Research for Intra-city Distribution Based on Improved Genetic Algorithm

Haoyu Zhen*[1,a], Xiang Xie[1,b]

[1] School of Economics and Management, Beijing Jiaotong University, Beijing 100044，China
*[a]21120631@bjtu.edu.cn, [b]xxie@bjtu.edu.cn
Corresponding Author : Haoyu Zhen

ABSTRACT

Aiming at the situation of strong timeliness and short delivery distance in intra-city delivery, a vehicle route optimization model with time window under intra-city delivery is designed. The model aims to minimize the delivery cost and takes delivery time and vehicle load as constraints. This problem is an NP-hard problem and cannot be solved accurately. Therefore, this paper uses genetic algorithm to solve the problem and improves the problem that the local search ability of genetic algorithm is not strong. The maximum retention principle is selected during the intersection, and a large neighborhood search algorithm is added to improve the quality of the algorithm. Finally, the algorithm test is carried out, and it is found that the improved algorithm has better results and higher stability of the solution.

Keywords: *Intra-city Distribution; Vehicle Route Optimization; Improved Genetic Algorithm*

1. INTRODUCTION

With the development of technology, online shopping has become one of the main shopping methods for people. Affected by the epidemic, the logistics between cities is sometimes seriously affected, which has promoted the development of intra-city logistics to a certain extent. Intra-city distribution is also called urban logistics and is known as the "last mile" in logistics. It refers to providing regular and quantitative distribution services for single or multiple commodities within a specified city by optimizing various distribution elements according to customers' order needs. The transportation distance of intra-city distribution is relatively short, which is limited by urban traffic and transportation network [1], and is dominated by user needs, and there are strict restrictions on delivery time.

The VRPTW (Vehicle Routing Problems With Time Windows) model adds a time limit to the traditional vehicle routing problem. VRPTW has two types of time constraints, hard time window constraints and soft time window constraints. The hard time window constraint means that there is a strict limit on the delivery time, if the delivery time is not satisfied, the service will be refused, and the delivery scheme is an infeasible solution. The soft time window constraint sets a certain limit on the delivery time. When the arrival time is earlier or later than the time expected by the customer, a certain penalty will be given, but the delivery scheme is still within the search range of feasible solutions.

At present, intra-city distribution has problems such as low distribution efficiency and poor customer satisfaction. Few scholars have studied the path planning problem of intra-city distribution. Therefore, according to the transportation characteristics of intra-city distribution, this paper establishes a vehicle routing optimization model with soft time window, designs an improved genetic algorithm to solve it, and conducts numerical experiments to verify the effectiveness and stability of the improved algorithm.

2. MATHEMATICAL MODEL

2.1. Problem Description

Intra-city distribution has the requirements of timeliness and is limited by the urban traffic network. It is required to meet the requirements of the time window during the distribution process, including the time to arrive at each customer node during the distribution process, the time to wait for the customer to pick up the goods, and the final completion of the work. Time to return to the distribution center. The problem of this paper can be described as follows: in the way of intra-city distribution, there are several distribution demand points

A. El-Hashash et al. (Eds.): IEIT 2022, ACSR 100, pp. 889–894, 2023.
https://doi.org/10.2991/978-94-6463-058-9_139

within a certain city, each demand point has corresponding restrictions on the distribution time, and the location of the distribution center and each demand point is known, the distribution demand is known, so it is required to reasonably arrange and dispatch vehicles, traverse the demand points and meet the distribution demand, so as to minimize the driving cost under the condition of meeting various constraints.

In the actual delivery process, the delivery task may be disturbed by other emergencies, so this paper makes the following assumptions:

(1) There is only one distribution center, and the distribution center is large enough to meet all customer needs.

(2) The location of the distribution center is known, and there is special sorting personnel. Regardless of the time required for sorting, the distribution personnel only need to carry out the distribution according to the established route and requirements.

(3) All distribution points' demand, location, and time window requirements are known and will not change during the distribution process.

(4) All delivery vehicles are the same, including load capacity, model, fuel consumption per unit time, etc., and the vehicle load does not consider the type and size of the commodity.

(5) During the delivery process, all vehicles keep the same speed and drive at a uniform speed

(6) Each demand point can only be served by one vehicle, and it can be delivered at one time.

(7) All vehicles start from the distribution center and finally return to the distribution center, and all vehicles cannot exceed the load during the distribution process.

(8) The delivery process only considers delivery and does not consider reverse logistics.

(9) The distribution process does not consider road conditions or emergencies.

2.2. Parameter Description

λ: Consumption cost per unit mileage;

V_{ijs}:0-1 variable, judging whether the vehicle S goes from customer i to customer j;

d_{ij}: Euclidean distance from customer i to j;

β: Time penalty cost weight

F_{is}: 0-1 variable, judging whether vehicle s serves customer i;

F_{js}: 0-1 variable, judge whether vehicle s serves customer j;

w_i: the demand of customer i;

C: the load capacity of the vehicle;

D: the maximum mileage of the vehicle;

T_i: the time when the vehicle arrives at customer i;

a_i: the left time window that the customer expects to arrive;

b_i: the right time window that the customer expects to arrive;

V_{0js}: 0-1 variable, judging whether the vehicle departs from the distribution center to customer j;

V_{i0s}: 0-1 variable, judging whether the vehicle departs from customer i to return to the distribution center;

I: I={i | i=1,2,3…,N} customer node with delivery demand;

J: J={j | j=1,2,3…,N} customer node with delivery demand;

S: S={s | s=1,2,3…,S} the set of vehicles;

2.3. Model Building

$$\min Z=\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda V_{ijs}d_{ij}+\beta\sum_{i=1}^{N}\left|\min\left[T_i-a_i,0\right]\right|+\left|\min\left[b_i-T_i,0\right]\right| \tag{1}$$

$$\sum_{s=1}^{S}\sum_{i=1}^{N}V_{ijs}=1 \tag{2}$$

$$\sum_{s=1}^{S}\sum_{j=1}^{N}V_{0js}=\sum_{s=1}^{S}\sum_{j=1}^{N}V_{j0s}\le S \tag{3}$$

$$\sum_{i=1}^{N}F_{is}w_i\le C \tag{4}$$

$$\sum_{i=1}^{N}\sum_{j=1}^{N}V_{ijs}d_{ij}\le D \tag{5}$$

$$a_i\le T_i\le b_i \tag{6}$$

$$\sum_{j=1}^{N}V_{ijs}=F_{is} \tag{7}$$

$$\sum_{i=1}^{N}V_{ijs}=F_{js} \tag{8}$$

$$\sum_{i=1}^{N}V_{iks}-\sum_{j=1}^{N}V_{kjs}=0 \tag{9}$$

$$V_{ijs}=\begin{cases}1 & From\ customer\ i\ to\ j\\ 0 & else\end{cases} \tag{10}$$

$$F_{is}=\begin{cases}1 & Customer\ i\ is\ served\ by\ s\\ 0 & else\end{cases} \tag{11}$$

$$i,j\in N,\ \ s\in S \tag{12}$$

Formula(1) represents the objective function, which is the minimum value of the weighted sum of the travel cost and the time penalty cost. Formula(2) means that every customer is guaranteed to be served. Formula (3) indicates that all vehicles need to depart from the distribution center and return to the distribution center, and the number of vehicles used is not allowed to exceed the available quantity. Formula(4) indicates that each vehicle cannot be overloaded. Equation (5) indicates that the mileage of each vehicle cannot exceed the maximum mileage of the vehicle. Formula (6) indicates that the arrival time of the vehicle should be within the time range specified by the customer. Formula (7) to Formula (8) indicate that if i and j are service points of vehicle s, they are both served by *s*. Formula (9) indicates that *i-k* and *k-j* are served by the same vehicle. Formula (10),(11),(12) define the ranges of V_{ijs}, F_{is}, *i, j, s*.

3. Algorithm Design

The vehicle path optimization problem with time window studied has high complexity and strong real-time requirements and has relatively high requirements for the global search ability and convergence speed of the algorithm. The genetic algorithm has good global search ability and fast search speed, but it is easy to fall into the local optimal solution. Therefore, this paper adopts the improved genetic algorithm to solve the problem. The specific improvement design is as follows:

(1) Chromosome coding

The encoding in this paper adopts the natural number encoding method, 0 represents the distribution center, and the customers are randomly sorted from 1 to N. Encode the vehicle's driving trajectory. For the convenience of encoding use a larger number to encode to separate different paths. Using this encoding method, each customer can be guaranteed to be served, and will not repeat.

(2) Initialize the population

Randomly select a customer, assign the customer to the first car, traverse the customer, and fill in the arrangement of the vehicle in the order of the left time window until the capacity limit of the vehicle or the maximum driving mileage limit of the vehicle is reached. Then change to the next vehicle until all customers have arranged for the vehicle. When these paths are arranged together, an initial solution is formed. Then the iterative cycle is carried out until the initial population is formed.

(3) Fitness function

The larger the fitness function is, the better the individual is, and the objective function is to find the minimum value of the distribution cost and the time penalty cost. In the evolution process, it is necessary to select individuals with large fitness, so the reciprocal transformation of the objective function is used as the fitness function.

(4) Selection operator

In this paper, the roulette operator is used to select chromosomes and select individuals with large fitness.

(5) Crossover operator

The VRPTW problem is subject to time constraints and capacity constraints. In this paper, the maximum retention crossover method is selected to cross the chromosomes. The selected individuals were paired with each other and randomly generated two numbers s and e between 1 ~ N + K-1. The s-e position gene in the two chromosomes was retained and deleted, which should ensure that the distribution center position remained unchanged, and then the deleted part of the other chromosome was inserted in sequence to form offspring. Specific operations are as follows, 7,8 represent the location of the distribution center.

parent 1	1	2	7	3	4	8	5	6

parent 2	3	2	7	1	5	8	4	6

Randomly select the intersection position s=3, e=6, and keep the distribution center 7 unchanged. The genes except 7, 3, 4, and 8 in parent 2 are filled into offspring 1, and the genes in parent 1 except 7, 1, 5, and 8 are filled into offspring 2.

Offspring1	2	1	7	3	4	8	5	6

Offspring2	2	3	7	1	5	8	4	6

Figure 1: Schematic diagram of chromosome crossover

(6) Mutation operator

The mutation operator adopts the method of inversion mutation, randomly selects two gene positions a and b in the individual and flips the gene segments from a to b to form a new individual after mutation.

(7) Local search

In order to solve the problem that the local search ability of the genetic algorithm is not strong, a large-scale neighborhood search algorithm is added to the algorithm, which mainly includes removal and repair operators [4][6]. It is mainly based on the correlation of the distance between customers.

1) Remove

Randomly select a customer a and remove it as the first element in the removal set A. Establish a correlation array between the remaining element set B and customer a, and select the customer with the greatest correlation to add it to the removal set A. Then randomly select a customer from the removal set A to perform the corresponding cycle until the number of elements in the removal set reaches the given requirement. The calculation formula of the correlation is as formula(13) where d_{ij}' is the normalized value of the distance function d_{ij}.

$$R_{ij} = \frac{1}{d'_{ij} + V_{ijs}} \tag{13}$$

2) Repair

Find all the locations in the set A where the customer satisfies the constraints in the new path scheme, traverse all the locations. Calculate the path distance after the customer is inserted, make a difference with the original distance, and find the distance increment after the customer is inserted in all locations. Then find the smallest distance increment and record it. Then find the maximum minimum distance increment of all customers in the set A, insert it into the path, and delete it from the removed customer set A, iterative cycle until there is no element in the set A.

4. Simulation Example

In order to verify the effectiveness of the algorithm, the algorithm needs to be tested. According to the characteristics of short distribution distance and strong timeliness in the same city, this paper selects the standard data of the Solomon RC1 series to test the algorithm [3][5]. All tests are implemented through MATLAB 2017 programming.

4.1. Algorithm Parameter Value Test

The size of the population and the number of iterations limit the diversity of the population and the efficiency of the solution. Crossover and mutation affect the diversity of the population and the efficiency of the solution. Since the mutation probability is small, it is directly set to 0.05. In order to determine the optimal combination of the other three parameters, tests were carried out using Solomon standard data RC101. An orthogonal experiment with three factors and two levels was designed, and the factors and levels are shown in Table 1.

Table1 Factors and levels

	A Population size	B Number of iterations	C Poor probability
1	100	100	0.8
2	500	500	0.95

Because it is an orthogonal experiment with three factors and two levels, the L_8 (2^7) orthogonal table is selected for experimental design. Eight experiments were conducted for 20 times, and the average value was obtained as the test result. The average values of the two levels under the six factors A, B, AB, C, AC and B are calculated respectively to obtain the values of K1 and K2 under this factor. The difference between the two factors is made to obtain the range of each factor level. The results are shown in Table 2.

Table 2 Range analysis

factor	A	B	AB	C	AC	BC
K1	1723	1743	1714	1682	1698	1679
K2	1677	1658	1678	1710	1694	1713
R	46	85	36	28	4	34

The result of extreme difference shows that among the three factors, factor B has the greatest influence on the optimal value, and the result is better when factor B is at level 2, followed by factor A, which is better when factor B is at level 2 and finally factor C. Based on the above analysis, the optimal combination is $B_2A_2C_2$. The final parameter selection is shown in Table 3.

Table 3 Parameter value table of genetic algorithm

Parameter	Parameter meaning	Value
NIND	population size	500
MAXGEN	max iterations	500
Pc	crossover probability	0.95
Pm	mutation probability	0.05
GGAP	generation gap	0.9

4.2. Results Analysis and Comparison

To verify the effectiveness of the improved algorithm, a simple comparative analysis is made between the traditional algorithm and the improved algorithm. The two algorithms both use the same coding strategy, and both use the flip mutation method when mutating. The difference is that the traditional algorithm uses mapping crossover, the improved algorithm chooses the principle of maximum retention, and adds a local search algorithm. Under the condition that all parameters are the same, the two algorithms are randomly tested 20 times respectively, and the obtained results are shown in Table 4.

Table 4 Comparison results

Type	optimal value	worst value	differen ce	average value
Basic	1973.13	2366.95	393.82	2202.17
Improve	1601.01	1688.69	87.68	1640.62
Effect	372.12	678.26	306.14	1029.74

When using the traditional algorithm to solve, the quality of the solution is poor, and the stability of the solution is insufficient and the difference is large. However, by using the improved algorithm to solve the problem, the quality of the solution will be better and the stability will be stronger. Using the improved genetic algorithm, the optimal driving distance is reduced by 18.%; the driving distance of the worst solution is reduced by 28.7%; the average value is reduced by 25%; the fluctuation range is reduced by 306.14, indicating that the optimal solution obtained by the improved algorithm has less fluctuation, more stable. Compared with the traditional genetic algorithm, the improved algorithm reduces the distribution cost, improves the distribution efficiency, and achieves a better purpose, which further illustrates the superiority of the improved algorithm.

5. Conclusion

This paper analyzes the characteristics of intra-city distribution and establishes a vehicle routing model with time window under intra-city distribution. The traditional genetic algorithm is improved, and the maximum retention principle is selected in the crossover to reduce the occurrence of infeasible solutions and improve the search efficiency; a large-scale neighborhood search algorithm is added, and the distance correlation between customers is used to improve the current solution and reduce the number of infeasible solutions. The emergence of the algorithm improves the solution quality of the algorithm. Then, the important parameters in the model are determined by orthogonal tests. Finally the basic data is used to test to verify the effectiveness of the algorithm, which provides a decision basis for route planning in intra-city distribution.

REFERENCES

[1] Crainic, T.G., Ricciardi, N., and Storchi, G. (2009).Models for Evaluating and Planning City Logistics Systems. Transportation Science 43(4), 432-454.

[2] Mohammed, M. A. , Gani, M. A. , Hamed, R. I. , Mostafa, S. A. , Ahmad, M. S. , & Ibrahim, D. A. . (2017). Solving vehicle routing problem by using improved genetic algorithm for optimal solution. Journal of Computational Science, 255–262.

[3] N. Zhu and C. Shao, "Vehicle Routing Problem with Simultaneous Delivery and Pick-up Based on the Improved Genetic Algorithm," 2010 Fourth International Conference on Genetic and Evolutionary Computing, 2010, pp. 312-316, doi: 10.1109/ICGEC.2010.84.

[4] S. Hao, "Electric Vehicle Routing Problem with Soft Time Window and Weight-related Discharging Using Adaptive Large-Scale Neighborhood Search Algorithm," 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT), 2021, pp. 711-715, doi: 10.1109/ICCASIT53235.2021.9633428.

[5] Solomon M.M (1987). Algorithms for the vehicle routing and scheduling problems with time window constraints. Operations Research, 35, 254–265.

[6] Stenger, A., Vigo, D., Enz, S., Schwind, M., 2013. An adaptive variable neighborhood search algorithm for a vehicle routing problem arising in small package shipping. Transp. Sci. 47 (1), 64–80.