Advance Journal of Food Science and Technology 10(1): 31-36, 2016
DOI: 10.19026/ajfst.10.1748
ISSN: 2042-4868; e-ISSN: 2042-4876

Submitted: April 19, 2015 Accepted: May 10, 2015 Published: January 05, 2016

Research Article

Multiple-depot Food Transport Vehicle Routing Genetic Algorithm Based on Two-stage Fuzzy Clustering

[1]Qiang Song, [1]He Feng and [2]Lingxia Liu
[1]Computer School, Anyang Institute of Technology,
[2]School of Software Engineering, Anyang Normal University, Anyang City, 455000, China

Abstract: Aiming to large-scale Multiple-Depot Food transport Vehicle Routing Problem (MDFVRP), this study proposed an improved genetic algorithm solution frame based on the two-stage fuzzy clustering. In the static upper stage, the k-means technology is used to divide the MDFVRP into several one-to-many sub-problems. From the perspective of improving the customer satisfaction and integrating logistics resource, the lower fuzzy clustering stage adopts fuzzy clustering algorithm to form the dynamic customer base based on customer's order distribution according to customers request attributes. Furthermore, the Genetic Algorithm (GA) of VRP is designed through the improvement of the selecting operator and the crossover operator. The stochastic simulation experiments show the proposed algorithm and solution strategy are efficient.

Keywords: Food transport vehicle routing, improved genetic algorithm, two-stage fuzzy clustering

INTRODUCTION

Multiple-Depot Food transport Vehicle Routing Problem (MDFVRP) refers to the food transport vehicles can start from more than one depot, traverse the appropriate routes under the certain constraints (customer demand, the number of food transport vehicles, food transport vehicle capacity, etc.), to achieve the goal of optimization (the minimum cost, the shortest distance, the largest customer satisfaction, etc.). The problem is the extension of a classical food transport Vehicle Routing Problem (VRP), is a more complex NP problem and with the increase of the depot number, the solution space of problem is seriously swelling, the efficiency of the existing algorithm will drop rapidly, greatly deducing the solution space and avoiding repetition search to become the key to solve the MDFVRP. But MDFVRP is more in line with the complex application background in the reality, the enterprise business is expanding continuously, customers have a wide distribution range, more and more enterprises choose the area where customers congregate to establish multiple regional distribution centers, which has formed a complex large-scale distribution network with multiple depot service for more customers.

At present, MDFVRP has become the difficult problems that are commonly concerned by the academia and the business circles, how to realize the rapid, efficient and intelligent solution of the problem, becomes one of the hard problem in the logistics field to be solved.

At present, the research on MDFVRP mainly focuses on the improved intelligent optimization algorithms. Ho *et al*. (2008) proposed two kinds of hybrid genetic algorithm to solve multiple-depot food transport vehicle routing problem. Mirabi *et al*. (2010) used random hybrid heuristic algorithm to solve the MDFVRP. Yvcenur and Demirel (2011) adopted the genetic algorithm based on clustering technology to solve MDFVRP. Tong *et al*. (2004) presented the MDFVRP model and used genetic algorithm to solve the problem. In order to optimize the goal, Liying and Haipeng (2009) did not consider the customer orders attributes and customer satisfaction in the optimization process. Along with the increase of the customer's number and the problem size, the solving efficiency of all kinds of existing heuristic algorithm must drop rapidly (Huang *et al*., 2009). In view of the actual large-scale MDFVRP, greatly reducing the problem solving space becomes one of the effective ways to solve the problem.

From the perspective of integration of logistics resources and improving customer satisfaction, this study puts forward the improved genetic algorithm based on two-stage fuzzy clustering, the upper static field division will transfer the many-to-many problems into multiple one-to-many sub-problems, the lower level fuzzy clustering adopts fuzzy clustering technology according to the customer requirement attributes to form a distribution client base, on the basis the improved genetic algorithm is used for solving food transport vehicle routing problem. The example

Corresponding Author: Qiang Song, Computer School, Anyang Institute of Technology, Anyang City, 455000, China

simulation proves that the solving strategy is effective for the large-scale MDFVRP.

MATERIALS AND METHODS

The framework of two-stage fuzzy clustering algorithm: Two-stage fuzzy clustering algorithm framework includes static field classification and fuzzy clustering two layers. At first k-means technology is used to form static client base to customer coordinate clustering and according to the distance the multiple-depot multiple-customer problem is transferred into multiple one-to-many sub-problems; then according to the multiple attribute characteristic of customer order demand to determine fuzzy clustering decision variables, the fuzzy clustering algorithm is used to cluster customers. The clustering framework can effectively solve the problems that in many-to-many large-scale logistics distribution network the batch is few, frequency is high and the part-load traffic cost is high, at the same time ensuring customer satisfaction, the logistics resources should be integrated to reduce distribution costs.

The fuzzy clustering algorithm: Hu and Sheu (2003) emphasized the importance of customer clustering in the increasingly complex logistics environment and according to customer's requirements variable the binary conversion is used for customers group clustering, on this basis, aiming to MDFVRP this study designs the fuzzy clustering algorithm including four stages:

- The determination of fuzzy clustering decision variables
- Dealing with the clustering decision variables
- Fuzzy similarity calculation
- Customer clustering procedure

The fuzzy clustering decision variables: In the optimization of logistics distribution network, the companies tend to focus on the indicators such as the loading rate of food transport vehicle, food transport vehicle travel cost and customer satisfaction and so on. The customer often concerns about the indicators such as delivery punctuality, service attitude, safety of goods delivery in the distribution process, these indicators directly decides the discretion of the customer satisfaction. Based on the principle of improving the loading rate of food transport vehicle distribution, reducing distribution cost and improving customer satisfaction, the five clustering decision variables are designed:

$$y_i^1, y_i^2, y_i^3, y_i^4, y_i^5$$

Hereinto, y_i^1 is the order quantity demanded for customer i, through the order integration to implement part-load traffic to full carload transport. y_i^2 represents the delivery deadline demanded by customers. y_i^3 represents the expiration date of order demand products of customer i. y_i^4 represents the requirement of customer i for the quality of distribution services, which is mainly reflected at both sides of the response time to customer's requirement and distribution service attitude. y_i^5 represents the external characteristics of the demanded product by customer i, if the product's exterior characteristics are more similar, the efficiency of unified batch loading and unloading process is higher (Changshi and Mingyong, 2010). y_i^1, y_i^2 and y_i^3 are quantitative decision variables, y_i^4 y_i^5 are qualitative decision variables.

The process of clustering decision variables: Both of customer demand decision variable y_i^1 and the food transport vehicle load capacity Q are one of the key conditions to judge whether the customer clustering stops, it doesn't need to be deal with and only the part-load traffic part is concerned, if the quantity demanded of customers exceeds the maximum loading capacity, the whole food transport vehicle part will adopt direct distribution.

The process of qualitative decision variable: Qualitative decision variables processing of y_i^4 and y_i^5 uses of five language scales to evaluate customer requirements attributes, the five language scales are "very high", "high", "middle", "low", "very low", the corresponding generalized trapezoidal fuzzy numbers are shown in Table 1, hereinto, ω_i^k represents the credibility of the data provided by customers and the value is between (0, 1) and thus the qualitative variable is transferred into a generalized trapezoidal fuzzy number, it is expressed in formula (1):

$$y_i^k = (\sigma_{i,1}^k, \sigma_{i,2}^k, \sigma_{i,3}^k, \sigma_{i,4}^k; \omega_i^k), \quad k = 4,5 \tag{1}$$

The process of quantitative decision variables: Because the dimension of y_i^2 and y_i^3 are different, the extremum method is adopted to standardize the data processing, which is shown in formula (2) and the standardized data is transferred into a generalized trapezoidal fuzzy numbers, such as formula (3):

$$\overline{y}_i^k = \frac{y_i^k - \min(y_i^k)}{\max(y_i^k) - \min(y_i^k)}, \quad k=2,3 \tag{2}$$

$$\tilde{y}_i^k = (\overline{y}_i^k, \overline{y}_i^k, \overline{y}_i^k, \overline{y}_i^k; \omega_i^k), \quad k=2,3 \tag{3}$$

The calculation of fuzzy similarity: Customer i and customer j adopt the method proposed in literature about the similarity calculation of decision variable k, the similarity S ($\tilde{y}_i^k$, $\tilde{y}_j^k$) of $\tilde{y}_i^k$ and $\tilde{y}_j^k$ can be represented by formula (4):

Table 1: The generalized trapezoidal fuzzy numbers corresponding to five levels

	The generalized trapezoidal fuzzy numbers				
Language scale	$\sigma_{i,1}^{k}$	$\sigma_{i,2}^{k}$	$\sigma_{i,3}^{k}$	$\sigma_{i,4}^{k}$	ω_{i}^{k}
Very high	0.8	0.9	0.9	1.0	(0, 1)
High	0.6	0.7	0.8	0.9	(0, 1)
Middle	0.5	0.6	0.6	0.7	(0, 1)
Low	0.2	0.3	0.4	0.5	(0, 1)
Very low	0.0	0.1	0.2	0.3	(0, 1)

$$S(\tilde{y}_i^k, \tilde{y}_j^k) = [1 - |\frac{\sum_{n=1}^{4}(\sigma_{i,n}^k - \sigma_{j,n}^k)}{4}|] \times \frac{\min(P(i^k), P(j^k)) + \min(\omega_i^k + \omega_i^k)}{\max(P(i^k), P(j^k)) + \max(\omega_i^k, \omega_j^k)} \tag{4}$$

In the expression, $P(i^k)$ and $P(j^k)$ are respectively the perimeters of the generalized trapezoidal fuzzy numbers $\tilde{y}_i^k$ and $\tilde{y}_j^k$. The following expressions can be concluded:

$$P(\mathrm{i}^k) = \sqrt{(\sigma_{i,2}^k - \sigma_{i,1}^k)^2 + (\omega_i^k)^2} + \sqrt{(\sigma_{i,4}^k + \sigma_{i,3}^k)^2 + (\omega_j^k)^2} + (\sigma_{i,3}^k - \sigma_{i,2}^k) + (\sigma_{i,4}^k - \sigma_{i,1}^k) \tag{5}$$

$$P(\mathrm{j}^k) = \sqrt{(\sigma_{j,2}^k - \sigma_{j,1}^k)^2 + (\omega_j^k)^2} + \sqrt{(\sigma_{j,4}^k + \sigma_{j,3}^k)^2 + (\omega_j^k)^2} + (\sigma_{j,3}^k - \sigma_{j,2}^k) + (\sigma_{j,4}^k - \sigma_{j,1}^k) \tag{6}$$

$S(\tilde{y}_i^k, \tilde{y}_j^k,) \in [0, 1]$, if its value is bigger, which shows the similarity of the decision variable k is higher between customer i and j, after the similarity is obtained among each decision variable about the customers, it is necessary according to the weight φ_k of each decision variable to determine the similarity z_{ij} among the customers, which are shown in formulas (7) and (8). According to the similarity between the customers a $n \times n$ fuzzy similar symmetric matrix can be constructed:

$$Z = (z_{ij})_{n \times n} \quad z_{ij} = \sum_{k=2}^{5} \varphi_k S(\tilde{y}_i^k, \tilde{y}_j^k) \tag{7}$$

$$\sum_{k=2}^{5} \varphi_k = 1 \quad \varphi_k \geq 0 \tag{8}$$

Customer clustering algorithm: Customer clustering algorithm is based on fuzzy similarity matrix, divides the customers with high similarity into the same distribution base; the termination conditions of algorithm are all customers are processed, the steps are as follows:

Step 1: Take any one of the upper clustering customer base, input fuzzy similar matrix Z, initialize the customer clustering number t = 1.

Step 2: Initialize the computing cycle and suppose the initial cycle number h = 1.

Step 3: Take the target customer j to start the cycle, remove the related line z_j with target customer j in fuzzy similar matrix Z.

Step 4: Suppose C is the customer set distributed into the same base with customer j in column z_j and $C = \{j\}$:

- Find the maxim element z_{ij} in column z_j and cluster according to the steps below.
- If $C = \{i, j\}$ and at the same time it meets the conditions $z_{ij} \geq \lambda$ and $\sum_{c \in C} y_c^1 \leq Q \times V_m$, customer i and customer j are distributed into the same customer base, the customer set $C = \{i, j\}$ and the line z_i^T, corresponding to customer i is deleted.
- $h = h + 1$, return back to step 1 to check the other elements in z_j until there is no elements to meet the upper conditions.
- Delete the column z_j from the matrix Z.

Step 5: If all clients are assigned, namely in fuzzy matrix there is no columns, the clustering program will stop, otherwise t = t+1 and return back to step 2 to do the next cycle. Hereinto, V_m is the food transport vehicle number in depot m. λ is the standard to judge the relative similarity among customers, it is also an important factor (Wei and Chen, 2009) to decide the customer clustering number t and cycle times h, if λ value is too small, the clustering is speed up due to easing the similarity conditions, a few of customer clustering number will generate, which will cause unrealistic clustering results, therefore, in generally its value is greater than 0.5.

Improvement of genetic algorithm: The two-stage fuzzy clustering forms the distribution customer base, the follows uses the proposed improved genetic algorithm to do food transport vehicle routing optimization for the distribution customer base. This study improves the selection operator and crossover operator of traditional genetic algorithm and the global optimization ability of this algorithm is improved. Improved selection operator will combine the best strategy and the roulette selection model, at the same time keeping the optimal solution ensures the fair opportunity that the rest solutions enter into the next generation, namely the bigger is the fitness value in the rest solutions, the bigger is the opportunity entering into the next generation.

The improved crossover operator randomly selects gene sections for cross through the front and shifts the parent residual gene in sequence to form new progeny, to effectively avoid the local optimum.

To generate the initial population: Chromosome encoding adopts the ordinal number coding method based on customer, chromosome G is expressed by vector (l_1, l_2, …, l_n), gene l_i is a natural number without repetition between (1, n), n is the customer number, the random sequence of chromosome structure generates the initial population with size of N_1.

Fitness function: The individuals with high fitness in genetic algorithm have greater probability to enter into the next generation, due to the objective function f (x) is to solve the minimum path cost, the fitness function is expressed by the reciprocal of the objective function, that is:

$$f = {}^{1}\!/\!{}_{f(x)}$$

The operator selection: The selection strategy adopts the combination of retaining the best and roulette wheel selection, N_1 individuals in each generation sort according to the fitness value f, the individual ranked at No. 1 directly copies and goes into the next generation, lists side by side at the first place, the (N_1 - 1) other individuals of the next generation will generate by using the roulette method, according to the fitness of N_1 individuals in the last generation. While keeping the best individual at the same time, the individuals with bigger fitness have greater chance to enter into the next generation.

Crossover operator: Randomly generate two intersection points between two parents' generation (parent 1 and parent 2) respectively, the gene between two junctions is taken as gene section for cross, which are remarked as $(C - GE)_1$ and $(C - GE)_2$, progeny 1 takes $(C - GE)_2$ as the start gene, from left to right moves the non-overlap genes with $(C - GE)_2$ in parent 1 to the progeny1, similarly the progeny 2 can generate. Adopting the method of front cross gene sections, even if both of the parents are same, also can produce a new progeny to do iterative optimization and overcome the shortcomings of local optimal and premature convergence.

The mutation operator: Using translocation mutation operation, with a certain mutation probability randomly select the chromosomes with mutation, on the chromosomes randomly select two genes and swap the genes.

RESULTS AND DISCUSSION

MATLAB 7.8 programming is used to realize random examples simulation, the program running environment is: 2.26 GHz Intel CPU, 512 MB memory. Within the Cartesian plane of 200×200, 62 customer points and three depots are randomly generated as shown in Fig. 1. The attribute information of 62 customer order requirement is randomly generated, the quantity demand of each customer obeys the uniform distribution U (1, 10), the delivery deadline is randomly selected in (0.5, 1, 1.5, 2, 3, 3.5, 2.5 and 4, respectively), the product warranty is generated from (30, 45, 60, 120 and 180, respectively), the customer demand for service quality and product exterior characteristic are randomly generated from (very high, high, low, very low), suppose that the credibility of the data provided by the customer is 1, when the fuzzy similarity calculation is executed, the delivery deadline, product warranty, customer demand for service quality and product exterior characteristic, the corresponding weights of these four decision variables are set to 0.3, 0.2, 0.3 and 0.2, respectively the similarity criterion λ value is 0.84, the division result of customer bases in the two-stage fuzzy clustering algorithm is shown in Table 2.

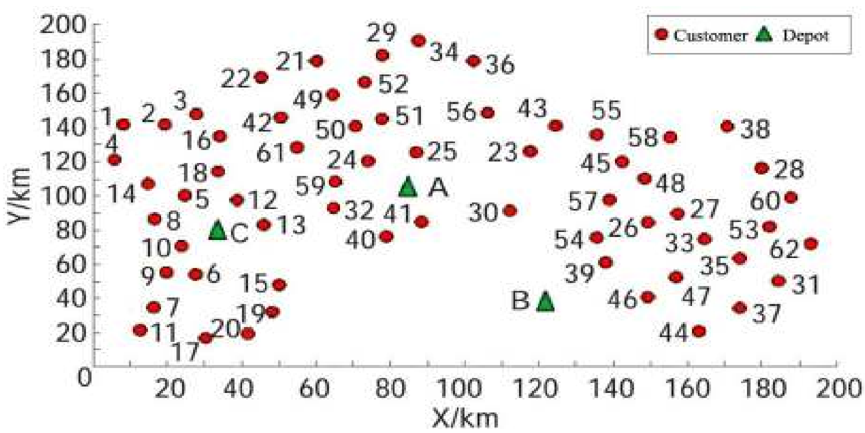

Fig. 1: The distribution diagram between depots and customers

Table 2: The divisions result of customer bases in the two-stage fuzzy clustering algorithm

Upper level static clustering	Low level fuzzy clustering	Distributed customer base	Food transport vehicle distance optimization
F1	F11	1, 2, 3, 4, 14, 16, 18	C, A, B
	F12	7, 11, 15, 17, 19, 20	
	F13	5, 6, 8, 9, 10, 12, 13	
F2	F21	21, 22, 29, 34, 36, 42, 49 50, 51, 2, 56, 61	A, B, C
	F22	24, 25, 32, 40, 41, 59	
F3	F31	23, 30, 43, 45, 54, 55, 57	B, A, C
	F32	26, 27, 33, 39, 46, 47	
	F33	28, 31, 35, 37, 38, 44, 48 53, 58, 60, 62	

Table 3: The improved genetic algorithm optimization results in fuzzy clustering strategy

Depots	Customer base	Optimal route	Food transport vehicle cubed out/ (%)	Distribution distance cost/km	Depot distance cost/km	Total distance cost/km
A	F21	A-61-42-22-21-49-50-A	95.0	178	489	1742
		A-51-52-29-34-36-56-A	91.0	191		
	F22	A-25-24-59-32-40-41-A	92.5	120		
B	F31	B-54-57-45-55-43-23-30-B	97.5	224	806	
	F32	B-39-26-27-33-47-46-B	90.0	144		
	F33	B-48-58-38-28-60-53-B	94.0	256		
		B-44-37-31-62-35-B	89.5	182		
C	F11	C-14-41-2-3-16-18-C	93.0	162	447	
	F12	C-7-11-17-20-19-15-C	90.0	160		
	F13	C-13-12-5-8-10-9-6-C	92.0	125		

Table 4: The results comparison of different strategy

	The improved genetic algorithm based on two-stage fuzzy clustering		The genetic algorithm based on distance clustering		
Depot	The total distribution distance cost/KM	The food transport vehicle cubed out rate section (%)	The total distribution distance cost/KM	The food transport vehicle cubed out rate section	The cost improvement rate / (%)
A	489	(91-95)	532	(61-73)	8.1
B	806	(89.5-97.5)	901	(70-82)	10.5
C	447	(90-93)	495	(60-72)	9.7

According to the division results of customer base in Table 2, the improved genetic algorithm is used for food transport vehicle routing optimization. The model of distribution food transport vehicle is single, the maximum loading capacity is 20t and the improved genetic algorithm parameters are set to: the population size is 50, the iteration time is 100, the crossover probability is 0.9, the mutation probability is 0.05, the food transport vehicle routing optimization results of each depot are shown in Table 3.

The results of the improved genetic algorithm under two-stage fuzzy clustering and the genetic algorithm based on distance clustering are compared and analyzed, the algorithm based on distance clustering is only according to the distance between customer to each depot, according to the principle of minimum distance the customer is assigned to each depot, only the shortest distance that each depot services customers is taken as the optimization goal, regardless of the food transport vehicle and container etc., logistics resources utilization and customers satisfaction level. The results are shown in Table 4, which shows that the solution results in this proposed strategy reduces about 10% of the distribution cost than the genetic algorithm based on distance clustering, the food transport vehicles cubed outs have obviously improved, under the two-stage fuzzy clustering the food transport vehicle cubed outs are all above 89.5%.

CONCLUSION

This study puts forward the improved genetic algorithm based on fuzzy clustering to solve the large-scale multiple-depot food transport vehicle routing problem, designs the two-stage fuzzy clustering algorithm frame, by the static upper clustering, the many-to-many problem could be divided into multiple one-to-many sub-problems, the fuzzy clustering in the lower level starts from the perspective of improving customer satisfaction and integration of logistics resources, according to the customer demand the property variables via fuzzy clustering algorithm form distribution customer base and use the improved genetic algorithm to optimize the food transport vehicle routing of each distribution customer base. The improved genetic algorithm has stronger optimal ability to prevent the solution trapped into a local optimum; the example simulation in this study demonstrates that the proposed strategy for solving large-scale multiple-depot food transport vehicle routing problem has better optimization effect.

ACKNOWLEDGMENT

This study is supported by the funds of the science and technology research key project of Henan province (2014), China. The project name is "The logistics distribution food transport vehicle scheduling model and its optimization based on the dynamic demand" and the project number is 142102210231.

REFERENCES

Changshi, L. and L. Mingyong, 2010. The stochastic food transport vehicle routing problem based on fuzzy clustering and food transport vehicle collaboration strategy. J. Manage. Eng., 24(2): 75-78.

Ho, W., G.T.S. Ho, P. Ji and H.C.W. Lau, 2008. A hybrid genetic algorithm for the multi-depot vehicle routing problem. Eng. Appl. Artif. Intel., 21(4): 548-557.

Hu, T.L. and J.B. Sheu, 2003. A fuzzy-based customer classification method for advanced demand-responsive logistical distribution operations. Fuzzy Set. Syst., 139(2): 431-450.

Huang, M., X. Hu, Z. Wang and A.Z. Zeng, 2009. A three-stage solution approach to vehicle routing problem. J. Manage. Sci., 22(3): 37-46.

Liying, Z. and J. Haipeng, 2009. The multiple-depot heterogeneous-food transport vehicle scheduling algorithm based on global search clustering. J. Lanzhou Jiao Tong Univ., 28(6): 19-22.

Mirabi, M., S.M.T.F. Ghomi and F. Jolai, 2010. Efficient stochastic hybrid heuristics for the multi-depot food transport vehicle routing problem. Robot. Cim-Int. Manuf., 26(6): 564-569.

Tong, Z., L. Ning and S. Debao, 2004. The genetic algorithm of multiple-depot food transport vehicle routing problem. Comput. Eng. Appl., 40(21): 81-83.

Wei, S.H. and S.M. Chen, 2009. A new approach for fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. Expert Syst. Appl., 36(1): 581-588.

Yvcenur, G.N. and N.C. Demirel, 2011. A new geometric shape-based genetic clustering algorithm for the multi-depot food transport vehicle routing problem. Expert Syst. Appl., 38(9): 11859-11865.