

METHODS

Application of genetic algorithm, a mega-heuristic approach, to solve a real-size vehicle routing problem: a case study

Phong Nguyen Nhu* and Thao Do Thi

HCMC University of Technology, Ho Chi Minh, Vietnam

***Correspondence:**

Phong Nguyen Nhu,
nnphong@hcmut.edu.vn

Received: 17 July 2023; **Accepted:** 19 July 2023; **Published:** 27 July 2023

Most of the 3PL companies that provide transportation services are handling thousands of orders per day. Vehicle routing problems (VRPs) help plan the distribution of goods with the optimum fleet of vehicles and delivery routes and play an important role in helping businesses reduce transportation costs while ensuring service level. VRPs are NP-hard combinatorial optimization problems. It is quite difficult to achieve an optimal solution for real-size problems with a mathematical modeling approach because of its NP-hard structure. Genetic algorithm (GA) plays a major role in searching for near-optimal solutions for NP-hard optimization problems. This article develops the GA model for VRPs. The result shows that the delivery cost is reduced by 17.88%, while the service level increase from 88.7 to 100%. It indicates that the model can be a good technique for VRPs.

Keywords: distribution planning, mathematical modeling, vehicle routing problems, genetic algorithm

1. Introduction

In today's fiercely competitive business environment, companies need to optimize their supply chains. Logistics is one of the key areas in supply chain management. Vehicle routing problem (VRP) is a core problem in logistics that refers to a class of combinatorial optimization in which customers are served by several vehicles. VRP is a problem in which there is a set of vehicles and a set of customers; each customer requires a certain number of goods.

This problem is a type of distribution problem that determines the transportation facilities and their delivery routes with specific objectives and constraints. The problem to be solved is a VRP with an assumption that the trucks start at the real warehouse and end at the virtual warehouse, the travel time between the points is fixed, the unloading time at the points is fixed, the speed of the trucks is the same and fixed, and the factors of failure, accident, and weather are not taken into account.

The objective of the problem is to minimize the transportation cost. The constraints are on the level of customer service and on the capacity of the transportation

facilities. The model of the problem is built based on the above assumptions, objectives, and constraints.

A genetic algorithm (GA) is built. Based on the model of the problem, the GA will find a good solution for the problem, and this solution will be compared with the solution of the currently used heuristic model to evaluate the effectiveness of the GA.

2. Literature review

2.1. Vehicle routing problems

Vehicle routing problem is about finding optimal routes for a fixed fleet of vehicles in order that they can meet the demands of a set of given customers by traveling through those paths (1). Vehicle routing is an important area in the field of supply chain management. Therefore, solving VRP efficiently plays a vital role in the effective implementation of supply chain practices for any organization (2).

Open VRP (OVRP) is a type of VRP, where routes finish after servicing the last client (3). The OVRP differs from the

classic VRP because either the vehicles are not required to return to the depot or they have to return by revisiting the customers assigned to them in the reverse order (4).

The capacitated VRP is a problem risen in the fields of transportation, distribution, and logistics in which a fleet of delivery vehicles must service known customer demands for a single commodity from a common depot at minimum cost (5). In the well-known VRP, a set of identical vehicles, based at a central depot, is to be optimally routed to supply customers with known demands subject to vehicle capacity constraints (6).

TABLE 1 | Loads, time limits, and fleet costs.

v	1	2	3	4	5	6
Q_v (kg)	300	300	300	400	400	500
f_v (M)	3	3	3	4	4	5
T_v (h)	8	8	8	8	8	8

TABLE 2 | Demand, unloading time at delivery points.

i	1	2	3	4	5	6	7	8
d_i (kg)	50	300	200	150	100	150	50	100
S_i (h)	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25

TABLE 3 | Travel time T_{ij} (h) between nodes.

i\j	0	1	2	3	4	5	6	7	8
0	-	0.068	0.137	0.133	0.143	0.345	0.203	0.261	0.202
1		-	0.08	0.065	0.076	0.345	0.228	0.265	0.236
2			-	0.062	0.049	0.301	0.221	0.231	0.238
3				-	0.019	0.362	0.27	0.289	0.283
4					-	0.349	0.265	0.279	0.28
5						-	0.165	0.084	0.189
6							-	0.088	0.03
7								-	0.117
8									-

TABLE 4 | The initial population $P^{(0)}$.

$P^{(0)}$	S1	S2	Routes	V	Z	F
C1	[1, 3, 2, 4, 7, 5, 6, 8]	[2, 3, 4, 5, 1, 6]	[1, 3]; [2]; [4, 7, 5]; [6, 8]	2, 3, 4, 5	14	1
C2	[1, 4, 3, 8, 2, 6, 7, 5]	[3, 1, 2, 4, 5, 6]	[1, 4]; [3, 8]; [2]; [6, 7, 5]	3, 1, 2, 4	13	2
C3	[6, 4, 1, 7, 8, 3, 5, 2]	[1, 3, 4, 5, 6, 2]	[6, 4]; [1, 7, 8]; [3, 5]; [2]	1, 3, 4, 5	14	1
C4	[8, 5, 1, 7, 6, 2, 3, 4]	[1, 2, 4, 6, 3, 5]	[8, 5, 1, 7]; [6]; [2]; [3, 4]	1, 2, 4, 6	15	0

The VRP is a challenging combinatorial optimization problem (7). Complexity in solving the problem increases exponentially as the problem size increases. GA is a widely applied search method to complex problems such as VRPs (2).

2.2. Genetic algorithm

Genetic algorithm, first introduced by Holland in 1975, is an artificial intelligence search method that uses the process of evolution and natural selection of individuals called chromosomes. It is an efficient tool for solving optimization problems. In order to apply GA to a problem, generally the solution space of the problem is represented by a population of chromosomes where each chromosome is a possible solution to the problem. Chromosomes are represented by strings. A method of coding is the selection of a string format for the chromosomes. A fitness value is associated with each chromosome. The fitness function is a measure of the extent to which the objective of the problem is achieved. The fitness function is derived from the objective function of the problem. A certain number of chromosomes are chosen to form the initial generation. The chromosomes of the next generation are generated by applying genetic operators, including selection, crossover, mutation, and replacement, to the chromosomes of the existing generation.

Starting from an initial population, the algorithm produces a new population of individuals, which are presumably more fit than their ancestors. The process is repeated until a pre-specified termination rule becomes true. At each generation, every new chromosome corresponds to a solution.

TABLE 5 | Values of F_i , P_i , and the cumulative probability function (CPF).

$P^{(0)}$	S1	S2	Z	F	P	CPF
C1	[1, 3, 2, 4, 7, 5, 6, 8]	[2, 3, 4, 5, 1, 6]	14	1	0.25	0.25
C2	[1, 4, 3, 8, 2, 6, 7, 5]	[3, 1, 2, 4, 5, 6]	13	2	0.5	0.75
C3	[6, 4, 1, 7, 8, 3, 5, 2]	[1, 3, 4, 5, 6, 2]	14	1	0.25	1
C4	[8, 5, 1, 7, 6, 2, 3, 4]	[1, 2, 4, 6, 3, 5]	15	0	0	1

TABLE 6 | The elite population $P_E^{(0)} = \{C2, C1, C3, C3\}$.

R_i	0.569	0.194	0.828	0.913
$P_E^{(0)}$	C2	C1	C3	C3

TABLE 7 | The crossover list $P_C = \{C2, C3\}$.

$P_E^{(0)}$	C2	C1	C3	C3
R_i	0.47	0.91	0.11	0.88
P_C	C2	-	C3	-

3. The vehicle routing problem

The problem to be solved is an OVRP, where vehicles after delivery to customers do not need to return to the starting point. The distribution system consists of a warehouse and a fleet of vehicles V_v , $v = 1 \div V$. Each vehicle V_v has max load Q_v , limited delivery time T_v , and operating cost f_v . The system distributes for N customers, and each customer is 1 point on the distribution network. The distribution network consists of $N + 1$ nodes, with node 0 being the warehouse, node i , $i = 1 \div N$ corresponding to the i th customer. Customers at node i have demand d_i , unloading time S_i , $i = 1 \div N$. On a distributed network, the travel time from node i to node j is T_{ij} , $i = 0 \div N-1$, $j = 1 \div N$.

TABLE 8 | The crossover population $P^C = \{C5, C6\}$.

P^C	S1	S2
C5	[6, 4, 3, 8, 1, 7, 5, 2]	[3, 1, 2, 4, 5, 6]
C6	[1, 4, 2, 8, 3, 5, 6, 7]	[1, 3, 4, 5, 2, 6]

TABLE 9 | The mutation list $P_m = \{C1\}$.

$P_E^{(0)}$	C2	C1	C3	C3
R_i	0.49	0.12	0.23	0.86
P_m	-	C1	-	-

TABLE 10 | The mutation population $P^M = \{C7\}$.

P^M	S1	S2
C7	[6, 8, 2, 4, 7, 5, 1, 3]	[4, 5, 2, 3, 1, 6]

TABLE 11 | The genetic population $P_G^{(0)}$.

$P_G^{(0)}$	S1	S2	Z
C5	[7, 4, 3, 8, 2, 6, 1, 5]	[3, 1, 6, 4, 5, 2]	13
C6	[2, 8, 7, 1, 4, 3, 6, 5]	[3, 4, 5, 1, 2, 6]	14
C7	[6, 8, 2, 4, 7, 5, 1, 3]	[4, 5, 2, 3, 1, 6]	14

TABLE 12 | The next population.

$P^{(1)}$	S1	S2	Z	F
C2	[1, 4, 3, 8, 2, 6, 7, 5]	[3, 1, 2, 4, 5, 6]	13	2
C5	[7, 4, 3, 8, 2, 6, 1, 5]	[3, 1, 6, 4, 5, 2]	13	2
C6	[2, 8, 7, 1, 4, 3, 6, 5]	[3, 4, 5, 1, 2, 6]	14	1
C7	[6, 8, 2, 4, 7, 5, 1, 3]	[4, 5, 2, 3, 1, 6]	14	1

The objective of the problem is to minimize the total delivery cost. The constraints of the problem include the following: Each route starts at the warehouse and ends at a customer point. The number of vehicles starting at the real warehouse and ending at the virtual warehouse is equal and does not exceed the limited number of vehicles. The number of cars arriving and leaving at each node is equal. Each customer is only served once by one vehicle. The total delivery time does not exceed the time limit. The total demand of the customers in each route does not exceed the vehicle tonnage.

The model is set up with the decision variable y_{ij}^v , $i = 0 \div N-1$, $j = 1 \div N$, $v = 1 \div V$. If vehicle v moves from node i to node j , then y_{ij}^v is 1; otherwise, y_{ij}^v is 0. The model is as follows:

$$\text{Min } Z, Z = \sum_{v=1}^V \sum_{j=1}^N f_v y_{0j}^v$$

$$\text{St. : } \sum_{v=1}^V \sum_{j=1}^N y_{0j}^v = \sum_{v=1}^V \sum_{i=1}^N y_{i0}^v \leq V$$

$$\sum_{v=1}^V \sum_{i=0}^N y_{ij}^v = 1, j \in N \setminus \{0\}, i \neq j$$

$$\sum_{i=0}^N y_{ie}^v = \sum_{j=0}^N y_{ej}^v, v \in V, e \in N \setminus \{0\}, e \neq i, e \neq j$$

$$\sum_{i=0}^N \sum_{j=1}^N (d_j y_{ij}^v) \leq Q_v, v \in V, i \neq j$$

$$\sum_{i=0}^N \sum_{j=0}^N [(t_{ij} + s_j) y_{ij}^v] \leq T, v \in V, i \neq j$$

$$y_{ij}^v \in \{0, 1\}, v \in V, j \in N, i \in N, i \neq j$$

4. The GA model for the pilot VRP

The pilot problem is a VRP problem with six trucks and eight customers ($V = 6$, $N = 8$). The parameters of fleet, customers, and travel time between nodes are as in **Tables 1–3**. The pilot problem model is set up according to the above parameters. The GA model is used to solve the problem by the following procedure:

Step 1: Initialize the GA model.

Step 2: Generate the initial population $P^{(0)}$. Set $k = 0$.

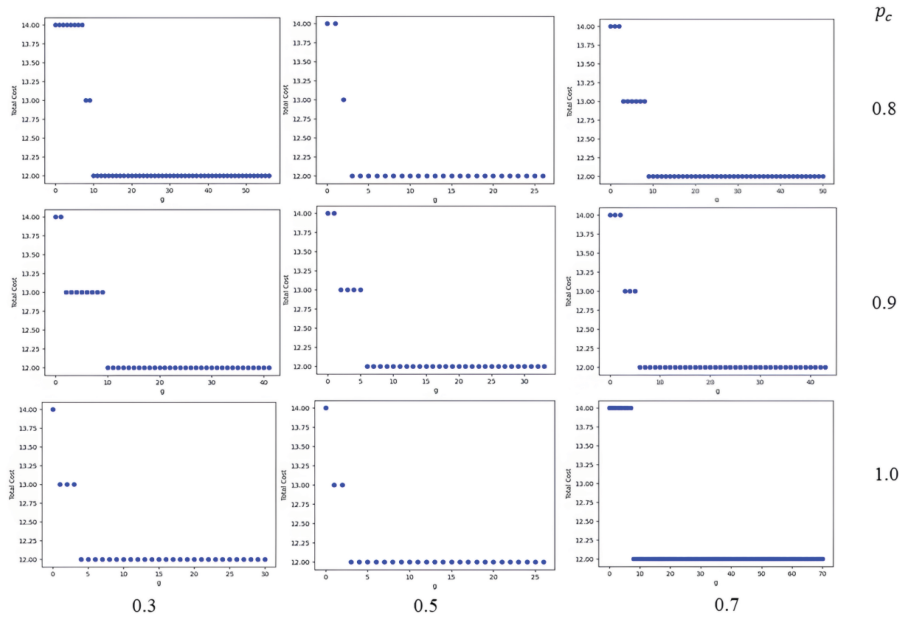


FIGURE 1 | The objective values Z in nine combinations of P_c , P_m .

- Step 3: Generate elite population $P_E^{(k)}$.
- Step 4: Generate the genetic population $P_G^{(k)}$.
- Step 5: Generate the next population $P^{(k+1)}$. Set $k = k + 1$.
- Step 6: Check the termination rule. If no, return to Step 3. If yes, finish the loop.

Step 1: Initialize the GA model

This step sets up factors of GA models, including the method of coding, the population size, the fitness function, the parameters of GA operators, and the termination rule.

The method of coding: Each chromosome C is a string of two sub-chromosomes, S1 is corresponding to the customer sequence, and S2 is corresponding to the vehicle sequence:

$$C = [S1, S2]$$

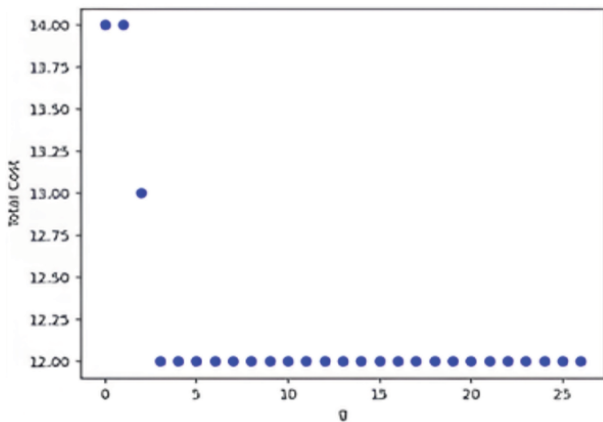


FIGURE 2 | The objective value Z in combination of 0.8, 0.5.

Customers are numbered from 1 to 8. Each sub-chromosome S1 is a string of eight genes, G_{ci} , $i = 1 \div 8$, corresponding to eight customers. The sequence of genes in a sub-chromosome represents the sequence of customers being served: $S1 = (G_{c1}, G_{c2}, G_{c3}, G_{c4}, G_{c5}, G_{c6}, G_{c7}, G_{c8})$. Vehicles are numbered from 1 to 6. Each sub-chromosome S2 is a string of six genes, G_{vj} , $j = 1 \div 6$, corresponding to six vehicles. The sequence of genes in a sub-chromosome represents the sequence of vehicles being used:

$$S2 = (G_{v1}, G_{v2}, G_{v3}, G_{v4}, G_{v5}, G_{v6})$$

The population size P is chosen to be 4:

$$P = 4.$$

The fitness function F is defined as $F_i = Z_{max} - Z_i$, where F_i and Z_i are the fitness and objective values of chromosome i and Z_{max} is the maximum objective value in the population.

The GA operators include selection, crossover, mutation, and replacement operators. The selection method is based on selection probabilities, determined by fitness values. The crossover method is OX (Order Crossover), the mutation method is SWAP, and the replacement method is acceptance threshold. The crossover probability P_c , the mutation probability P_m , and threshold K are chosen as follows:

$$P_c = 0.8; P_m = 0.2; K = 2.$$

The termination rule: The best objective value of the population Z_{min} does not decrease after 15 consecutive iterations.

Step 2: Generate the initial population $P^{(0)}$, set $k = 0$.

The initial population consists of four chromosomes. Each initial individual is created including customer sequence and vehicle sequence, by randomizing the customer delivery points and available vehicles. The initial population $P^{(0)}$ is shown in **Table 4**: $P^{(0)} = \{C1, C2, C3, C4\}$.

From customer sequence S_1 , vehicle sequence S_2 , based on the constraints, the vehicles used, and their corresponding routes are determined, as shown in **Table 4**. From that, the objective values Z and fitness values F of each chromosome can be determined. In the initial population $P^{(0)}$, the best chromosomes are C2, with the best objective value Z of 13 (M).

Step 3: Generate elite population $P_E^{(k)}$.

This step uses the selection operator to generate elite population $P_E^{(k)}$ from $P^{(k)}$. Each chromosome in the current population has a corresponding fitness value F_i and is selected for inclusion in the elite population $P_E^{(k)}$ with selection probability P_i determined as follows:

$$P_i = F_i / \sum_{i=1}^4 (F_i).$$

With population $P^{(0)}$, the values of F_i , P_i , and the cumulative probability function (CPF) are calculated as shown in **Table 5**. Random numbers R_i are generated 4 times. Based on CPF, the chromosomes selected into the population $P_E^{(0)}$ are as **Table 6**.

Step 4: Generate the genetic population $P_G^{(k)}$.

This step uses the crossover and mutation operators to generate genetic population $P_G^{(k)}$ from the elite population $P_E^{(k)}$. The genetic population $P_G^{(k)}$ includes the new chromosome generated from the crossover and mutation operators.

The chromosomes of $P_E^{(0)}$ are selected to be included in the crossover list P_c with the crossover probability of 0.8. After generating four random numbers R_i , the set P_c is determined as shown in **Table 7**. Each pair of chromosomes in P_c is selected to cross over by the OX method, on both sub-chromosomes, resulting in two new chromosomes in population P^C . C2 and C3 are crossed over to each other and generate two children, C5 and C6. The population P^C is shown in **Table 8**.

The chromosomes of $P_E^{(0)}$ are also selected to be included in the mutation list P_m with the mutation probability of 0.2. After generating four random numbers R_i , the set P_m is determined as shown in **Table 9**. Each chromosome in P_m is selected to mutate by the SWAP method, on both sub-chromosomes, resulting in one new chromosome in population P^M . C1 is mutated and generates C7. The population P^M is shown in **Table 10**.

After crossover and mutation, three new chromosomes are created in the population $P_G^{(0)}$. Chromosomes in $P_G^{(0)}$ with their genes and objective values are shown in **Table 11**.

$$P_G^{(0)} = \{C5, C6, C7\}$$

Step 5: Generate the next population $P^{(k+1)}$

This step uses the replacement operator to generate the next population $P^{(k+1)}$ from the populations $P^{(k)}$ & $P_G^{(k)}$. The chromosomes from $P_G^{(k)}$ will be added to the current population $P^{(k)}$ to make the next population $P^{(k+1)}$ if their objective values are better than the acceptable threshold, defined by the objective value of the threshold chromosome. The threshold chromosome is a chromosome in $P^{(k)}$ with position n defined by population size P and threshold parameter K : $n = P/k$. With a population size of 4, and the replacement parameter K of 2, the acceptable threshold is selected as the value of the second chromosome of $P^{(k)}$ in the top-down ranking.

On the other side, in order to keep the next population size constant, some chromosomes in $P^{(k)}$ with the lowest fitness values are removed. In $P^{(0)}$, C1 is the second chromosome in the top-down list, and the acceptable threshold is 14. Looking at the objective values of the chromosomes in population $P_G^{(0)}$, C5, C6, and C7 move into $P^{(1)}$ and C1, C3, and C4 have to move out to keep $P^{(1)}$ population size constant. **Table 12** shows the chromosomes in the next population $P^{(1)}$. In $P^{(1)}$, the best chromosomes are C2 and C3, with the best objective value of 13.

$$P^{(1)} = \{C2, C5, C6, C7\}$$

TABLE 13 | Loads, time limits, and fleet costs.

V	1÷6	7÷9	10	11÷12
Q_v (kg)	1000	1400	1900	2000
f_v (M)	1.150	1.600	1.800	2.000
T_v (h)	8	8	8	8

TABLE 14 | Demand, unloading time at delivery points.

i	1	2	3	4	...	97
d_i (kg)	14.5	17	9	14.5	...	1000
S_i (h)	0.25	0.25	0.25	0.25	...	0.25

TABLE 15 | Travel time T_{ij} (h) between nodes.

	0	1	2	3	4	...	97
0	-	0.313	0.214	0.124	0.521	...	0.365
1		-	0.651	0.145	0.689	...	0.287
2			-	0.09	0.871	...	0.697
3				-	0.416	...	1.010
4					-	...	0.723
...						-	...
97							-

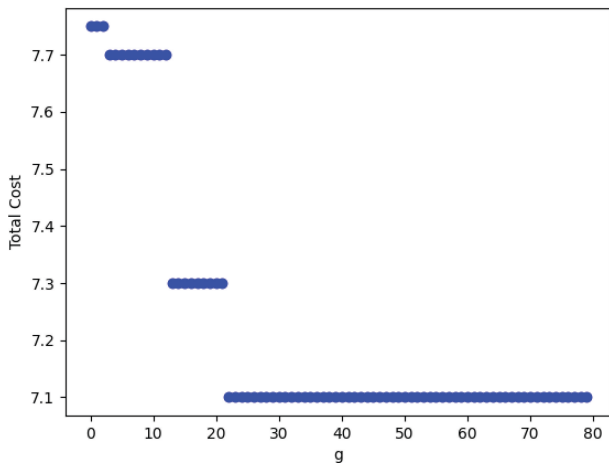


FIGURE 3 | Total cost Z.

Step 6. Check the termination rule.

After iteration 1, the best objective value is 13, appearing only twice. The termination rule is not satisfied, so iteration 2 is executed.

5. The GA model with experiments for the pilot VRP

Experiments are performed on the above GA model with a population size P of 20, and the two input factors are the probability of crossover P_c and the probability of mutation P_m . The probability P_c has three levels 0.8, 0.9, and 1. The probability P_m has three levels 0.3, 0.5,

TABLE 16 | The vehicles used and the corresponding route by genetic algorithm.

v	Route
2	[0, 12, 77, 82, 64, 50, 60, 42, 69, 93, 28, 87, 30, 11, 21, 32, 96, 25, 56, 91, 53, 5, 23]
5	[0, 6, 75, 66, 57, 9, 72, 73, 22, 16, 41, 89, 81, 13, 19, 36, 20, 79, 92, 74, 24, 43, 2, 90]
7	[0, 97, 26, 44, 95, 48, 15, 4]
8	[0, 71, 94, 45, 18, 51, 14, 3, 59, 84, 70, 65, 88, 52, 8, 38, 61, 63, 86, 58, 10, 54, 62]
9	[0, 7, 31, 47, 1, 29, 80, 40, 76, 67, 39, 49, 37, 35, 68, 46, 83, 17, 78, 34, 55, 27, 33, 85]

TABLE 17 | Total cost between the current method and the genetic algorithm (GA) model.

Model	Total cost (million)	Number of late delivery nodes	Service level (%)
Current model	7.55	11	88.7
GA	7.1	0	100

and 0.7. The run charts of the objective values Z in nine combinations of P_c, P_m are shown in Figure 1. The nine combinations (P_c, P_m) give the same objective value, but the combination (0.8, 0.5) has the smallest number of iterations. This combination is used to solve real-world problems.

The nine combinations (P_c, P_m) give the same objective value, but the combination (0.8, 0.5) has the smallest number of iterations, as shown in Figure 2. This combination is used to solve real problems.

6. The GA model for the real VRP

The real problem is a VRP problem with 12 trucks and 97 customers ($V = 12, N = 97$). The parameters of fleet, customers, and travel time between nodes are shown in Tables 13–15.

The real problem model is set up according to the above parameters. The GA model is used to solve the problem with the parameter:

$$P = 20, P_C = 0.8, P_M = 0.5$$

The run chart of the total cost Z is as Figure 3. The result has a total cost of 7.1 million with the vehicles used and the corresponding route as shown in Table 16.

The comparison between the current method and the GA model is shown in Table 17. Compared with the current model, the GA model has a cost reduction of 7.55 million to 7.1 million (17.88%), the number of late delivery points reduced from 11 to 0, and the service level increased from 88.7 to 100%.

7. Conclusion

The GA model has been used to solve the VRP in a distribution system with one warehouse, 12 trucks, and 97 customers. The results show that the GA model gives lower cost and higher service level than the heuristic method, being used. However, the parameters of the model are only selected empirically, so the results are not very good. Future research is to use experimental design DOE to determine the model parameters to get suboptimal results.

Author contributions

PN is the thesis advisor of TT. PN and TT contributed to the thesis conception and design. PN has developed the research models for the thesis. TT has collected and analyzed data and

written program to run the algorithms, based on the models. PN has composed the article, based on the thesis. Both authors read and approved the final manuscript.

Acknowledgments

We extend our heartfelt appreciation to everyone who has contributed to the completion of this research article, especially our families, HCMC University of Technology, and the scientific community for their invaluable support.

Conflict of interest

The research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Foroutan RA, Rezaeian J, Mahdavi I. Green vehicle routing and scheduling problem with heterogeneous fleet including reverse logistics in the form of collecting returned goods. *Appl Soft Comput.* (2020) 94:106462.
2. Kumar V, Panneerselvam R. A study of crossover operators for genetic algorithms to solve VRP and its variants and new sinusoidal motion crossover operator. *Int J Comput Intell Res.* (2017) 13:1717–33.
3. Ruiz E, Soto-Mendoza V, Ruiz Barbosa AE, Reyes R. Solving the open vehicle routing problem with capacity and distance constraints with a biased random key genetic algorithm. *Comput Ind Eng.* (2019) 133:207–19.
4. Sariklis D, Powell S. A heuristic method for the open vehicle routing problem. *J Oper Res Soc.* (2000) 51:564–73.
5. Liu S, Huang W, Ma H. An effective genetic algorithm for the fleet size and mix vehicle routing problems. *Transp Res E.* (2008) 45:434–45.
6. Baldacci R, Battarra M, Vigo D. Routing a heterogeneous fleet of vehicles. In: Golden B, Raghavan S, Wasil E editors. *The vehicle routing problem: latest advances and new challenges.* Boston, MA: Springer (2008).
7. Kan H, Lee AHL. An enhanced approach for the multiple vehicle routing problem with heterogeneous vehicles and a soft time window. *Symmetry.* (2018) 10:650.