

# An Integrated Route Planning Approach for Driverless Vehicle Delivery System

Tianyang Li<sup>1,2</sup>, Zhangyi He<sup>1</sup>, and Yanling Wu<sup>3</sup>

<sup>1</sup>School of Computer Science, Northeast Electric Power University, Jilin, China.

<sup>2</sup>Jiangxi New Energy Technology Institute, Xinyu, China.

<sup>3</sup>Hebei Hanguang Industry Co., Ltd, Handan, China.

Corresponding author:

Zhangyi He<sup>1</sup>

Email address: 626372333@qq.com

## ABSTRACT

With the rapid growth of express delivery in urban areas, the use of driverless vehicles as an alternative to traditional human delivery can reduce costs and improve efficiency. The route planning of driverless vehicles is crucial in realizing autonomous navigation, which improves the working level and ensures the efficiency improvement. However, it is difficult to reasonably organize the real-time delivery taking into account several factors that influence the planning of routes, such as load capabilities, power limits and traffic conditions. To deal with this concern, we propose an integrated approach including a multistage model and improved genetic algorithm to obtain the optimal delivery plan for driverless vehicles. The experimental results in an urban scenario with a realistic delivery service show the superiority of our proposition in the delivery efficiency. The code and data are available at <https://doi.org/10.5281/zenodo.7046194>.

## INTRODUCTION

Driverless vehicles, also known as autonomous vehicles, are broadly defined as driverless robotic vehicles Kaur K (2018). They can be used in urban areas, industrial parks and various terminal distribution scenarios that require short-distance logistics, which can effectively reduce transportation costs and improve the efficiency of logistics industry. Benefiting from the advance of artificial intelligence and new energy technology, the using of driverless vehicles, especially electric driverless vehicles (EDVs) will make delivery system more clean, efficiency and security Xie G (2017) Feng J (2021). Our goal is to investigate the route planning approach for Driverless Vehicle Deliver System (DVDS) to improve distribution efficiency of the "last-kilometer" in an urban scenario.

In recent years, with the rise of online commerce, logistics industry has developed rapidly. However, current delivery system is still based on man power to deal with a large number of packages. There is a bottleneck on delivery efficiency and service quality on the "last-kilometer". Therefore, the DVDS becomes a promising way to replace manual distribution, and achieves costs reduction and efficiency improvements. That may upgrade the logistics industry with a new business model and further promote the development of online commerce.

Despite the advantages of DVDS, there are still some challenges for the improvement of delivering efficiency. These challenges mainly come from the remnant of batteries, payload, weather, traffic conditions, etc. The key to improve the delivering efficiency relies on the reasonable arrangement of distribution routes between the Distribution Center (DC) to each Express Station (ES) or to the end user. In addition, for most logistics companies, their variable costs and delivery efficiency are mainly related to the specific distribution path selected for vehicles. Obviously, providing an effective delivery path that considers all the factors including remnant of batteries, payload, weather, traffic conditions and can assist in saving cost and improving efficiency is important. Therefore, an effective route planning approach for DVDS is indispensable for the sustainable development of the logistic industry.

Typically, there are many mathematical models for solving the route planning problem, such as the model of heterogeneous Fleet Vehicle Routing Problem (HFVRP), traditional Vehicle Routing Problem (VRP) and Dynamic Vehicle Routing Problem (DVRP). Here HFVRP refers to the fact that the whole distribution task should be completed by different types of vehicles and the number and cost of each type of vehicle may be different. And DVRP is a VRP that satisfies certain dynamic constraints. However, considering to use multiple EDVs to deliver packages in a real world context, many factors, such as the payload, remaining battery capabilities, road conditions and delivery priorities, should be taken into account when solving the problem. The traditional model lacks the expression of these crucial factors. For example, the HFVRP model doesn't consider the dynamic traffic situation and the DVRP model doesn't consider the different characteristics of multiple vehicles.

Figure.1 shows an example of DVDS with multiple driverless vehicles. There are three EDVs for the DC to deliver packages to ESs. Allocating ESs for each EDV is the first and foremost problem to be solved in the consideration of the rated payload and remaining battery capabilities. That means, with the target of maximum delivery quantity, each EDV should deliver more ESs within the distance based on its remaining capacity of batteries. When suitable ESs are identified for each EDV, the route of delivery should be re-planned according to the road condition. Obviously, this problem cannot be explained by a simple model, which makes the efficient solving of this problem more difficult. This may be an obstacle for the development of delivery industry.

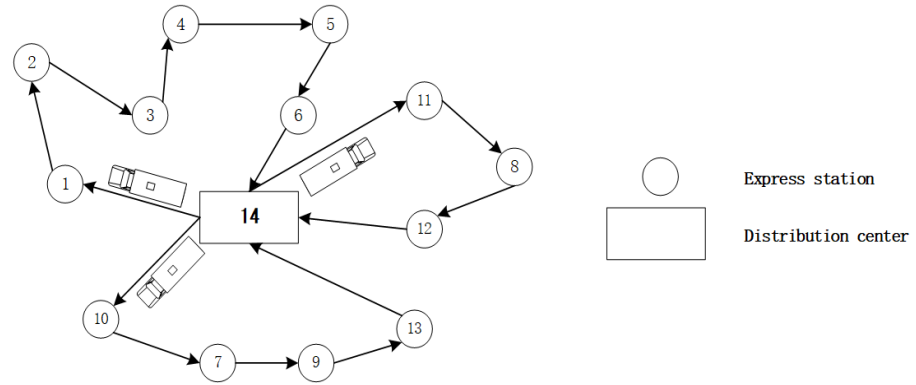


Figure 1. An example of a distribution routes.

Currently there is no comprehensive model that describes DVDS in the literature. The lack of literature precedent raises two questions: i) How might these key factors be described in a uniform model? ii) How to solve the uniform model effectively?

Therefore, an integrated approach including a multistage model and an improved GA (Genetic Algorithm) is proposed to describe and solve the route planning problem for DVDS with multiple EDVs.

The main contributions of this paper may be threefold:

- (i) designs an multistage model to describe the route planning problem with multiple EDVs;
- (ii) proposes a GA-based multi-stage optimization strategy to solve the multistage model;
- (iii) uses the data set from a real-world case to illustrate the application of proposed integrated approach, which may provide insights for delivery industry.

The paper is structured as follows. “Related Works” section introduces the related literature. “Problem Definition” section describes the model entity, multistage model, assumptions and problem definition. Then in “Improved GA for The Multistage Model” section, we introduce the improved GA and its application with the multistage model. “Experiment And Result” section validate the proposed approach with a real-world case. Our conclusions, research limitations and further developments are presented in the last section.

## RELATED WORKS

According to a traffic technical report in the United States, 94% of road traffic accidents are caused by human error, such as fatigue driving and drunk driving [Yurtsever E \(2020\)](#). If driverless cars replace conventional cars in the future, problems caused by human driving errors may be reduced. Results obtained from questionnaire survey in Germany have shown that Autonomous delivery vehicles are believed to have the potential to revolutionize last-mile delivery in a more sustainable and customer-focused way [Kapsler S \(2020\)](#).

Moreover, driverless vehicles also have many advantages. They can alleviate congestion and reduce exhaust emissions through intelligent travel strategies [Thurner T \(2022\)](#) and new energy technology. Therefore, there are lots of literature on the various areas involved in driverless driving (including Drones) for delivery. For example, the new Air and Ground Autonomous Vehicle that has lower energy consumption and carbon dioxide emission when delivering goods was discussed [A. \(2020\)](#). Autonomous vehicles travel on circular networks for pickup and delivery was also studied [Trotta M \(2022\)](#). Moreover, vehicle routing problem with drones [Brunner G \(2019\)](#) [Lei D \(2022\)](#), such as the Drone-as-a-Service (DaaS) [Shahzaad B \(2019\)](#) and Swarm-based Drone-as-a-Service (SDaaS) [Alkouz B \(2020\)](#) [Alkouz B \(2021\)](#) were discussed. These studies have widely investigated the potential of using driverless vehicles to delivery.

Obviously, driverless vehicles have got lots of attention despite they have not been put into use on a large scale. However, with the maturity of relevant technologies in the future, it is believed that they will be put into commercial use soon. Therefore, the research on driverless vehicle distribution is of great significance.

At the same time, many scholars have conducted multi-level research on vehicle routing in delivery distribution problem. Some literature have studied this problem with different models and presented many novel algorithms. Table 1 shows more details about these studies.

Table 1. Studies with Different Models

Models	Algorithms
	A developed Tabu Search algorithm <a href="#">Meliani Y (2019)</a> .
HFVRP	Exact algorithms based on logic-based Benders decomposition and a variant <a href="#">Fachini R F (2020)</a> .
	a hybrid meta-heuristic approach based on a developed Tabu Search (TS) algorithm <a href="#">Meliani Y (2022)</a> .
VRP	Genetic algorithm combined with ant colony optimization algorithm <a href="#">Veronika L (2021)</a> .
	Ant colony algorithm combined to firefly algorithm <a href="#">Goel R (2018)</a> .
	Clustering algorithm is combined with ant colony algorithm <a href="#">Z. (2020)</a> .
	Hybrid algorithm with neighborhood reduction <a href="#">He P (2022)</a> .
DVRP	Best-cost Route Crossover (BCRC) based on Genetic Algorithm <a href="#">Alwabli A (2020)</a> .
	Evolutionary algorithms(EAs) <a href="#">Sabar N R (2019)</a> .
	Discrete particle swarm optimization algorithm <a href="#">Daely P T (2021)</a> .
	A meta-heuristic algorithm of requirement coverage diversity based on ant colony algorithm framework <a href="#">Xiang X (2020)</a> .

As we all know, Tabu search (TS), artificial bee colony (ABC), ant colony optimization (ACO) and particle swarm optimization (PSO) are also meta-heuristic algorithms like genetic algorithms. Obviously, these algorithms can overcome slow convergence and excessive randomness for effectively handling periodic demands, but still require a tremendous amount of hand labour to adjust mode for coping with the random surge demands properly. When the iteration reaches a certain level, the genetic algorithm and the above algorithms will converge prematurely due to the loss of population diversity. But we can know that the effectiveness of GAs

extremely depends on genetic operators, we can achieve the goal of enriching population diversity by improving the genetic operators, so we proposed an improved mutation operator to help achieve this goal.

Meanwhile, according to above literature, we can find that evolutionary algorithms and swarm intelligence algorithm are widely used to a single model for solving VRP with certain circumstances. For example, Meliani Y (2022) uses TS algorithm to allocate loading plans for multiple vehicles under the condition that loading constraints are met. Goel R (2018) uses firefly algorithm to achieve the overall goal of minimizing the driving distance while minimizing the number of vehicles required for delivery. Xiang X (2020) is based on the ant colony algorithm framework to effectively respond to customers' dynamic requests. However, for considering more factors of DVDS, a single model is not enough and existing methods may fail to solve the more complicated problem effectively. And we know that genetic algorithm can solve the vehicle routing problem and its deformation very well and play a cohesive role among the models with chromosome as the carrier. Therefore, we propose a multistage model combined to an improved genetic algorithm to solve the distribution problem of DVDS efficiently.

## PROBLEM DEFINITION

In this section, we will introduce the proposed multistage model and define the problem in detail.

### Definition of Entities

Firstly, five entities involved in the multistage model are introduced as follows.

- (i) DC: Each DC can be represented as a triplet,  $DC = \langle dcl, edvn, on \rangle$ , where  $dcl$  represents the geographical location of DC located by GPS,  $edvn$  represents the number of EDV in DC and  $on$  represents the number of orders received by logistics companies from customers. DC is the starting node of the DVDS transportation network. Companies design reasonable distribution routes based on ESs that need to be distributed and then select the appropriate EDVs for delivery.
- (ii) EDV: Each EDV can represent a quad,  $EDV = \langle vt, edvd, lw, lv \rangle$ , where  $vt$  represents the vehicle type,  $edvd$  represents the distance that the EDV can drive under the remaining battery capacity,  $lw$  represents the loading weight of EDV,  $lv$  represents the loading volume of EDV. The EDV types can be roughly divided into small, medium and large and they have different load capacities. Big EDV has a large load and a long distance to travel, so more ESs or farther ES can be arranged for distribution.
- (iii) ES: Each ES can be expressed as a binary group,  $ES = \langle esl, t \rangle$ , where  $esl$  refers to the geographical location of ES located by GPS and  $t$  refers to the type of ES, which mainly includes express station and delivery locker. For each delivery, all the packages stored in the ES will be delivered together.
- (iv) PA: The package delivered by DVDS from DC to ES and can be expressed as a quintuple,  $PA = \langle n, pal, es, pw, pv \rangle$ , where  $n$  represents the specific package number of each PA,  $pal$  represents the receiving place of PA,  $es$  represents the ES where the package is deposited,  $pw$  represents the weight of the package and  $pv$  represents the volume of the package. All this information is an important basis for package delivery.
- (v) TN: The transportation network of a DVDS, which is a key factor in logistics and transportation, can be represented as a triplet,  $TN = \langle s, d, c \rangle$ , where  $s$  represents all nodes in the network (including ES and DC),  $d$  represents the distance between all nodes and  $c$  represents the road congestion coefficient between all nodes. We can use  $d_{ij}$  to represent the distance between station  $i$  and station  $j$  and use  $c_{ij}$  to represent the road congestion coefficient between station  $i$  and station  $j$ . In this network, the paths between nodes are bidirectional and symmetric.

163 The Multistage Model

164 Based on traditional VPR models and defined entities, we further proposed an multistage model  
165 to characterize the route planning problem of DVDS.

166 In general, multistage model has three stages. The first stage is HFVRP model, which is  
167 mainly for ESs allocation; the second stage is VRP model, which is mainly to generate basic  
168 distribution routes; the third stage is the DVRP model, which is responsible for optimizing the  
169 distribution route.

- 170 (i) Package Arrangement Model: DC has several different types of EDV available for delivery  
171 and each vehicle has different battery power and load capacity. Assuming that various  
172 types of EDVs are fully charged, there are the following rules: the battery capacity of  
173 EDV is positively correlated with the vehicle loading capacity, i.e. the farther the EDV  
174 travels, the more packages it can carry. This model mainly arranges EDV's driving routes  
175 according to the driving distance of EDV. For describing this problem, we employ the  
176 HFVRP that is used to find a fine-grained arrangement plan.
- 177 (ii) Static Route Planning Model: With the shared information including EDVs status and  
178 road condition, the optimal driving path between ESs (or end users) should be calculated.  
179 In this part, the driving route of each EDV can be obtained through solving the Static  
180 VRP model in a relatively short time.
- 181 (iii) Dynamic Route Planning Model: There may be uncertain factors in the distribution  
182 process, such as rush hour, special weather, traffic control and other situations, which  
183 may cause road congestion. If the EDV still follows the route derived from static path  
184 planning and drives into a congested road, it will increase the travel time for delivery,  
185 which ultimately affects the efficiency of delivery and users' satisfaction. Hence, a dynamic  
186 route planning model combined with Global Positioning System (GPS) is used to obtain  
187 the traffic situation between each ES and path planning, which enables the delivery path  
188 to be reasonably planned on the basis of new realistic conditions. When each EDV finishes  
189 the delivery for a certain ES and is about to drive to the next station, if congestion is  
190 found in the next road section, it will re-plan the route according to the current road  
191 conditions, so as to complete the delivery in the shortest time with small delivery distance.  
192 Since there are dynamic constraints, we can consider this problem as DVRP.

193 These three models are independent and correlated. Any single of them is not able to  
194 describe the VRP problem for DVDS. Consequently, multistage model is proper solution. For  
195 further detailing the formulation of this multistage model, there are several assumptions that  
196 are needed to introduce.

197 Assumption of The multistage Model

198 Assumptions about DC:

- 199 (i) Only consider the case of single DC for each delivery routing planning.
- 200 (ii) The time for EDV to load packages in DC and unload packages in each ES are not  
201 considered, but the driving distance and driving time of EDV with its remnant power of  
202 batteries are mainly considered.

203 Assumptions about EDV:

- 204 (i) Each route is planed for only one EDV.
- 205 (ii) Ignore the power loss generated by each EDV when it stops or waits in congested roads.
- 206 (iii) Temporary faults on EDV or delivery errors are not considered during delivery.

207 Assumptions about delivery routes:

- 208 (i) The path between any two ESs can be regarded as an undirected arc, i.e. an EDV can  
209 drive from  $ES_i$  to  $ES_j$  or from  $ES_j$  to  $ES_i$ .

210 (ii) The path between any two ESs is bidirectional but the congestion coefficient only has  
 211 one-way weight, i.e. the coefficient matrix is a symmetric matrix.

212 Our focus is to plan the route of EDVs. Under the above restrictions and assumptions, we  
 213 should make reasonable distribution arrangements.

#### 214 Problem Formulation

215 The weighted network graph is used to describe the delivery network. Given a simple undirected  
 216 graph  $G = \langle v, e, c \rangle$  with  $N + 1$  nodes, where  $v = \{v_1, v_2, \dots, v_n, v_{n+1}\}$  is the node set in the graph.  
 217  $v_{n+1}$  represents DC and the other nodes ( $v_1 - v_n$ ) present ESs. The  $e$  is the edge set of the  
 218 graph and  $c$  is the traffic congestion coefficient matrix. Other symbols are described as follows:  
 219  $K$  is the number of EDV,  $q_k$  represents the package quantity (including loading volume and  
 220 loading weight) to be loaded by the  $k$ th EDV in DC and  $Q_k$  refers to the rated maximum loading  
 221 capacity of the  $k$ th EDV (also including loading volume and loading weight).  $D_k$  represents the  
 222 maximum driving distance of the  $k$ th EDV and  $dd_k$  denotes the driving distance traveled by the  
 223  $k$ th EDV to complete all deliveries.  $T_k$  represents the time taken for the  $k$ th EDV to complete  
 224 all deliveries.  $x_{ijk}$  refers to whether the  $k$ th EDV travels from  $v_i$  to  $v_j$ , which is a variable of 0  
 225 and 1.

Aiming at minimizing the total route length and delivery time of delivery vehicles, we design  
 a uniform mathematical model to describe the multistage model. The mathematical model  
 established in this paper is as follows:

$$\min Z = \sum_{k=1}^K (dd_k + T_k) \quad (1)$$

$$s.t. \quad \sum_{k=1}^K \sum_{i=1}^{N+1} x_{ijk} = 1 \quad (j = 1, 2, \dots, N) \quad (2)$$

$$\sum_{k=1}^K \sum_{j=1}^{N+1} x_{ijk} = 1 \quad (i = 1, 2, \dots, N) \quad (3)$$

$$\sum_{j=1}^N x_{N+1,jk} = \sum_{i=1}^N x_{iN+1,k} \leq 1, \quad k \in \{1, 2, \dots, K\} \quad (4)$$

$$\sum_{k=1}^K \sum_{j=1}^N x_{N+1,jk} \leq K \quad (5)$$

$$Q_k - q_k \geq 0, \quad k \in \{1, 2, \dots, K\} \quad (6)$$

$$D_k - dd_k \geq 0, \quad k \in \{1, 2, \dots, K\} \quad (7)$$

$$x_{ijk} = \begin{cases} 1, & \text{if the } k\text{th EDV delivers from } v_i \text{ to } v_j \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

226 In the above model, the objective function (1) considers minimizing the length of delivery  
 227 path and delivery time. Constraints (2) and (3) indicate that each ES's packages can only  
 228 be delivered by one EDV, i.e. there is no such thing as multiple vehicles delivering packages



229 to the same ES. Constraint (4) means that EDV no longer return to DC to pick up packages  
 230 when performing the delivery service but pick up all packages that need to be delivered at the  
 231 beginning of EDV's delivery. Constraint (5) indicates that the number of EDVs departing from  
 232 DC doesn't exceed the number of EDVs available for transport services. Constraint (6) refers  
 233 to that the package quantity delivered by each EDV is less than the maximum loading capacity  
 234 of the EDV, which includes loading weight and loading volume. Constraint (7) indicates that  
 235 the remaining driving distance of each EDV is greater than the distance to be traveled for the  
 236 distribution of all packages. Constraint (8) indicates that for the  $k$ -th EDV, if it drives from  
 237  $ES_i$  to  $ES_j$ , then  $x_{ijk} = 1$ , otherwise  $x_{ijk} = 0$ .

238 As we can see that the uniform mathematical model is more complicated and has a large-  
 239 scale solution space. For obtaining a routing planing for all EVDs, there are at least three times  
 240 of optimization of the objective function with different constraints. Of course, an effective opti-  
 241 mization algorithm is needed. It is well known that GA can be used to all kinds of optimization  
 242 problem. But it still has its own drawbacks, such as pre-mature and easily trapped into local  
 243 optimization, especially with the large-scale solution space. Consequently, we further improve  
 244 the GA to effectively solve the uniform model with different situations.

## 245 IMPROVED GA FOR THE MULTISTAGE MODEL

246 In this section, we will use the improved genetic algorithm to solve the proposed uniform model.  
 247 Firstly, the improvements of genetic algorithm and the application of it in different stages are  
 248 detailed.

### 249 Improved GA

250 Compared with other heuristic algorithms searching from a single point, GA searches from  
 251 multiple points at the same time. However, traditional GA is easy to pre-mature when solving  
 252 large-scale problems and cannot solve the uniform model effectively, so the optimal results can  
 253 be efficiently obtained only by improving the GA.

254 GA usually consists of four stages: initial population, chromosome evaluation selection,  
 255 crossover and mutation. The latter three stages need many iterations, then the new popula-  
 256 tion will be better than the previous generation and gradually approach the global optimal  
 257 solution. Next, the improvements of GA will be introduced below.

258 Encoding Settings Genetic coding mainly adopts the set of natural number to represent the  
 259 driving routes of  $K$  EDV, numbers  $1, 2, \dots, N$  represents the ESs to be accessed by  $K$  EDVs. In  
 260 order to represent  $K$  routes with one chromosome,  $K - 1$  interruption points are added for route  
 261 segmentation and these points are  $N + 2, \dots, N + K$ . Each interruption point represents DC, so  
 262 that the chromosome can be split into  $K$  loops and each of which represents the driving route  
 263 of an EDV. A chromosome code is assigned to three EDVs and delivery routes are shown in  
 264 Figure.2, where  $1, 2, \dots, 13$  represent ESs, 14 represents DC and 15, 16 represent interruption  
 points.

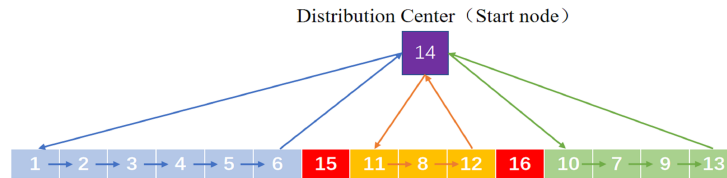


Figure 2. A chromosome code.

265

266 Selection Operation At the beginning of GA, some particularly excellent individuals are usually  
 267 produced. If the optimal individuals are selected according to the fitness value from high to  
 268 low or according to the proportion, these abnormal individuals will control the selection process  
 269 because their competitiveness is too prominent. As a result, at the later stage of evolution, these

abnormal individuals occupy the whole population and their potential for further optimization is reduced. In the end, we may not get the global optimal solution but only get the local optimal solution. Hence, we use the Tournament Selection to select individuals by randomly selecting  $n$  individuals from the entire population (usually  $n=2$ ) and pitting them against each other in a series of competitions, leaving only the best of them. This method doesn't need to sort all the fitness values, is not easy to fall into the local optimum and has a small time complexity Shukla A (2015).

**Crossover Operation** GA crossover operation exchanges some genes between two chromosomes in some way to produce a new superior individual. In this paper, Partial-Mapped Crossover (PMX) GGGoldberg D E (1985) combined with Order Crossover (OX) is used, where the probability of executing PMX is  $\delta$  and the probability of executing OX is  $1-\delta$ . Both methods involve setting up two intersections, intercepting the gene fragments between the two intersections and recombining the genes after the crossover. After the crossover operation, it can better enhance the ability of GA to open up new search space, so that the final obtained offspring will be closer to the optimal solution of the problem.

**Mutation Operation** The mutation operation can maintain the diversity of the population to some extent, increase the local search ability of the algorithm and prevent the premature convergence of the algorithm. The mutation operations adopted in this paper include exchange mutation, inversion mutation and insertion mutation. Here we transform inversion mutation and insertion mutation into selective inversion mutation (SIM) and selective insertion mutation (SIM') to change the search ability of the whole algorithm, where the probability of executing SIM is  $\mu$  and the probability of executing SIM' is  $\sigma$ .

Inversion mutation refers to the 180-degree turnover of a gene segment in the chromosome. When it is applied to VRP, it means to reverse the driving route from  $ES_A$  to  $ES_B$  to produce a new distribution route. For the operation of SIM, randomly select a  $ES_A$  then find the last  $ES_{AEnd}$  in the distribution route where the  $ES$  is located. Between  $ES_A$  and  $ES_{AEnd}$ , select a location nearest to  $ES_A$  and then reverse the whole distribution route from  $ES_{ARight}$  to  $ES_A'$  to generate a new distribution route. For example, the process of SIM is in Figure.3.

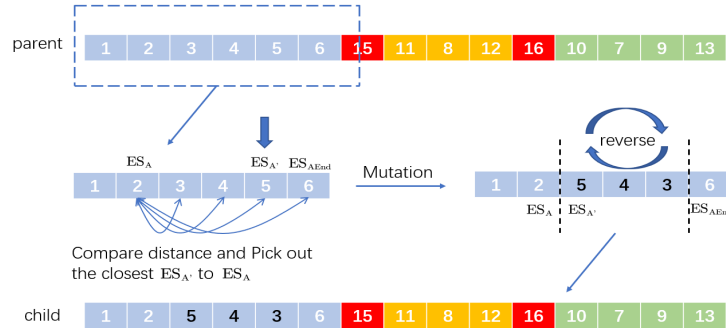


Figure 3. An example of selective inversion mutation.

We randomly select a  $ES_2$  and we know  $ES_6$  is the last station of the distribution route according to interruption point 15. If the distance between  $ES_5$  and  $ES_2$  is the shortest, the new delivery route after selective inversion mutation is 'child'.

On the whole, selective inversion mutation can find the good evolutionary characteristics of the gene sequence without disrupting the good chromosome sequence.

Insertion mutation refers to randomly taking two positions on the chromosome and then the second gene is inserted after the first one. Corresponding to the VRP,  $ES_B$  is taken as the next driving destination of  $ES_A$ . The newly proposed SIM' is to randomly select two ESs ( $ES_A$ ,  $ES_B$ ) and then find the last  $ES_{BEnd}$  in the distribution route where  $ES_B$  is located. Between  $ES_B$  and  $ES_{BEnd}$ , select a  $ES_A'$ , which is closest to  $ES_A$  and compare the distance between them and the distance between  $ES_A$  and  $ES_{ARight}$ . If the former is less than the latter, remove A' and insert it



after A to form a new distribution route; otherwise, the original distribution route will not be changed. For example, the process of SIM' is in Fig.4.

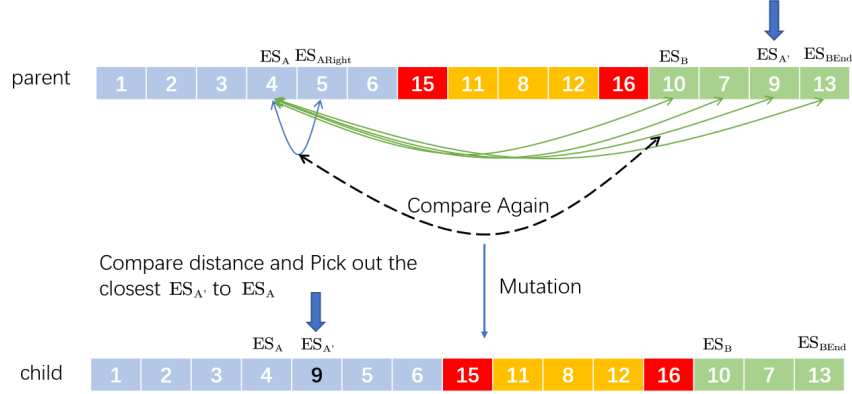


Figure 4. An example of selective insertion mutation.

We randomly select two ESs:  $ES_4$  and  $ES_{10}$ . Behind  $ES_{10}$  (including  $ES_{10}$ ), we search which station is closest to  $ES_4$ . If the distance from  $ES_9$  to  $ES_4$  is closest and less than the distance from  $ES_4$  to  $ES_5$ , the new delivery route after selective insertion mutation is 'child'.

This method expands the diversity of chromosomes to a certain extent and helps to find more effective solutions.

Next we will introduce each sub model in the multistage model in turn. The first model is centered on the farthest distance traveled by the EDV to get the ESs that EDV needs to pass through and then gives a preliminary driving route. According to the route of each EDV, each vehicle will pick up the packages to be delivered at DC. The second model is based on the first model to carry out detailed static route planning for each EDV and work out a more reasonable route to complete delivery. The third model is to obtain dynamic traffic congestion to carry out route planning, develop a more time-saving route scheme and effectively avoid the delay of delivery due to entering the congested road section.

#### The Application of Improved GA to Multistage Model

**Improved GA with Clustering Algorithm for Package Arrangement** Here, in order to solve the delivery problem in a large area efficiently, the improved GA needs to generate a better initial population and complete path planning more quickly. Considering that ESs with similar locations can be arranged to the same EDV. We employ k-means++ clustering algorithm David A (2006) to enhance the searching ability of the improved GA called "KGA-based allocation algorithm". This algorithm can group stations with similar locations and generate better individuals, which make the package arrangement more excellent in accuracy and speed.

Under complex requirements, the GA usually adopts the scheduling strategy and objective that can satisfy all demands as soon as possible. Therefore, we choose to minimize the driving distance and minimize the gap between the actual driving distance and the maximum driving distance of the EDV. The purpose is to evenly distribute delivery routes of different lengths to each EDV, i.e. the longer delivery route is allocated to EDV with longer driving distance. So we transform the multi-objective optimization problem into a single-objective optimization problem, whose objective function is defined as follows:

$$\min Z = \omega_1 * \sum_{k=1}^K dd_k + \omega_2 * \sum_{k=1}^K (D_k - dd_k) \quad (9)$$

where  $\omega_1$  can be called as the distance weight,  $\omega_2$  can be called as the "gap" weight and  $\omega_3$  can be called as time weight appearing in below. The KGA-based allocation algorithm can

334 quickly obtain the ESs that each EDV should pass through and then we can arrange the  
 335 delivery for each EDV.

Algorithm 1 shows the pseudocode of KGA-based allocation algorithm.

---

Algorithm 1: KGA-based allocation algorithm

---

Input:  $d, edvd, P, c, c\_rate, m\_rate$ .

Output:  $TDD, DR, dd_k, k \in 1, 2, \dots, K$ .

for i in range P:

    Initialize the initial population by K-means++.

j=1

while j < I:

    selection: select P excellent individuals by tournament selection.

    crossover: cross two individuals with PMX or OX to generate new individuals when  $p \leq c\_rate$ .

    mutation: mutate the individual with SIM or SIM' when  $p \leq m\_rate$ .

    Gets the distance traveled by all individuals.

    fitness evaluation: calculate the fitness of all individuals in the population.

    Record the optimal individuals of each generation.

    if fitness of the optimal individual in each generation is greater than the global optimal individual.

        Update the information of the global optimal individual.

    for m in range len(edvd):

        if  $bd[m] > edvd[m]$  then flag=False.

    if ( flag == False ) then j++.

    else I=j and return  $TDD, DR, dd_k$ .

    Update the population.

---

336 Where I is the maximum number of iteration, P is the size of population, c\_rate and m\_rate  
 337 are the crossover and mutation rates, TDD is total driving distance, i.e. sum of distances traveled  
 338 by five EDVs,  $bd[m]$  is the distance that the m-th EDV in the global optimal individual needs  
 339 to drive and DR is the driving route of all EDVs. So far, the first stage of express delivery is  
 340 completed and the detailed route planning will be realized by the second and third stages.  
 341

**Improved GA for Static Route Planning** In the phase of solving the static route planning model, stations that each EDV needs to pass through have been obtained. Next, we have to decide the routing plan for each EDV with assigned stations. Hence clustering algorithm is no longer used for population initialization. We take the driving routes of each EDV generated in the previous stage and shuffle the sequence of the routes to form a new individual.  $P$  individuals are combined into a new population, where each individual is represented as a different driving route of the EDV. It should be noted that  $d_{ij}$  is the distance from the  $ES_i$  to the  $ES_j$  and the driving speed of vehicles on the road is not only affected by road speed limit but also by traffic congestion or special weather. Therefore, we introduce the traffic congestion coefficient  $c$  to estimate the driving time (DT) taken by vehicles from station  $i$  to station  $j$ , and the formula is as follows:

$$DT_k = \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} (c_{ij} * \frac{x_{ijk} * d_{ij}}{v}), k \in \{1, 2, \dots, K\} \quad (10)$$

We mainly focus on the urban logistic distribution, where  $v$  is the driving speed of EDV in the process of distribution in the city,  $c$  indicates the congestion coefficient of the road between different ESs and  $c \in [1, 5]$ . The larger value of  $c$  is, the more serious the road congestion is and

the calculated driving time is greater than the time that it drives at a constant speed  $v$ . For this circumstance, the routing plan should take the shortest driving time as the objective condition that EDV's driving distance may be more longer for preventing driving into congested areas and ensuring that each EDV can spend the least time to complete delivery. Therefore, we also transform the multi-objective problem into a single-objective optimization problem in order to shorten the traveling time and distance at the same time and its objective function is defined as follows.

$$\min Z = \omega_1 * dd_k + \omega_3 * DT_k \quad (11)$$

342 The improved GA is used to solve the static route planning model, the pseudocode of GA-based static route planning algorithm is shown in algorithm 2.

---

Algorithm 2: GA-based static route planning algorithm

---

Input:  $d, edvd, P, c, DR, c\_rate, m\_rate$ .

Output:  $TDD, DR, dd_k, k \in 1, 2, \dots, K$ .

for  $i$  in range(len(DR)):

$pop = []$ .

$pop.append(DR[i])$ .

    for  $k$  in range(P):

        shuffle  $DR[i]$  and add it to population.

    Take  $pop[0]$  as the global optimal individual and calculate its fitness.

$j=1$ .

    while  $j < I$ :

        selection, crossover and mutation.

        Obtain driving distances of all individuals and calculate their times according to  $c$ .

        fitness evaluation: calculate fitness of individuals of population.

        Record the optimal individual of each generation.

        if fitness of the optimal individual in each generation is greater than the global optimal individual.

            Update the information of the global optimal individual.

        Update the population.

$j++$ .

    return  $TDD, DR, dd_k$ .

---

343

344 Where  $pop$  represents the whole population. This algorithm obtains the traffic congestion  
345 coefficient between each ES in the transportation network after the vehicles complete packages  
346 loading and then makes the route planning. Once the route of each EDV is determined, it will  
347 not be changed.

348 Improved GA for Dynamic Route Planning According to solving the static route planning model,  
349 a route planning is obtained for each EDV. However, road information collected at a certain  
350 time point is time-sensitive. If we only follow the distribution route obtained before departure,  
351 it's likely to cause vehicles to enter the congested road and delay the delivery time. Therefore,  
352 we want to establish dynamic route planning according to real-time traffic conditions. In order  
353 to search the distribution route dynamically, we need to consider how to deal with the dynamic  
354 model. Generally, there are two dynamic calculation methods. One is to calculate at a fixed time  
355 interval. Starting from the departure of ESs, every time interval, the system will automatically  
356 obtain the traffic information between stations in the transportation network and re-plan the  
357 route. The characteristics of this method are as follows: when the time interval is too large, the  
358 traffic information will be out of date; when the time interval is too short, the route planning  
359 will be more accurate but the calculation time will be longer.

360 The other method is to calculate at the node. After arriving at any node, the EDV will  
 361 re-obtain the traffic information in the transportation network to re-planing the distribution  
 362 route according to the remaining nodes until it finally returns to the DC. Here, we adopt the  
 363 second method and make corresponding improvements.

364 When EDV arrives at a certain  $ES_i$  and finishes unloading, it still follows the distribution  
 365 route plan made at the previous node: drive to  $ES_j$ . At the same time, according to the  
 366 comparison between  $c_{ij}$  (represents the traffic congestion coefficient between  $ES_i$  and  $ES_j$ ) and  
 367  $\beta$  (the threshold determined by experience), if  $c_{ij} < \beta$ , it means that the congestion between  
 368 the two stations is not very serious and EDV can still drive according to the distribution route,  
 369 i.e. from  $ES_i$  to  $ES_j$ . Otherwise, it is necessary to re-planing the distribution route for the  
 370 remaining distribution nodes. At this circumstance, according to the new distribution route,  
 371 the EDV should drive from  $ES_i$  to  $ES_k$  (other node with a lower congestion coefficient). The  
 372 objective function of this model is consistent with the static route planning model.

373 The improved GA is used to solve the dynamic route planning model, the pseudocode of  
 GA-based static route planning algorithm is shown in algorithm 3.

---

Algorithm 3: GA-based dynamic route planning algorithm

---

Input:  $d, edvd, P, c, DR, c\_rate, m\_rate$ ;

Output:  $TDD, DR, dd_k, k \in 1, 2, \dots, K$ ;

for  $i$  in range(len(DR)):

    chrome=DR[i].

    for  $k$  in range(len(chrome)):

        Obtain real-time road congestion coefficient matrix: conMatrix.

        if (conMatrix[chrome[k]-1,chrome[k+1]-1] <=  $\beta$ ):

            continue.

        pop = [ ].

        Initialize the population.

        j=1

        while  $j \leq (\text{len}(\text{chrom})-k)*2$ :

            selection, crossover and mutation

        Obtain driving distances of all individuals and calculate their driving times according  
 to  $c$ .

        fitness evaluation: calculate the fitness of all individuals in the population.

        if fitness of the optimal individual in each generation is greater than the global optimal.

            Update the information of the global optimal individual.

        Update the population.

        j++

        obtain the next stop of the  $i$ -th EDV.

    obtain the  $DR_i$  of the  $i$ -th EDV.

return  $TDD, DR, dd_k$

---

374  
 375 Where chrome represents an individual in the population. This algorithm re-plans the route  
 376 when the vehicles encounter serious traffic jams until each EDV completes the delivery.

## 377 EXPERIMENTS AND RESULTS

378 In this section, we will evaluate the performance of the integrated approach that includes both  
 379 the proposed multistage model and the improved GA. For the former, we mainly evaluate it  
 380 according to the total driving distance and driving time of EDVs, while for the latter, we mainly  
 381 evaluate it according to the average number of iterations of the algorithm. We conduct a series

of experiments based on a dataset collected from a real-world delivery services. The details of dataset is shown in table 2. There are one DC for sorting and temporarily storage packages. 1000 ESs randomly around the DC. Five EDVs are assigned to the DC for delivery. We assume that the average driving speed of EDV in the city is 30km/h and the maximum driving distance of these five EDVs is different, which are 2000, 2200, 1900, 2400 and 1800. We take the threshold of traffic congestion coefficient as 3, which means that when the congestion coefficient of a road is greater than the threshold value, path planning will be redone.

Table 2. Details Of Synthesized Data Set

Data Item	Value
The number of ES	1000
The number of DC	1
The number of EDV	5
Average speed of EDV	30 km/h
Threshold of traffic congestion coefficient	3
Maximum driving distance of five EDVs	[2000,2200,1900,2400,1800]

Three experiments were implemented using Python language on personal notebook: AMD Ryzen 7 4800U with Radeon Graphics (1800MHz) @16GB memory and 64-bit. For each experiment, we analyze the results according to the total distance and driving time of EDVs in each model. Meanwhile, finding the best solution in finite iterations is also a main criterion for evaluating the effectiveness of the algorithm. We did the following experiments mainly according to each stage of the multistage model in order to better reflect the progressive relationship between each layer.

In the first experiment, we compare the performance difference between ordinary genetic algorithm and improved genetic algorithm in route allocation. Given the maximum distance of each EDV, we compare which method can allocate the driving route of each EDV in the shortest time (the least number of iterations). Due to the different EDVs have different maximum driving distance, the allocated area blocks (clusters classified according to the clustering algorithm) should also be different. The result is shown in Figure.5.

As shown in the figure, under the condition of the maximum driving distance of five EDVs :2000, 2200, 1900, 2400 and 1800, the results of one experiment are shown as follows: 70 iterations based on the improved genetic algorithm can meet the requirements of the condition, while it takes 202 iterations based on the traditional genetic algorithm to determine the delivery stations for each EDV. According to the experimental results, the average time of the first stage based on the improved genetic algorithm is 12.14 seconds, while the average time based on the traditional genetic algorithm is 41.3 seconds. The traditional genetic algorithm only uses OX for exchange operator, insertion mutation and inversion mutation for mutation operator, but it's also based on k-means++ algorithm. If the traditional GA is used alone, thousands of iterations are required. It can be seen that the improved genetic algorithm has better convergence effect.

In the second experiment, we optimize the allocation route derived from the improved GA in experiment 1 based on the traditional VRP model. This experiment proves that the algorithm we use can optimize the distribution route well and ensure that the driving distance of each EDV is as short as possible under the shortest driving time. Optimized routes are shown in Figure.6.

As shown in Fig.6, route optimization was carried out for each EDV and the delivery distance was greatly reduced. Compared with the total driving distance of each EDV in experiment 1, the final total driving distance was reduced by 66.7% by optimizing the delivery route of each EDV in experiment 2. The parameters of the objective function and improved genetic algorithm are shown in Table 3.

In the second experiment, we performed 5000 iterations for each EDV route update, taking

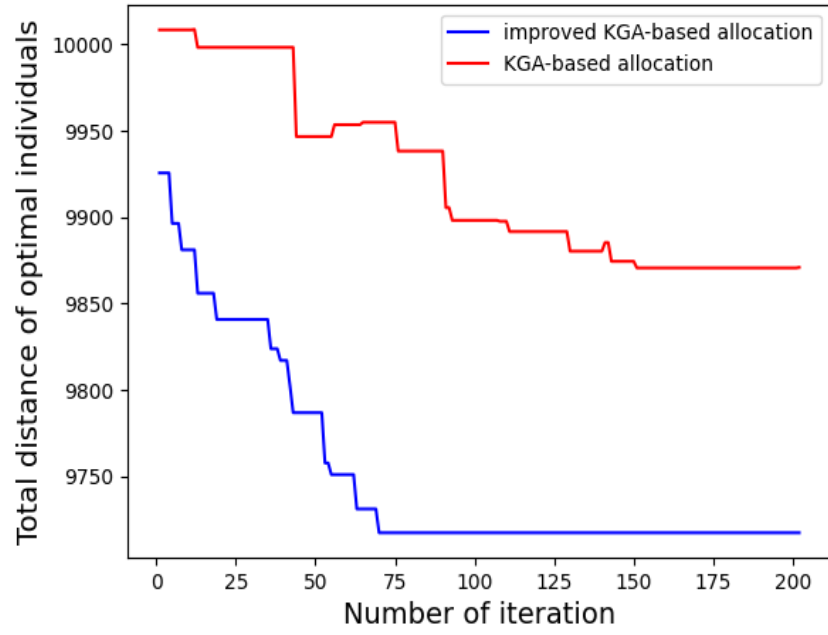


Figure 5. Total distance of different algorithm.

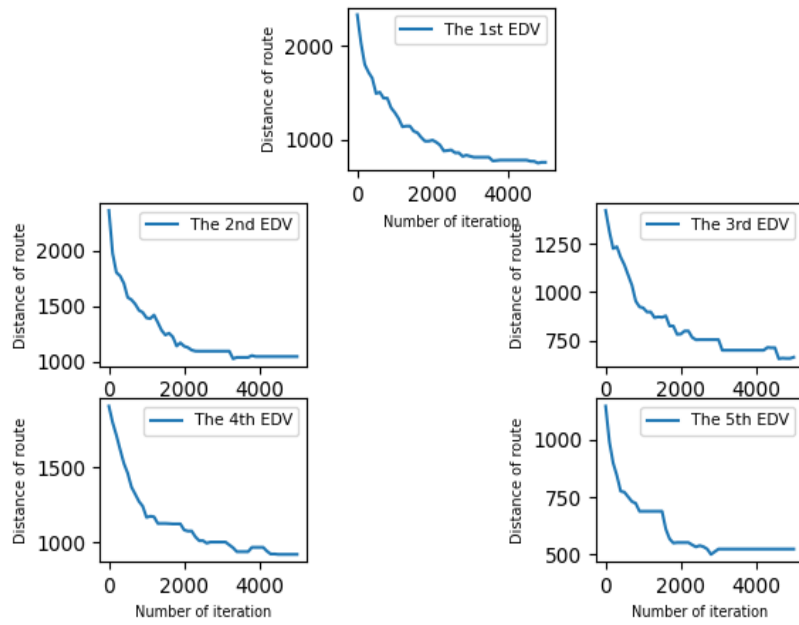


Figure 6. Optimization of different distribution routes.



Table 3. Settings Of Object Function and Algorithm

Variable	Value
Objective function	
$\omega_1$ , $\omega_2$ and $\omega_3$	1, 1500 and 250
Improved Genetic algorithm	
mutation rate	0.25
crossover rate	0.8
population size $P$	30
PMX rate $\delta$	0.5
SIM and SIM' rate $\mu$ , $\sigma$	0.5, 0.5
maximum iteration number $I$	1000 and 5000

an average of 230 seconds. However, if the experiment is done on a computer with higher computing power, it is believed that the time will be greatly reduced. Since the algorithm has a good effect in static model, so we continue to study the effect of the algorithm in dynamic model.

In the third experiment, we adopt dynamic VRP model to deal with changing traffic conditions. The route planning shall be carried out again after the vehicles reach each ES and the final results (delivery time and delivery distance) shall be compared with the results obtained from the static route planning.

In order to calculate the delivery time under the static model (HFVRP+SVRP), we use the delivery route of each EDV obtained in experiment 2. Because the delivery route has been fixed, so we only need to obtain the traffic congestion coefficient (We use random numbers to replace the traffic congestion coefficient obtained by GPS) between the  $ES_A$  and the next  $ES_B$  after the EDV arrives at  $ES_A$  and then calculate the time required to reach the next  $ES_B$ . Finally, we add up each section of driving time to get the total delivery time. The time calculation of dynamic model (HFVRP+SVRP+DVRP) also accumulates the driving time between the two ESs, which is similar to the static model, but the distribution path will be adjusted dynamically. In addition, we calculate the driving distance of each EDV in the two models and analyze it with the delivery time. By comparing delivery distance and time, the results of dynamic model have both reduced in some different degree. Specifically, the delivery distance has an average decrease of 7.54% and delivery time declines 19.9%. These results are shown in Table 4.

Because the DC receives the packages frequently, in order not to pile too many packages, it is necessary to arrange new delivery in time and deliver the packages to the customers as early as possible. Therefore, we prefer EDV to finish each delivery earlier, so as to start the next delivery earlier. It can be seen from the results that the driving time decreased significantly and the driving distance as a whole shows a downward trend. This shows that the current algorithm will still fall into local optimization when there are too many data points and a small number of iterations, so we apply the algorithm to the dynamic model to obtain the optimal solution by continuously reducing the size of the dataset. In the third stage, there are only few factors that need to consider, which makes it easier to achieve global optimization on the basis of the optimized route when re-planning the route. Obviously, the results can meet our goal, that is to complete delivery in the shortest time.

At the end of the experiments, all the results were summarized and the results of the three stages of route planning were displayed. It can be seen that after route planning, the delivery route was greatly reduced and the delivery time was also greatly reduced, as shown in Figure.7.

Table 4. Experimental Results

Lines	Driving Distance			Estimation of driving time		
	Static model	dynamic model	Proportion of reduce	Static model	dynamic model	Proportion of reduce
Distribution line1	541.4	484.6	10.5%	51.7	43.9	15.1%
Distribution line2	722.6	735.1	-	72.7	63.9	12.1%
Distribution line3	599.7	617.4	-	62.1	53.7	13.5%
Distribution line4	913.9	717.4	21.5%	91.6	68.4	25.3%
Distribution line5	547.1	405.4	25.9%	55.0	36.6	33.5%

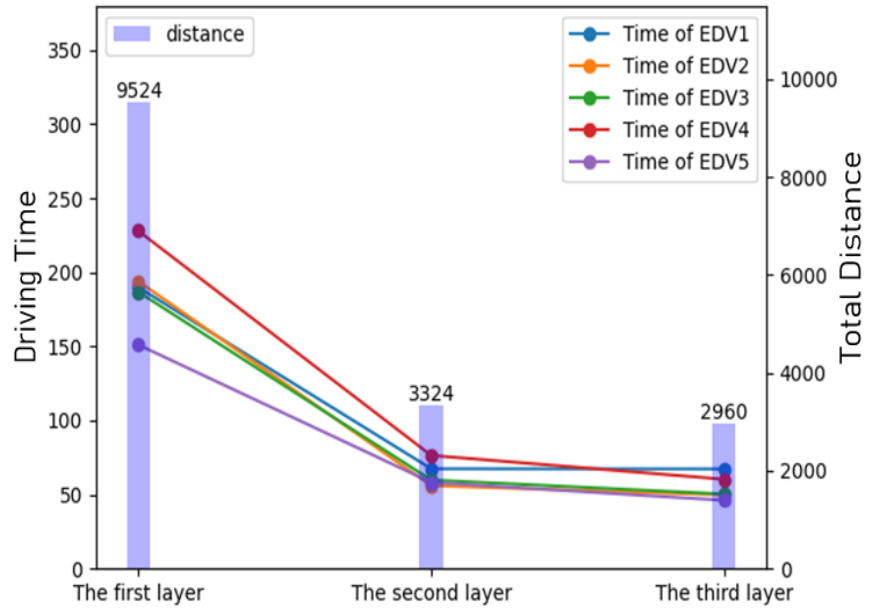


Figure 7. The result of optimization of different stages.

## 458 CONCLUSIONS

459 With the rapid development of the logistics industry and the the rise of online commerce, using  
460 DVDS is a promising way to improve delivery efficiency and users' satisfaction. In this paper,  
461 we propose an integrated approach with a multistage model and an improved genetic algorithm  
462 to solve the problem of distribution route planning for DVDS. To be specific, the model in the  
463 first stage is to determine the package arrangement, then generate a preliminary distribution  
464 route in the second stage and finally optimize the distribution route plan for each EDV in the  
465 third stage considering the traffic congestion situation. In short, the goal of this study is to  
466 improve delivery efficiency with driverless vehicles, i.e. to complete the delivery in the shortest  
467 time. Experimental results have shown that our proposal can improve the delivery efficiency.  
468 This work may provide some practical guidance and valuable insights for designers of delivery  
469 services.

470 In the future work, we will consider the utilization of real-time GPS information, improve  
471 the efficiency and performance of the algorithm, and achieve more accurate route planning.

## 472 ACKNOWLEDGMENT

473 This work has been supported by the Scientific Research Funds of Northeast Electric Power  
474 University (No.BSZT07202107).

## 475 REFERENCES

- 476 A., F. M. (2020). Carbon emissions reductions in last mile and grocery deliveries utilizing air and  
477 ground autonomous vehicles. Transportation Research Part D: Transport and Environment,  
478 85:102443.
- 479 Alkouz B, B. A. (2021). Provider-centric allocation of drone swarm services. In 2021 IEEE  
480 International Conference on Web Services (ICWS), pages 230–239.
- 481 Alkouz B, Bouguettaya A, M. S. (2020). Swarm-based drone-as-a-service (sdaas) for delivery.  
482 In 2020 IEEE International Conference on Web Services (ICWS), pages 441–448.
- 483 Alwabli A, K. I. (2020). Dynamic route optimization for waste collection using genetic algorithm.  
484 In 2020 International Conference on Computer Science and Its Application in Agriculture  
485 (ICOSICA), pages 1–7.
- 486 Brunner G, e. a. (2019). The urban last mile problem: Autonomous drone delivery to your  
487 balcony. In 2019 International Conference on Unmanned Aircraft Systems (ICUAS), pages  
488 1005–1012.
- 489 Daely P T, e. a. (2021). Dynamic vrp optimization using discrete pso in edge computing envi-  
490 ronment. In 2021 International Conference on Information and Communication Technology  
491 Convergence (ICTC), pages 654–656.
- 492 David A, S. V. (2006). k-means++: The advantages of careful seeding. Technical Report  
493 2006-13, Stanford InfoLab.
- 494 Fachini R F, A. V. A. (2020). Logic-based benders decomposition for the heterogeneous fixed fleet  
495 vehicle routing problem with time windows. Computers & Industrial Engineering, 148:106641.
- 496 Feng J, e. a. (2021). Research on path planning and control of driverless logistics train. In 2021  
497 5th CAA International Conference on Vehicular Control and Intelligence (CVCI), pages 1–6.
- 498 GGoldberg D E, L. R. (1985). Alleleslociand the traveling salesman problem. In Proceedings  
499 of the 1st International Conference on Genetic Algorithms, page 154–159, USA. L. Erlbaum  
500 Associates Inc.
- 501 Goel R, M. R. (2018). A hybrid of ant colony and firefly algorithms (hafa) for solving vehicle  
502 routing problems. Journal of Computational Science, 25:28–37.
- 503 He P, H. J. K. (2022). Hybrid search with neighborhood reduction for the multiple traveling  
504 salesman problem. Computers & Operations Research, 142:105726.
- 505 Kapser S, A. M. (2020). Acceptance of autonomous delivery vehicles for last-mile delivery in  
506 germany—extending utaut2 with risk perceptions. Transportation Research Part C: Emerging  
507 Technologies, 111:210–225.
- 508 Kaur K, R. G. (2018). Trust in driverless cars: Investigating key factors influencing the adoption  
509 of driverless cars. Journal of Engineering and Technology Management, 48:87–96.

510 Lei D, Cui Z, L. M. (2022). A dynamical artificial bee colony for vehicle routing problem with  
 511 drones. Engineering Applications of Artificial Intelligence, 107:104510.  
 512 Meliani Y, e. a. (2019). A developed tabu search algorithm for heterogeneous fleet vehicle routing  
 513 problem. IFAC-PapersOnLine, 52(13):1051–1056.  
 514 Meliani Y, e. a. (2022). A tabu search based approach for the heterogeneous fleet vehicle routing  
 515 problem with three-dimensional loading constraints. Applied Soft Computing, 126:109239.  
 516 Sabar N R, e. a. (2019). A self-adaptive evolutionary algorithm for dynamic vehicle routing  
 517 problems with traffic congestion. Swarm and evolutionary computation, 44:1018–1027.  
 518 Shahzaad B, e. a. (2019). Composing drone-as-a-service (daas) for delivery. In 2019 IEEE  
 519 International Conference on Web Services (ICWS), pages 28–32.  
 520 Shukla A, Pandey H M, M. D. (2015). Comparative review of selection techniques in genetic  
 521 algorithm. In 2015 International Conference on Futuristic Trends on Computational Analysis  
 522 and Knowledge Management (ABLAZE), pages 515–519.  
 523 Thurner T, Fursov K, N. A. (2022). Early adopters of new transportation technologies: At-  
 524 titudes of russia’s population towards car sharing, the electric car and autonomous driving.  
 525 Transportation Research Part A: Policy and Practice, 155:403–417.  
 526 Trotta M, e. a. (2022). Pickup and delivery problems with autonomous vehicles on rings.  
 527 European Journal of Operational Research, 300(1):221–236.  
 528 Veronika L, e. a. (2021). A case study of vehicle route optimization. CoRR, abs/2111.09087.  
 529 Xiang X, e. a. (2020). Demand coverage diversity based ant colony optimization for dynamic  
 530 vehicle routing problems. Engineering Applications of Artificial Intelligence, 91:103582.  
 531 Xie G, e. a. (2017). Vehicle trajectory prediction by integrating physics-and maneuver-based  
 532 approaches using interactive multiple models. IEEE Transactions on Industrial Electronics,  
 533 65(7):5999–6008.  
 534 Yurtsever E, e. a. (2020). A survey of autonomous driving: Common practices and emerging  
 535 technologies. IEEE access, 8:58443–58469.  
 536 Z., C. (2020). Research on multi-center route planning based on improved ant colony algo-  
 537 rithm. In 2020 19th International Symposium on Distributed Computing and Applications  
 538 for Business Engineering and Science (DCABES), pages 24–27.