174, which is about 7.2% lower than the optimal HPSO-HGA (\$33409) in other algorithms, the number of algorithm iterations is about 4.7%, and the algorithm execution time is about 25.6%. As can be seen from Table 9, on the premise of considering the complex urban road network structure, the distribution region partitioning method proposed in this paper can well realize the division of logistics distribution regions, so as to optimize the vehicle path in each region and minimize the overall cost. The experimental results show that the proposed method has some advantages in scheme division, total cost and iteration times. Through comparison, it is proved that the logistics distribution region partitioning method has the ability to optimize the logistics distribution network. The proposed method is easy to implement in practice, can effectively divide the urban logistics distribution region, and helps logistics operators reduce operating costs and improve customer service.

VI. CONCLUSION

To solve the problem of difficult region partitioning and routing planning in large-scale logistics distribution under complex road network conditions, this paper establishes a mathematical model of logistics distribution based on consideration of traffic elements and road network topology, and proposes a two-stage vehicle routing optimization scheme for logistics and distribution based on HSA-HGBS algorithm. In the first stage, the overall distribution region is divided into independent sub-regions by HSA algorithm to balance the vehicle task load. In the second stage of routing decision, with the goal of minimizing the total cost of logistics and distribution, the HGBS-based routing search method is proposed to reduce the randomness of the model in the initial search path by heuristic genetic algorithm, and then combined with Beam search method to reduce the space and time occupied by the search.

In this paper, 10 experiments are conducted on the standard data set for sub-region routing decisions. When HGBS algorithm is used for vehicle route planning, the average total cost is the lowest, about \$28912.6. Compared with TS-MOEA (\$29025.2), which is the best performing algorithm in other algorithms, the average total cost is reduced by about 0.39%. The results show that the HGBS algorithm can effectively improve the effectiveness of independent sub-region routing optimization. Through the experimental verification of the proposed algorithm in two practical cases, the average total cost of the HSA-HGBS algorithm in Case1 and Case is about 1.8% and 7.2% lower than the best result in the comparison algorithm, and the number of iterations of the algorithm of the algorithm are reduced by about 5.4% and 4.7% respectively. The results show that the two-stage algorithm proposed in this paper can effectively divide the urban logistics distribution region, reduce the region scale and route search cost, and improve the efficiency and rationality of the overall physical distribution route selection.

Although the two-stage decision algorithm proposed in this paper can better optimize the large-scale urban logistics distribution considering the complex road network, there are still some defects. For example, the complex road network considered in this paper is limited to the complex road network problem composed of detailed and complex traffic elements and the topological structure of the road network. Although it reflects the actual road network situation to a certain extent, further analysis is still needed in terms of traffic rules, population density, and measures of motor vehicles (or more accurately, road parameters causing problems). In addition, building a more realistic distribution data set is also the focus and difficulty of the next research.

REFERENCES

- [1] G. B. Dantzig and J. H. Ramser, "The truck dispatching problem," *Manage. Sci.*, vol. 6, no. 1, pp. 80–91, 1959.
- [2] T. Wagner, "Regional traffic impacts of logistics-related land use," *Transp. Policy*, vol. 17, no. 4, pp. 224–229, Aug. 2010.
- [3] N. Christofides, "Fixed routes and areas for delivery operations," *Int. J. Phys. Distrib.*, vol. 1, no. 2, pp. 87–92, Dec. 2013.
- [4] Y. Wang, X. Ma, M. Xu, Y. Liu, and Y. Wang, "Two-echelon logistics distribution region partitioning problem based on a hybrid particle swarm optimization–genetic algorithm," *Exp. Syst. Appl. Int. J.*, vol. 42, no. 12, pp. 5019–5031, Jul. 2015.
- [5] H. Zhang, H. Ge, J. Yang, and Y. Tong, "Review of vehicle routing problems: Models, classification and solving algorithms," *Arch. Comput. Methods Eng.*, vol. 29, no. 1, pp. 195–221, 2021.
- [6] S. Akyo and B. Alatas, "Plant intelligence based metaheuristic optimization algorithms," *Artif. Intell. Rev.*, vol. 47, no. 4, pp. 417–462, 2016.
- [7] B. Alatas and H. Bingol, "Comparative assessment of light-based intelligent search and optimization algorithms," *Light Eng.*, vol. 28, no. 6, pp. 51–59, Dec. 2020.
- [8] B. Alatas and H. Bingol, "A physics based novel approach for travelling tournament problem: Optics inspired optimization," *Inf. Technol. Control*, vol. 48, no. 3, pp. 373–388, Sep. 2019.
- [9] H. Bingol and B. Alatas, "Chaos based optics inspired optimization algorithms as global solution search approach," *Chaos, Solitons Fractals*, vol. 141, Dec. 2020, Art. no. 110434.
- [10] K. Heßler and S. Irnich, "A branch-and-cut algorithm for the soft-clustered vehicle-routing problem," *Discrete Appl. Math.*, vol. 288, pp. 218–234, Jan. 2021.
- [11] R. Lahyani, M. Khemakhem, and F. Semet, "Rich vehicle routing problems: From a taxonomy to a definition," *Eur. J. Oper. Res.*, vol. 241, no. 1, pp. 1–14, Jul. 2015.
- [12] J. Wang, T. Weng, and Q. Zhang, "A two-stage multiobjective evolutionary algorithm for multiobjective multi-depot vehicle routing problem with time Windows," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 1–12, Apr. 2018.
- [13] C.-J. Ting, K.-C. Wu, and H. Chou, "Particle swarm optimization algorithm for the berth allocation problem," *Exp. Syst. Appl.*, vol. 41, no. 4, pp. 1543–1550, Mar. 2014.
- [14] S. Akpınar, "Hybrid large neighbourhood search algorithm for capacitated vehicle routing problem," *Exp. Syst. Appl.*, vol. 61, pp. 28–38, Nov. 2016.
- [15] S. Karakatić and V. Podgorelec, "A survey of genetic algorithms for solving multi depot vehicle routing problem," *Appl. Soft Comput.*, vol. 27, pp. 519–532, Feb. 2015.
- [16] Y. Wang, L. Wang, G. Chen, Z. Cai, Y. Zhou, and L. Xing, "An improved ant colony optimization algorithm to the periodic vehicle routing problem with time window and service choice," *Swarm Evol. Comput.*, vol. 55, Jun. 2020, Art. no. 100675.
- [17] A. Bortfeldt and Y. Junmin, "The split delivery vehicle routing problem with three-dimensional loading constraints," *Eur. J. Oper. Res.*, vol. 282, no. 2, pp. 545–558, Apr. 2020.
- [18] L. Zhen, C. Ma, K. Wang, L. Xiao, and W. Zhang, "Multi-depot multi-trip vehicle routing problem with time Windows and release dates," *Transp. Res. E, Logistics Transp. Rev.*, vol. 135, Mar. 2020, Art. no. 101866.

- [19] V. Y. E. Medeiros, M. B. R. De Albergaria, and S. J. O. Soares Da, "Multi-depot vehicle routing problem for large scale disaster relief in drought scenarios: The case of the Brazilian northeast region," *Int. J. Disaster Risk Reduction*, vol. 58, May 2021, Art. no. 102193, doi: 10.1016/j.ijdr.2021.102193.
- [20] M. M. Solomon, "Algorithms for the vehicle routing and scheduling problems with time window constraints," *Oper. Res.*, vol. 35, no. 2, pp. 254–265, Apr. 1987.
- [21] F. Li, B. Golden, and E. Wasil, "Very large-scale vehicle routing: New test problems, algorithms, and results," *Comput. Oper. Res.*, vol. 32, no. 5, pp. 1165–1179, May 2005.
- [22] D. Mester and O. Bräysy, "Active-guided evolution strategies for large-scale capacitated vehicle routing problems," *Comput. Oper. Res.*, vol. 34, no. 10, pp. 2964–2975, Oct. 2007.
- [23] L. Accorsi and D. Vigo, "A fast and scalable heuristic for the solution of large-scale capacitated vehicle routing problems," *Transp. Sci.*, vol. 55, no. 4, pp. 832–856, Jul. 2021.
- [24] C. Duhamel, P. Lacomme, and C. Prodhon, "Efficient frameworks for greedy split and new depth first search split procedures for routing problems," *Comput. Oper. Res.*, vol. 38, no. 4, pp. 723–739, Apr. 2011.
- [25] S. Vonolfen, M. Affenzeller, A. Beham, and S. Wagner, "Solving large-scale vehicle routing problem instances using an island-model offspring selection genetic algorithm," in *Proc. 3rd IEEE Int. Symp. Logistics Ind. Informat.*, Budapest, Hungary, Aug. 2011, pp. 27–31.
- [26] G. Wasil, "Computerized vehicle routing in the soft drink industry," *Oper. Res.*, vol. 35, no. 1, pp. 6–17, Feb. 1987.
- [27] Y. M. Cheong, H. L. Ong, and H. C. Huang, "Modeling the vehicle routing problem for a soft drink distribution company," *Asia-Pacific J. Oper. Res.*, vol. 19, no. 1, pp. 17–34, May 2002.
- [28] J. E. Beasley and N. Christofides, "Vehicle routing with a sparse feasibility graph," *Eur. J. Oper. Res.*, vol. 98, no. 3, pp. 499–511, May 1997.
- [29] L. Huo, J. Zhu, G. Wu, and Z. Li, "A novel simulated annealing based strategy for balanced UAV task assignment and path planning," *Sensors*, vol. 20, no. 17, p. 4769, Aug. 2020.
- [30] E. B. Tirkolaee, P. Abbasian, M. Soltani, and S. A. Ghaffarian, "Developing an applied algorithm for multi-trip vehicle routing problem with time Windows in urban waste collection: A case study," *Waste Manage. Res.*, vol. 37, no. 1, pp. 4–13, Jan. 2019.
- [31] H. Fan, Y. Zhang, P. Tian, Y. Lv, and H. Fan, "Time-dependent multi-depot green vehicle routing problem with time Windows considering temporal-spatial distance," *Comput. Oper. Res.*, vol. 129, May 2021, Art. no. 105211.
- [32] J. H. W. Iv and T. M. Cavalier, "A genetic algorithm for the split delivery vehicle routing problem," *Amer. J. Oper. Res.*, vol. 2, no. 2, pp. 207–216, Jan. 2012.
- [33] B. Rabbouch, F. Saâdaoui, and R. Mraïhi, "Efficient implementation of the genetic algorithm to solve rich vehicle routing problems," *Oper. Res.*, vol. 21, no. 3, pp. 1763–1791, Sep. 2021.
- [34] O. J. Ibarra-Rojas and Y. Silva-Soto, "Vehicle routing problem considering equity of demand satisfaction," *Optim. Lett.*, vol. 15, no. 6, pp. 2275–2297, Jan. 2021.
- [35] L. Zhou, J. Dang, H. Zhang, Y. Wang, and D. Gao, "Path optimization of stereoscopic garage stacker based on IGA-beam search," in *Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC)*, Chongqing, China, May 2019, pp. 119–124.
- [36] I. Araya, M. Moyano, and C. Sanchez, "A beam search algorithm for the biobjective container loading problem," *Eur. J. Oper. Res.*, vol. 286, no. 2, pp. 417–431, Oct. 2020.
- [37] M. Yavuz, "Iterated beam search for the combined car sequencing and level scheduling problem," *Int. J. Prod. Res.*, vol. 51, no. 12, pp. 3698–3718, Jun. 2013.
- [38] U. Golle, F. Rothlauf, and N. Boysen, "Iterative beam search for car sequencing," *Ann. Oper. Res.*, vol. 226, no. 1, pp. 239–254, Mar. 2015.
- [39] P.-Y. Yin, S.-S. Yu, P.-P. Wang, and Y.-T. Wang, "A hybrid particle swarm optimization algorithm for optimal task assignment in distributed systems," *Comput. Standards Interfaces*, vol. 28, no. 4, pp. 441–450, Apr. 2006.
- [40] T. Shima, S. J. Rasmussen, A. G. Sparks, and K. M. Passino, "Multiple task assignments for cooperating uninhabited aerial vehicles using genetic algorithms," *Comput. Oper. Res.*, vol. 33, no. 11, pp. 3252–3269, Nov. 2006.
- [41] H. Chen, A. M. K. Cheng, and Y.-W. Kuo, "Assigning real-time tasks to heterogeneous processors by applying ant colony optimization," *J. Parallel Distrib. Comput.*, vol. 71, no. 1, pp. 132–142, Jan. 2011.
- [42] M. Okulewicz and J. Mańdziuk, "The impact of particular components of the PSO-based algorithm solving the dynamic vehicle routing problem," *Appl. Soft Comput.*, vol. 58, pp. 586–604, Sep. 2017.
- [43] M. A. Hannan, M. Akhtar, R. A. Begum, H. Basri, A. Hussain, and E. Scavino, "Capacitated vehicle-routing problem model for scheduled solid waste collection and route optimization using PSO algorithm," *Waste Manage.*, vol. 71, pp. 31–41, Jan. 2018.



QI SUN received the master's degree in railway signal from Lanzhou Jiaotong University, Lanzhou, China.

He is currently a Teacher with Lanzhou Jiaotong University. His research interests include intelligent transportation and traffic information engineering and control.



HAIFEI ZHANG (Graduate Student Member, IEEE) received the master's degree in traffic information engineering and control from Lanzhou Jiaotong University, Lanzhou, China, in 2019. He is currently pursuing the Ph.D. degree in control science and engineering with Jiangnan University, Wuxi, China.

His research interests include intelligent transportation and optimization algorithm.



JIANWU DANG received the Ph.D. degree in electrification and automation of railway traction from Southwest Jiaotong University, Chengdu, China, in 1996.

He is currently a Professor, a Doctoral Supervisor, and the Vice President of Lanzhou Jiaotong University, China. He is also a National Candidate for the New Century Ten Million Talent Project and one of the first batch of special science and technology experts in Gansu Province. He has published five monographs and more than 170 academic articles. His research interests include intelligent information processing, intelligent transportation, and image processing. He is an expert with outstanding contributions from the Ministry of Railways and won the Sixth Zhan Tianyou Railway Science and Technology Award.

...