

A Conservation Genetic Algorithm for Optimization of the E-commerce Logistics Distribution Path

Rui FU*, Mohammed Abdulhakim Al-Absi*, Ahmed Abdulhakim Al-Absi**, Hoon Jae Lee***

*Department of Ubiquitous IT, Graduate School, Dongseo University, 47 Jurye-ro, Sasang-gu, Busan 47011, Republic of Korea

**Department of Smart Computing, Kyungdong University 46 4-gil, Bongpo, Gosung, Gangwon-do, 24764, Republic of Korea

***Division of Information and Communication Engineering, Dongseo University, 47 Jurye-ro, Sasang-gu, Busan 47011, Republic of Korea

furui.qilienteng@gmail.com, Mohammed.a.absi@gmail.com, absiahmed@kduniv.ac.kr, hjlee@dongseo.ac.kr

Abstract— E-commerce is a business activity that uses modern information technology to process cash flow and logistics to achieve transactions. Under the e-commerce environment, an item can reach to the customer accurately is a key factor to enhance the company's image. Therefore, how to efficiently select the algorithm to solve the vehicle routing problem is an urgent problem to be solved. Aiming at the problems existing in the logistics distribution of e-commerce, this paper combines the genetic algorithm with the conservation algorithm based on the maximum travelable distance of the vehicle, and uses the conservation algorithm to generate the initial solution of the genetic algorithm, and constructs the conservation genetic algorithm to solve the e-commerce environment logistics distribution path optimization problem. The simulation results show that the conservation genetic algorithm is more globally optimal than the genetic algorithm.

Keywords— logistics and distribution, vehicle routing problem, saving algorithm, conservation genetic algorithm based on genetic algorithm

I. INTRODUCTION

In the traditional logistics distribution, there are unreasonable factors such as low automation, untimely information reception, and low network level, which hinder the efficiency of logistics distribution. The logistics distribution under e-commerce is a new logistics distribution mode that can shorten the distribution cycle, optimize the service quality and improve the competitiveness of enterprises by combining information network technology and logistics distribution in an open network environment. This kind of distribution mode is conducive to the improvement of logistics distribution efficiency. Logistics distribution under the e-commerce environment has many characteristics such as a large number of customers, relatively small customer demand, obvious customer dispensability, and time requirements.

E-commerce logistics and distribution is to meet the customer's satisfaction, and it is the basic goal of the enterprise to find ways to minimize the transportation cost during the operation [1], [2]. Characteristics of E-commerce logistics distribution are a large number of customers, size of each customer's demand is relatively small, dispersion of customers is obvious, customer has high requirements for the

timeliness of arrival of the goods, high customer service needs, and suppliers strive to save costs as showing in the Figure.1.

The Vehicle Routing Problem (VRP) [3] is a typical NP-hard problem. It has a wide range of applications in computer science, operations research and engineering optimization. The algorithms of solving VRP problems mainly include precise algorithms, traditional heuristic algorithms, simulation algorithms and artificial intelligence algorithms. The exact algorithm [4] uses the mathematical programming method to obtain the optimal path solution. The exact algorithm is suitable for solving small-scale VRP problems. The traditional heuristic algorithm [5] refers to a feasible solution to the combinatorial optimization problem to be solved under the premise of acceptable computation time and space. The simulation method [6] obtains the optimized solution of the vehicle route through computer simulation experiments. The artificial intelligence optimization algorithm [7] finds the optimal solution of the problem by revealing and simulating natural phenomena. It is global optimization performance, strong robustness, and strong versatility and is suitable for parallel processing. It is widely used to solve large-scale VRP problems. [8] E-commerce is a business activity that uses modern information technology to process cash flow and logistics to achieve transactions. Under the e-commerce environment, whether an item can reach to the customer accurately is a key factor to enhance the company's image. Therefore, how to efficiently select the algorithm to solve the vehicle routing problem is an urgent problem to be solved. In this paper, the genetic algorithm and the conservation algorithm [9], [10]-[11] are combined to construct the conservation genetic algorithm to study the logistics distribution path optimization problem under the e-commerce environment and the experimental results are verified in the simulation environment.

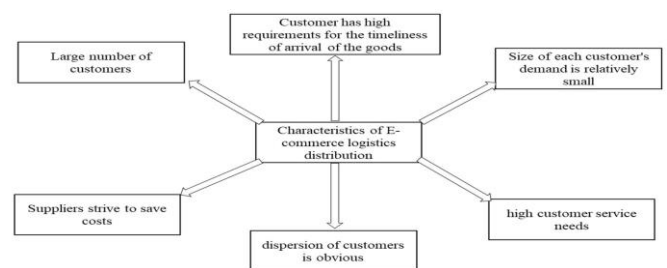


Figure. 1 Characteristics of E-commerce logistics distribution

II. PROBLEM DESCRIPTION

In the e-commerce environment, the logistics distribution path optimization problem can be described as follows: According to the customer's order status of the e-commerce enterprise, obtain the demand quantity of each customer's goods, determine the actual distribution network of the time period, and optimize the design of a vehicle. The delivery route completes the delivery tasks required by the customer under various constraints, so that the total cost of delivery is minimized.

In this paper, the maximum distance traveled by the vehicle is used as a limiting condition, and the saving algorithm and the genetic algorithm are combined to solve the optimal path. The delivery personnel take the list from the distribution center and deliver the goods to the customer points of known coordinates. When the total distance exceeds the maximum distance traveled by the truck, return to the distribution center to refuel, and then continue from the distribution center to the next waiting. The customer service department delivers the goods until all the customers are served, and finally the vehicle travels back to the distribution center and can quickly effectively meet customer needs.

A) Model hypothesis

- There is only one central site that provides delivery services and provides fueling services;
- The location coordinates of the central site and each customer are known;
- The demand for each customer point is known;
- Only one truck is available in the entire area;
- The sum of the distance traveled by the vehicle during the delivery process shall not exceed its maximum travelable distance;
- Each customer has one and only one visit;
- Each customer's distribution needs must be met.

B) Model establishment

Each customer's representation is

$$p_i = p(x_i, y_i), i = 1, 2, \dots, N$$

In the formula, x_i is the abscissa of the point where the customer is located; y_i is the ordinate of the point where the customer is located; N is the number of customers in the area who need the service. The coordinate point of each customer is known, so the distance between any two customers is known.

Define the decision variable as

$$x_{ij} = \begin{cases} 1, & \text{Passing the vehicle on the line}(i, j) \\ 0, & \text{Otherwise} \end{cases}$$

In the formula, (i, j) shows the car travelling from i to j .

1. The objective function for calculating the shortest path to travel is

$$D = \min \sum_{i=1}^N \sum_{j=1}^N d_{ij} \times x_{ij} + d_{1o} + s_o + d_{no} \quad (1)$$

In the formula, d_{ij} is the straight line distance between adjacent customers i and j ; d_{1o} is the distance from the distribution center to the first customer point; s_o indicating that when the vehicle travel route is limited by the maximum travel distance, the extra travel time is added back and forth; d_{no} is the distance from a customer to the distribution center.

2. Each customer is visited once.

$$s.t. \sum_{i=1}^N x_{ij} = 1, \quad (i \neq j; j = 1, 2, \dots, N) \quad (2)$$

$$\sum_{j=1}^N x_{ij} = 1, \quad (i \neq j; i = 1, 2, \dots, N) \quad (3)$$

In the formula, $\sum_{j=1}^N x_{ij}$ represents the total number of times from the customer i to the customer j (the value ranges from 1

to N); $\sum_{i=1}^N x_{ij}$ shows the total number of times from the customer i to the customer j (the value ranges from 1 to N). The combination of formula (2) and formula (3) means that each customer can only be served once, and the service cannot be repeated.

3. The maximum distance that a truck can travel at one time during the delivery process and cannot exceed the maximum distance that can be traveled that L is

$$\sum_{i=1}^k x_{ik} + x_{k0} \leq L \quad (4)$$

In the formula, x_{ik} means that the vehicle travels from the customer i to the customer k ; $\sum_{i=1}^k x_{ik}$ indicates the total distance that the vehicle has traveled to the point of k ; x_{k0} indicates the distance from the point k to the origin.

III. SAVING GENETIC ALGORITHM

A) Conservation algorithm

Conservation algorithm, also known as the C-W, which was proposed by Clarke and Wright [12] in 1964. Its basic idea is to first connect each point to the origin separately, c_{ij} the distance between the point i and the point j , from 1 to the total distance is

$$z = \sum_{i=1}^l c_{oi} + \sum_{i=1}^l c_{io} \quad (5)$$

Then define the savings of the distance between the points of i and j

$$s(i, j) = c_{oi} + c_{io} + c_{oj} + c_{jo} - (c_{oi} + c_{ij} + c_{jo}) = c_{io} + c_{oj} - c_{ij} \quad (6)$$

$$s(j, i) = c_{jo} + c_{oi} - c_{ji} \quad (7)$$

In the formula, the distance between the point and the point; for the path saving value, the steps of the saving algorithm are as follows:

1. Calculate the savings value of the route, arranged in a tabular form according to the value of the savings. And choose the maximum savings.

2. Check whether the two points corresponding to the saved value satisfy the following points:

- If the corresponding two points in the saved value are not on the already constructed line, get the line segment and turn to step 3.
- If the corresponding two points in the saved value are on the formed line and not the inner point of the line (i.e., not directly connected to the origin), the connection is obtained or the line is turned to step 3.
- If the corresponding two points in the saved value are on the different lines that have been constructed, and they are not internal points, the line is obtained and turned to step 3.
- If the corresponding two points in the saved value are on the formed line, then the connection cannot be made again, and turn to step 3.

3. If the row or column is removed, the point can no longer be connected to other points, and the point cannot be reached by other points.

4. After all elements are crossed, a qualifying line is obtained and the algorithm terminates. Otherwise, select the largest element from the elements that are not crossed, and turn to step 3.

B) Conservation genetic algorithm

The genetic algorithm [13] is an optimization algorithm formed during the natural evolution of Darwin's simulated organisms. The combination of the conservation algorithm and the genetic algorithm [14] can improve the efficiency of the genetic algorithm. The specific steps of the algorithm are as follows:

- Required information.
- Initial population

- Fitness function.
- Select.
- Cross.
- Variation.
- control parameter

IV. SIMULATION

It is assumed that within a certain period of time, a distribution demand order issued by 30 customers randomly selected in a region is received, and the maximum number of iterations is 200. Each order is analyzed by a conservation genetic algorithm, and the distribution center performs the vehicle on the vehicle loading to achieve logistics and distribution.

Each customer is numbered starting with 0 in decimal. And the demand for each customer is known.

Set the $N = 200$ $p_c = 0.9$ $p_m = 0.1$ $p_s = 0.6$ maximum evolution algebra gen_{max} to increase with the number of customers, assuming that the maximum driving distance of the vehicle is 30km. In order to verify the validity of the algorithm and obtain the optimal path, the results obtained by the two algorithms of genetic algorithm and conservation genetic algorithm are compared respectively. The results are shown in Table 1.

TABLE 1 SIMULATION CALCULATION RESULTS

Method	Customer	Average Best Value	Best Result	Relative Deviation /%	Average Iterations
GA	30	427.7	412.7	3.6	162
CWGA	30	410.4	401.3	2.2	72

Note: Relative deviation = $[(\text{Single measurement} - \text{Measurement average}) / \text{Measurement average}] * 100\%$ It can be seen from Table 1 that the average optimal value and the best result of the GA algorithm are larger than the CWGA algorithm when the number of customers is 30, which shows that the CWGA obtains the optimal path better; CWGA algorithm The relative error of the error is smaller than the relative error of GA, which indicates that the fluctuation of the CWGA result is smaller and more stable. The comparison of the average number of iterations shows that the convergence speed of the CWGA algorithm is much faster than that of GA. In summary, under a limited number of iterations, the algorithm CWGA as a new type of optimization algorithm is easier and faster to determine an optimized distribution path and reduce the delivery distance. The optimization process when the number of customers $N=30$ is as shown in Figure. 1.

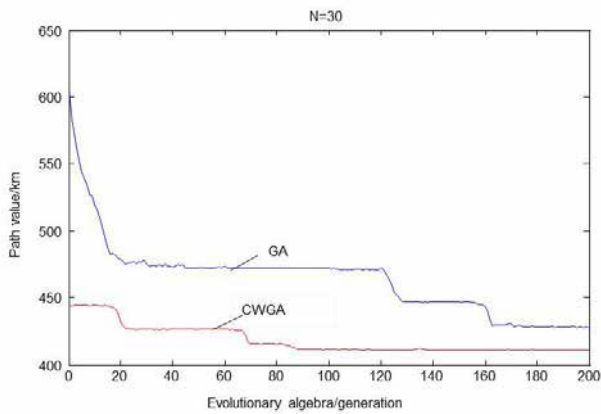


Figure 2 Optimization process when the number of customers is N=30

Figure 2 reflects the convergence of the path lengths of the two algorithms in the iterative process. The graph trends are gradually convergent. The CWGA algorithm requires less convergence than the GA algorithm, and runs to 200 generations. The shortest path finally obtained by CWGA is significantly improved compared to the shortest path obtained by the GA algorithm. Therefore, using the conservation algorithm to generate the initial population combined with the genetic algorithm makes the CWGA evolutionary starting point high and easy to find the optimal solution. 30 customer points are randomly selected. When considering the distance limit and not considering the distance limit, the route of the optimal path is obtained. One unit in the table indicates 1 km. The results are shown in Table 2.

TABLE 2. SIMULATION CALCULATION RESULTS

Customer	$N = 30$
Consider the distance limit	Order : 2 12 7 0 23 4 18 19 17 24 20 11 25 29 6 3 13 22 9 10 14 27 28 5 1 26 15 8 21 16 Distance : 413.7km
No distance limit	Order : 2 12 7 0 23 4 17 18 19 5 24 20 6 11 25 16 29 3 13 9 10 14 27 28 22 1 26 15 8 21 Distance : 76.7km

In the conservation genetic algorithm, the population size (PS) is fixed, and this process does not conform to the actual biological evolution in the process of biological evolution. The population size cannot be constant in the process of human and biological evolution. Only one fixed population size in an algorithm cannot visually compare the results of the comparison. Poor population size will affect the quality of the solution. When the population size increases, if the set evolutionary algebra increases, the efficiency and accuracy of the conservation genetic algorithm do not necessarily increase.

The evolutionary algebra of the conservation genetic algorithm is the number of cycles set in the program. A certain number of cycles will converge to a certain interval, which is the convergence algebra (EGN). The more the number of

loops, the more values the algorithm traverses, and the results obtained are not necessarily closer to the optimal value.

3) $N=30$, $P_c=0.9$, $P_m=0.1$, the program runs 5 times to find the shortest path, convergence algebra (EGN) and population size (PS) are shown in Table 3 and Figure 3:

TABLE 3. CHANGES IN POPULATION SIZE AND CONVERGENCE ALGEBRA

PS	EGN	Optimal path
30	150	298.17
60	142	239.09
90	146	227.19
120	132	218.97
150	124	209.61
180	126	228.69
210	107	299.15
240	113	311.80
270	105	293.80
300	104	291.09

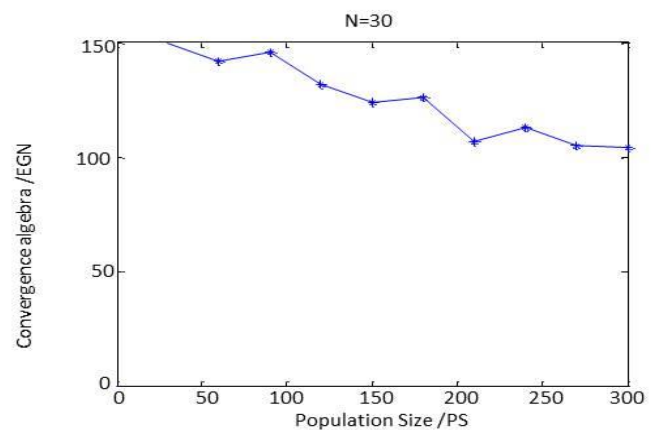


Figure. 3 Changes in population size and convergence algebra

V. CONCLUSION

Taking the e-commerce B2C model as the research background, the different needs of different customers in a certain area are the research objects, and the mathematical model is established. Combined with the saving genetic algorithm, the influence of the maximum distance that the truck can travel on the vehicle distribution problem is detailed. Processing and obtaining an optimal path that is more realistic than previous studies. Future work, there are still a lot of deficiencies where the problem algorithm itself still needs a long calculation time, and the calculation degree is more complicated.

ACKNOWLEDGMENT

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science, and Technology (grant number: NRF2016R1D1A1B01011908). And it was also supported by the Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2018-0-00245).

REFERENCES

- [1] Li Weijian. Research on logistics distribution route optimization under B2C e-commerce mode [D]. Beijing: Beijing Jiaotong University. 2007.
- [2] Ding Liyan, Zhang Wei. Logistics Foundation [M], Beijing: Tsinghua University Press, 2000, 78-79.
- [3] Cui Xueli, Ma Liang, Fan Bingquan. Ant colony algorithm for vehicle routing problem (VRP) [J]. Systems Engineering, 2004, 19(4): 418-422.
- [4] Sun Lijun, Hu Xiangpei, Wang Zheng. Research progress on vehicle path planning problems and solving methods [J]. Systems Engineering, 2006 (24): 31-36.
- [5] Li Yufeng, Li Jun. Dynamic programming heuristic algorithm for solving time-varying vehicle scheduling problems [J]. Systems Engineering Theory and Practice, 2012 (32): 1712-1718.
- [6] Liu Xiang, Li Yanhui. Cross-regional VRP model for e-commerce distribution and its heuristic algorithm [J]. Journal of Tsinghua University, 2006, 46(z1): 1014-1018.
- [7] Ratliff H D, Rosenthal A S. Order picking in a rectangular warehouse: A solvable case of the traveling salesman problem [J]. Operations research, 2002, 2 (31): 507-521.
- [8] Wang Dingwei. Modeling and optimization in e-commerce [M]. Science Press, 2008.
- [9] Yuan Jian, Liu Jin, Lu Houqing. Annealing network solution for VRP in stochastic demand situation [J]. Systems Engineering Theory and Practice, 2002, 22(3): 109-114.
- [10] Liu Jiali, Ma Zujun. There are vehicle rental and sharing and there are time windows and multiple distribution centers open loop VRP [J]. Systems Engineering Theory and Practice, 2013, 33(3): 666-672.
- [11] Lin Yuqin, Wu Peiru. Research on the traditional storage picking path of logistics center [D]. Supplementary city name: Taiwan Mingxin University of Science and Technology, 2005.
- [12] Clarke G, Wright J W. Scheduling vehicle from a central delivery depot to a number of delivery points[J]. Operations Research, 1964, 32(8): 568-581.
- [13] Xie Binglei, Sun Yi, Li Rongxi. Genetic algorithm for solving the problem of collecting travel salesman [J]. Journal of Shaanxi Institute of Technology, 2002, 18(1): 70-75.
- [14] Lang Maoxiang, Hu Siji. Research on solving the problem of logistics distribution route optimization by hybrid genetic algorithm [J]. Chinese Management Science, 2002, 10(5): 51-56.
- [15] R.Fu, Mohammed Abdulhakim Al-Absi, A. A. Al-absi, M. Sain, H. J.Lee., "General study of Secure E-commerce Logistics Distribution", International Conference on Culture Technology, Zhejiang College of Zhejiang University of Technology Shaoxing, China, ICCT Nov. 15~18 pp 404-407, 2018



Rui FU received his (MS) degree in System Theory from Qingdao University -China in 2012-2015. Currently, She is a Ph.D. student in the Department of Information and Communication Engineering at Dongseo University, Korea. Her research interests include Logistics Transportation and Mathematics.

Mohammed Abdulhakim Alabsi received his BS degree in Computer Application from Bangalore University in India. He earned his (MS) degree at Dongseo University- South Korea in 2018. Currently, he is a Ph.D. student in the Department of Information and Communication Engineering at Dongseo University, Korea. His research interests include IoT, VANET, Cryptology, Network Security, Computer Networks and Digital Communications.



Ahmed Abdulhakim Al-Absi is an assistant professor and head of smart computing department at Kyungdong University - Global Campus in South Korea. He earned his Ph.D. degree in ubiquitous computing at Dongseo University- South Korea in 2016. His research interests include Database Systems, Big Data, Hadoop, Cloud computing, Distributed systems, Parallel computing, High-performance computing, VANET, and bioinformatics. He received a Master of Science (MS) degree in information technology at University Utara Malaysia-Malaysia in 2011, and a Bachelor of Science (BS) degree in computer applications at Bangalore University- India in 2008.



HoonJae Lee received his BS, MS, and Ph.D. degrees in electronic engineering from Kyungpook National University, Daegu, Rep. of Korea, in 1985, 1987, and 1998, respectively. He is currently a professor in the Department of Information Communication Engineering at Dongseo University. From 1987 to 1998, he was a research associate at the Agency for Defense Development (ADD). His current research interests include Password Theory, Network Security, Side-Channel Attack, and Information Communication/Information Network.

