https://doi.org/10.31449/inf.v48i18.6436 Informatica **48** (2024) 140–154 **140**

Optimization of Cigarette Logistics Paths Using Hybrid GA-A* Algorithm

Delun Shi, Guangjun Dong, Enbo Chen, Ming Dai, Ni Xiao, Yi Zhang, Wei Chu*
Wuhan Cigarette Factory, China Tobacco Hubei Industrial LLC, Wuhan 430040, China
E-mail: chuwei830826@163.com

Keywords: GA, A* algorithm, cigarette logistics, Storage, distribution path

Received: June 19, 2024

The logistics distribution and storage of finished cigarette products is a key link to ensure the stable supply of the tobacco market and the healthy development of the industry. Aiming at the loss problem of finished cigarette products during transportation, this paper proposes a method for optimizing the logistics distribution and storage path of finished cigarette products based on the improved Genetic Algorithm (GA) and A-star(A) algorithm. This method first introduces a cost calculation model to calculate the loss of finished cigarette products during transportation, and uses the A* algorithm to solve the distribution in different areas. Then, the A* algorithm is combined with GA to construct an optimal path planning model based on minimum cost. Through experiments on the Solomn dataset and the Gehring dataset, the proposed method reached the minimum objective function value at 41 and 32 iterations, and showed a fast convergence speed. In the performance evaluation, the area under the ROC curve values of the research method reached 0.985 and 0.967, respectively, showing high accuracy. In addition, the path planning error analysis showed that when the iteration was carried out to the 27th time, the error value dropped to 0.06, which met the performance requirements. In practical applications, the system began to stabilize and reached the optimal state after about 47 iterations. The above results show that the research method has a faster convergence speed, smaller planning error and higher accuracy in the logistics distribution path planning of finished cigarette products, with good feasibility and effectiveness.*

Povzetek:Raziskava predlaga metodo za optimizacijo logističnih poti shranjevanja končnih cigaretnih izdelkov z uporabo hibridnega algoritma GA-A. Uvedena je metoda izračuna stroškov, ki določa izgube med transportom, in kombinacija A* algoritma za iskanje najkrajših poti ter genetskega algoritma (GA) za iskanje optimalnih rešitev. Metoda izboljšuje hitrost dostave, zmanjšuje napake in poveča uporabnost pri distribuciji cigaret.*

1 Introduction

Optimizing the logistics and distribution storage path can improve the efficiency of distribution [1]. Nowadays, the distribution of finished cigarette products is facing many challenges and reforms. Traditional logistics methods are no longer suitable for modern market demand, and optimizing logistics paths can improve logistics delivery efficiency. Therefore, tobacco logistics and distribution storage path optimization has become an important issue [2]. Meanwhile, there are many problems with optimizing the logistics and storage routes for finished cigarettes, including the difficulty of standardizing distribution costs, the high complexity of transport routes and the huge statistical workload. Due to the delicate nature of the production process, the transportation and storage of finished cigarettes in the distribution process deserves attention to ensure that finished cigarettes are delivered to the customer's point of demand at a faster rate [3-4]. Some scholars believe that the crucial point to enhancing the efficiency and reducing the cost of finished cigarette logistics is to reduce the transportation costs, losses and distribution routes of finished cigarettes during the distribution process. This requires that cigarette distribution needs to be arranged in a reasonable and efficient way that meets the customer's time requirements. Currently, the usual approach to solving this problem is to transform the route optimization issue into a mathematical model issue and then use modern heuristic algorithms to find the optimal solution. Modern heuristic algorithms include Genetic algorithm (GA), A-star(A*) algorithm ant colony algorithms, particle swarm optimization algorithms, neural network algorithms, etc. [5]. These algorithms have their own advantages and disadvantages, such as the particle swarm optimization algorithm has a strong global optimization capability, but no proof of convergence. The ant colony algorithm is robust but has a slow search time. The neural network algorithm is adaptive but needs quantities of data. Considering that the GA can quickly find the combined optimal solution and the A * algorithm can search for the shortest path in the grid, the study proposes a storage path method based on the improved GA-A* algorithm for the distribution of finished cigarette logistics, with a view to effectively providing some promotion for the construction as well as the development of China's finished cigarette distribution system.

2 Background review

As cigarette consumption in China increases year by year, the demands for the distribution and storage of cigarettes are getting higher. To enhance the distribution efficiency of finished cigarette products and reduce storage costs, many experts studied logistics, distribution, and warehousing. Liu proposed a method that combines RFID technology with logistics and distribution storage management to improve the management efficiency of logistics and distribution storage in warehouses. It was verified that the method had higher warehouse storage capacity and faster outgoing and incoming storage speed compared with the traditional method [6]. Shen proposed to establish a dynamic road network logistics and distribution path optimization model for the dynamic uncertainty of urban road networks. The model usedGA to handle the optimization model. The model could greatly enhance the transportation efficiency of urban logistics [7]. Jardas et al. proposed a method to optimize the transportation process to reduce the burden of the city's transportation network. The method used a quantity of data analysis and the centre of gravity method to determine the location of distribution centre and thus optimize transport routes. It was verified that the method could reduce local transport pressure [8]. Yang et al. solved the Vehicle Routing Problem (VRP) by proposing an improved ant colony algorithm. The method adaptively adjusted key parameters in the application and searches for optimal paths. The algorithm could minimize paths to reduce computational costs [9]. Liu et al. constructed an integer programming model to address the problems of backward technology, poor management and serious energy consumption in cold chain logistics, and applied a hybrid model to find the optimal solution. The hybrid model had obvious advantages over GA [10]. Peng et al. proposed a method for adding a global performance index of smoothness at joint acceleration level to address the functional redundancy problem arising when planning robot paths. The method used a sequential linearized planning approach to improve on the traditional method and provided an initial solution that is then used for robot path planning. It was verified that the model generated smoother robot paths [11].

The advantages of the A* algorithm include the ability to handle search problems in high-dimensional spaces and can be used in many application areas. Min et al. addressed the shortcomings of the A* algorithm in autonomous driving path planning applications and proposed a vehicle local motion planning algorithm. The algorithm setup a safety space and considers path curvature. The algorithm was easier to obtain better constrained paths for vehicles [12]. Beed et al. proposed a hybrid GA-a* algorithm for optimizing the carpool path selection problem. The algorithm could provide the shortest route between any two points. Practice showed that the algorithm could greatly improve vehicle utilization and reduce the diversions distance and cost [13]. Yue et al. addressed the problem of inaccurate navigation positioning and more path folds for mobile robots. They proposed a mobile smooth navigation strategy. The strategy enhanced the A* algorithm to a bidirectional mode and then used the Bézier curve to optimize the path. The results showed that the strategy could reduce the running time [14]. Meng et al. suggested an enhanced hybrid A* algorithm. This algorithm integrated Voronoi field potential into the path search stage and dynamically optimized it at each stage. It was verified that this algorithm significantly improved the efficiency of path search [15]. Zhang et al. proposed to construct a dynamic network radar model to improve the path safety and penetrating path search efficiency of UAVs. The model used the proposed penetrating path planning method to plan UAV way-points. The results showed that the method had optimal path cost and higher safety [16].

From the above research, many experts and scholars have designed quantities of improved algorithms for optimizing logistics distribution storage paths. Currently GA and A* algorithms have been widely used in many fields, with most of the research revolving around tourism and other areas of related research, with few areas analyzing the logistics distribution paths of finished cigarette products in the tobacco industry. In this regard, the research proposes an improved GA-A*-based method for optimizing the distribution and storage paths of finished cigarette logistics, in the expectation that the rational planning of the distribution and storage paths will facilitate the development of transportation in the tobacco industry. Summary of related work is shown in Table 1.

Table 1: Summary of related work

Author(s)	Method/Algorithm	Shortcomings	Observations
Liu [6]	A method of combining RFID technology with logistics distribution storage management	Time-consuming	Higher storage capacity and faster entry and exit speeds
Shen [7]	A dynamic road network logistics distribution path optimization model	High vehicle exhaust emissions	Improving the transportation efficiency of urban logistics
Jardas et al. [8]	A method for optimizing	More data are required	Reduce transportation

	the transportation process		stress
Yang et al. [9]	An improved ant colony algorithm	Limited	Reduce computing costs
Liu et al. [10]	A mixed integer programming model	Complicated solution	Obvious advantages
Peng et al. [11]	A method for adding a global performance indicator of smoothness at the level of joint acceleration	High accuracy requirements for robot operation	Smoother path
Min et al. [12]	Vehicle local motion planning algorithm based on improved A* algorithm	Path curvature must be considered	Better vehicle constrained paths
Beed et al. [13]	A hybrid GA-A* algorithm	The shortest path between two points are required	Vehicle utilization is improved, detour distances and costs are reduced
Yue et al. [14]	Mobile smooth navigation strategy based on A* algorithm	A* algorithm is improved to bidirectional mode	The planned path length is shortened and the running time is reduced
Meng et al. [15]	An improved hybrid A* algorithm	Time-consuming	The path search efficiency is improved significantly
Zhang et al. [16]	Building a dynamic network radar model	The penetration path planning method needs to be adopted	Optimal path cost and higher security

3 Construction of an improved GA-A*-based storage path optimization model for finished cigarette logistics distribution

3.1 Cigarette finished logistics transportation cost path optimization modeling

The finished cigarette logistics route optimization problem is a complex problem derived from the VRP. Before planning the path for the finished cigarette logistics, all costs incurred during transportation need to be considered. Based on the constraints of customer demand, time window requirements and the maximum mileage of refrigerated trucks, a path optimization model for the logistics transportation cost of finished cigarettes is constructed [17]. The distribution route schematic of the VRP problem is shown in Figure 1.

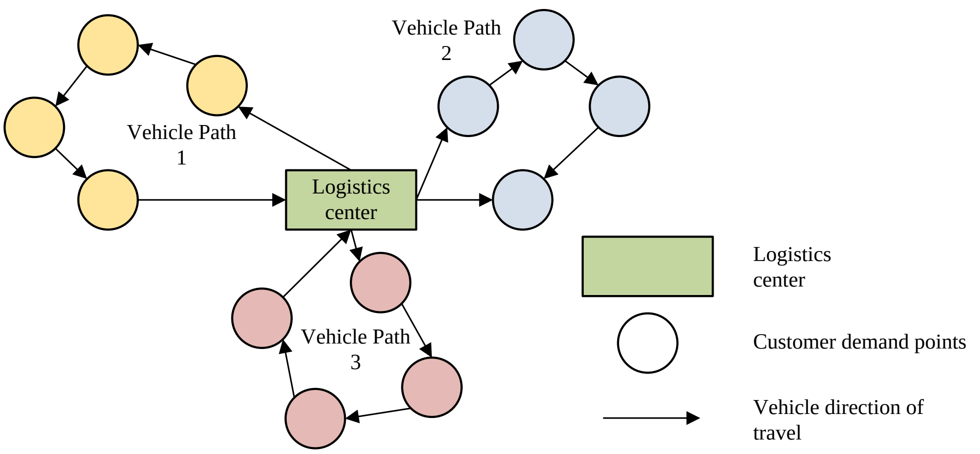


Figure 1: VRP distribution roadmap

The location and coordinates of the customers are known, because the demand for finished goods varies from customer to customer. The goods carried by each delivery vehicle must not exceed its specified capacity and must be delivered within the time required by each customer. Therefore, the distribution route and sequence must be scientifically and reasonably optimized to ultimately achieve the purpose of reducing distribution costs. The key to optimizing the logistics distribution route is in the optimization of logistics routes and logistics costs. Logistics cost optimization includes reducing the fixed cost of vehicles and transport, optimizing vehicle delivery time and reducing the cost of goods damage. The calculation of fixed costs of vehicles and transport is shown in equation (1).

$$TC_1 = K * \delta + \sigma \sum_{i=0}^{n} \sum_{j=1}^{n} \sum_{k=1}^{M} d_{ij} x_{ijk} \quad (1)$$

In equation (1), TC_1 is the fixed cost of the vehicle and transport; $i = [0.1.2.3....n]$ is the distribution point number including the distribution centre; $j = [0.1.2.3....n]$ is the distribution point number excluding the distribution centre; $K = [0.1.2.3....n]$ is the number of vehicles participating in this distribution; δ is the fixed cost of each vehicle; σ is the cost of each vehicle participating in the distribution service; d_{ij} is the distance from point i to point j. There are overtime compensation costs for the delivery of finished cigarettes en route. The overtime compensation cost is the cost incurred by the customer when the delivery vehicle fails to reach the delivery point on time, causing the customer to penalize the company for the loss. The study uses a hybrid time window to classify delivery times into desired and acceptable times [18]. Figure 2 illustrates the relationship between customer satisfaction and delivery time.

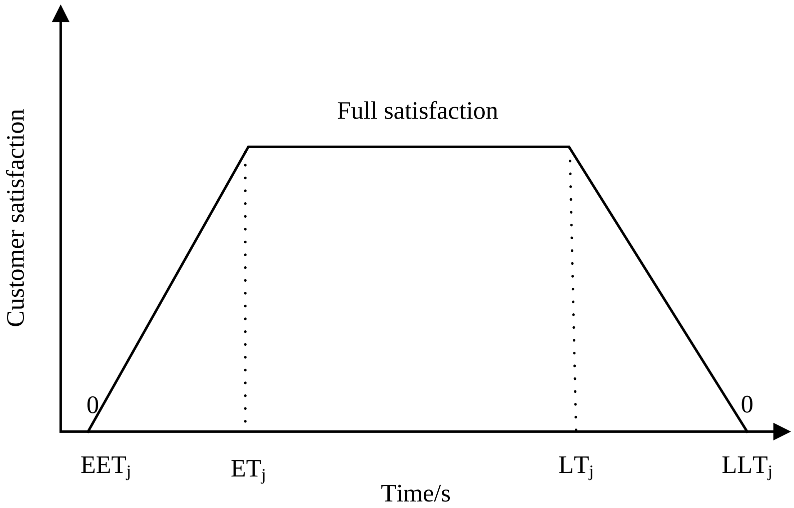


Figure 2: Relationship between customer satisfaction and delivery time

In Figure 2, EET_j is the earliest point at which the customer is acceptable; ET_j is the earliest point at which the customer is satisfied; LT_j is the latest point at which the customer is satisfied; and LLT_j is the latest point at which the customer is acceptable. If ET_j and LT_j are delivered, the customer is very satisfied, while if EET_j and LLT_j are delivered outside the time period, the customer satisfaction is 0. Let the penalty cost for early arrival be u_1 and the penalty cost for late arrival be u_2. Based on the relationship between customer satisfaction and delivery time and the penalty cost in Figure 2, a time window penalty cost segmentation function can be constructed to calculate the formula in equation (2).

$$\begin{cases} M & t_j < EET_j \\ u_1\left(ET_j - t_i\right) & EET_j < t_j < ET_j \\ 0 & ET_j < t_j < LT_j \\ u_2\left(t_j - LT_j\right) & LT_j < t_j < LLT_j \\ M & t_j > LLT_j \end{cases} \quad (2)$$

In equation (2), M is an infinite amount. The total cost of the time window penalty can be obtained from equation (2) in equation (3).

$$TC_2 = 2M + u_1 \sum_{j=0}^{n} \max\left\{ET_i - t_i, 0\right\} + u_2 \sum_{j=0}^{n} \max\left\{t_i - LT_i, 0\right\} \quad (3)$$

In equation (3), TC_2 is the time window penalty cost. During transportation and unloading, the cost of damage to the finished cigarette product is very likely to be caused by crushing and deformation, moisture deterioration and other problems. Assuming a constant rate of goods damage, the longer the time, the greater the rate of goods damage. Equation (4) indicates the calculation.

$$\begin{cases} E_1 = \sum_{k=1}^{m}\sum_{j=0}^{n} q_j x_{jk}\left(1-e^{-\partial_1\sum_{i=0}^{n}\sum_{j=0}^{n}T_{ijk}^{t}Z_{ik}}\right) \\ E_2 = \sum_{k=1}^{m}\sum_{j=0}^{n} Q_j x_{jk}\left(1-e^{-\partial_2 t_{sj}^{k}}\right) \\ TC_3 = A\left(E_1+E_2\right) \end{cases} \quad (4)$$

In equation (4), TC_3 is the cost of damage; is the cost of transport damage; E_1 and E_2 are the cost of damage during loading and unloading; q_j is the site demand; t_j is the time for the vehicle to arrive at the site j; Q_t is the cargo load; A is the price of the product; T_{ijk}^{t} is the time for the first k vehicle to travel from i to j at t; and t_{sj}^{k} is the service time for the vehicle k at the site j. The objective function model for the total cost of transportation of finished tobacco products based on the fixed cost of vehicles and transportation, the total cost of time window penalties, and the cost of damage is shown in equation (5).

$$\min TC = TC_1 + TC_2 + TC_3 \quad (5)$$

This total cost minimization function is subject to the constraints of the condition in equation (6).

$$\begin{cases} \sum_{j=1}^{n} y_{jk} = 1 & k\in M \\ \sum_{j=1}^{n} x_{ijk} = \sum_{j=1}^{n} x_{ijk} & i=0 \quad k\in M \\ \sum_{i=0}^{n}\sum_{j=0}^{n} x_{ijk} \le 1 & k\in M \\ \sum_{i=0}^{n}\sum_{j=1}^{n} x_{ijk} - \sum_{i=0}^{n}\sum_{j=1}^{n} x_{i(j-1)k} = 1 & k\in M \\ \sum_{j=1}^{n} y_{jk} q_j \le Q & k\in M \\ \sum_{i=0}^{n}\sum_{j=1}^{n} x_{ijk} d_{ij} \le L & k\in M \\ ET_j < t_j < LT, or EET_j < t_j < ET, or LT_j < t_j < LLT_j \end{cases} \quad (6)$$

In equation (6), $\sum_{j=1}^{n} y_{jk} = 1$ constrains all vehicles not to deliver to the same distribution point; $\sum_{j=1}^{n} x_{ijk} = \sum_{j=1}^{n} x_{ijk}$ constrains all distribution vehicles to start and end at a distribution centre; $\sum_{i=0}^{n}\sum_{j=0}^{n} x_{ijk} \le 1$ constrains all vehicles to start from a fixed location and deliver to all sites from that location; $\sum_{i=0}^{n}\sum_{j=1}^{n} x_{ijk} - \sum_{i=0}^{n}\sum_{j=1}^{n} x_{i(j-1)k} = 1$ constrains vehicles to schedule the next site immediately after delivering to a site; $\sum_{j=1}^{n} y_{jk} q_j \le Q$ constrains all vehicles to deliver to the maximum number of sites; $\sum_{i=0}^{n}\sum_{j=1}^{n} x_{ijk} d_{ij} \le L$ constrains vehicles to deliver over a range of distances. $ET_j < t_j < LT, or EET_j < t_j < ET, or LT_j < t_j < LLT_j$. The constraint vehicle must complete the distribution within the specified time. Logistics distribution route optimization means that the total transport distance for the entire distribution process is required to be minimized. With k car distribution; car load for Q_k ; vehicle a distribution of the most travel distance for D_k ; site cargo demand for q_i ; site i to site j distance for d_{ij} ; distribution centre to site i distance for d_{oi} ; the k car distribution site number n_k ; k path for R_K ; site in the route k order i for r_{ki} . A functional model of the shortest distribution path is shown in equation (7).

$$Z = \sum_{k=1}^{k}\left[\sum_{i=1}^{n_k} d_{r_{k(i-1)}r_{ki}} + d_{r_{knk}r_{ko}} \bullet sign\left(n_k\right)\right] \quad (7)$$

The constraints of equation (7) are shown in equation (8).

$$\begin{cases} \sum_{i=1}^{n_k} d_{r_{k(i-1)}r_{ki}} + d_{kn_k} r_{ko} \bullet sign\left(n_k\right) \le D_k \\ 0 \le n_k \le 1 \\ \sum_{i=1}^{k} n_k = L \\ R_k\left\{r_{ki} \middle| r_{ki} \in \{1,2,...,n_k\}, i=1,2,3...,n_k\right\} \\ R_{k1} \cap R_{k2} = \varnothing, \forall k_1 \ne k_2 \\ sign\left(n_k\right)\begin{cases} 1 & n_k \ge 1 \\ 0 & \end{cases} \end{cases} \quad (8)$$

In equation (8), $\sum_{i=1}^{n_k} d_{r_{k(i-1)}r_{ki}} + d_{kn_k} r_{ko} \bullet sign\left(n_k\right) \le D_k$ constrains the distance of each distribution route not to exceed the maximum distance travelled by the vehicle transport; $0 \le n_k \le 1$ constrains that each route stop does not exceed the total stops; $\sum_{i=1}^{k} n_k = L$ constrains that each distribution point is delivered; $R_k\left\{r_{ki} \middle| r_{ki} \in \{1,2,...,n_k\}, i=1,2,3...,n_k\right\}$ represents the stop composition of each route; $R_{k1} \cap R_{k2} = \varnothing, \forall k_1 \ne k_2$ constrains that each stop is delivered by one vehicle; $sign\left(n_k\right)\begin{cases} 1 & n_k \ge 1 \\ 0 & \end{cases}$ constrains whether the vehicle is involved in the delivery.

3.2 Fusing Improved GA-A* algorithm for finished cigarette logistics route transport storage path

After planning for the transportation costs of finished cigarette logistics, the specific distribution and storage paths of the finished cigarette products in the transportation process also need to be analyzed. Compared to other algorithms, GA is able to manipulate parameter codes and find the optimal solution quickly. However, the algorithm also suffers from the problem of prematurely falling into a local optimum solution. The study therefore applies the A* algorithm to logistics distribution route planning, incorporating the GA to solve for the paths [19]. The general flow of the A* algorithm is shown in Figure 3.

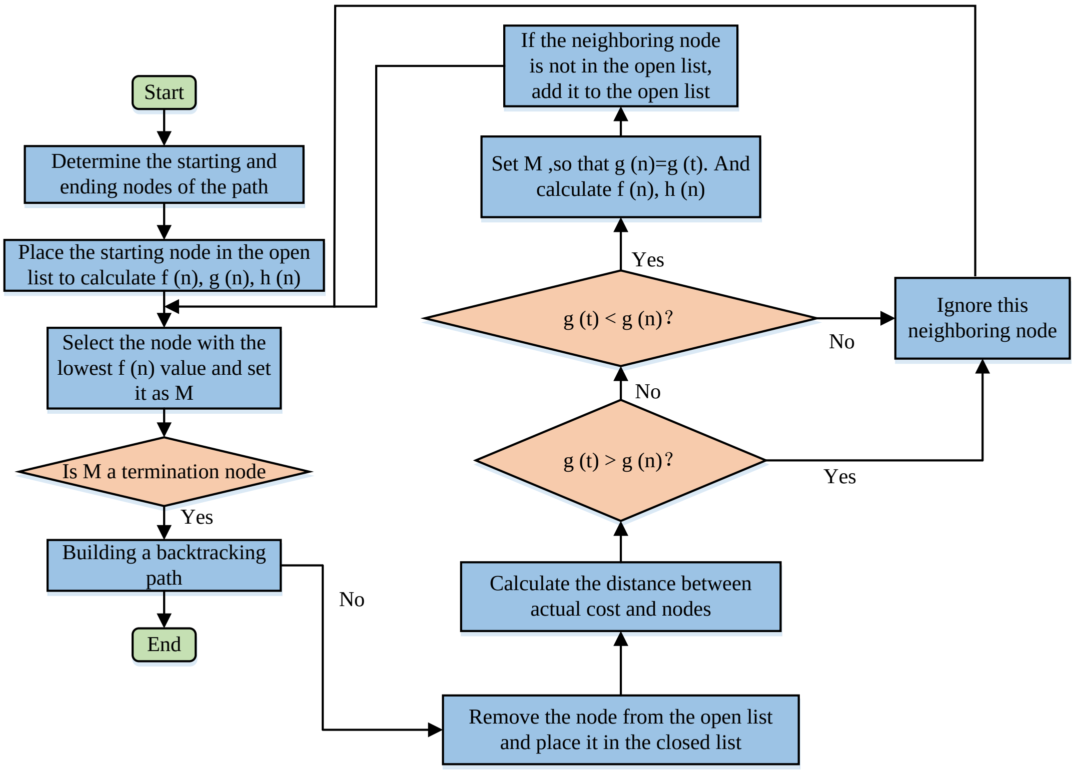


Figure 3: A * Algorithm flow diagram

As shown in Figure 3, the general process of the A* algorithm is to grid mark the target area with target points in the run path as A_1, A_2, A_3, ... , A_n, the grid map is set for R rows C columns, then the node coordinates are calculated in the way shown in equation (9).

$$\begin{cases} x = a \times \left[\mathrm{mod}(N, C) - \dfrac{a}{2} \right] \\ y = a \left[R + \dfrac{a}{2} - ceil\left(\dfrac{N}{R} \right) \right] \end{cases} \tag{9}$$

In equation (9), a is the length of each grid; mod is the remainder function; $ceil$ returns the smallest integer of the values. The start and end points are then determined and the starting point is placed in an open list to calculate the evaluation function $f(n)$, the actual cost of the node to the starting point $g(n)$ and the estimated cost of the target point to the node $h(n)$. Equation (10) gives the details of the calculation.

$$\begin{cases} f(n) = g(n) + h(n) \\ g(n) = \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2} \\ h(n) = \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2} \end{cases} \tag{10}$$

It then determines if the node is the termination point and, if so, returns to the previously passed node and constructs the path.Otherwise, the node M is placed in the closing list, and the actual cost of the node M is calculated with the distance from the node to the adjacent nodes at this point $g(t)$; Finally, whether the adjacent nodes are qualified is determined. For the traditional A* algorithm to construct paths with a large number of redundant turns and invalid points, the study uses the setting of parameters to limit the search direction of neighboring nodes to improve. The heuristic function is a key part of the A* algorithm, which is used to evaluate the distance from the current node to the target node. The study introduces weights to the heuristic function and adds a new evaluation metric based on this $h(p)$. Equation (11) shows the detailed calculation process.

$$f(n) = g(n) + \alpha\left[h(n) + h(p) \right] \tag{11}$$

In equation (11), α is the weight value; $h(p)$ is the distance from the parent node of the current child node to the target. The general flow of the GA is shown in Figure 4.

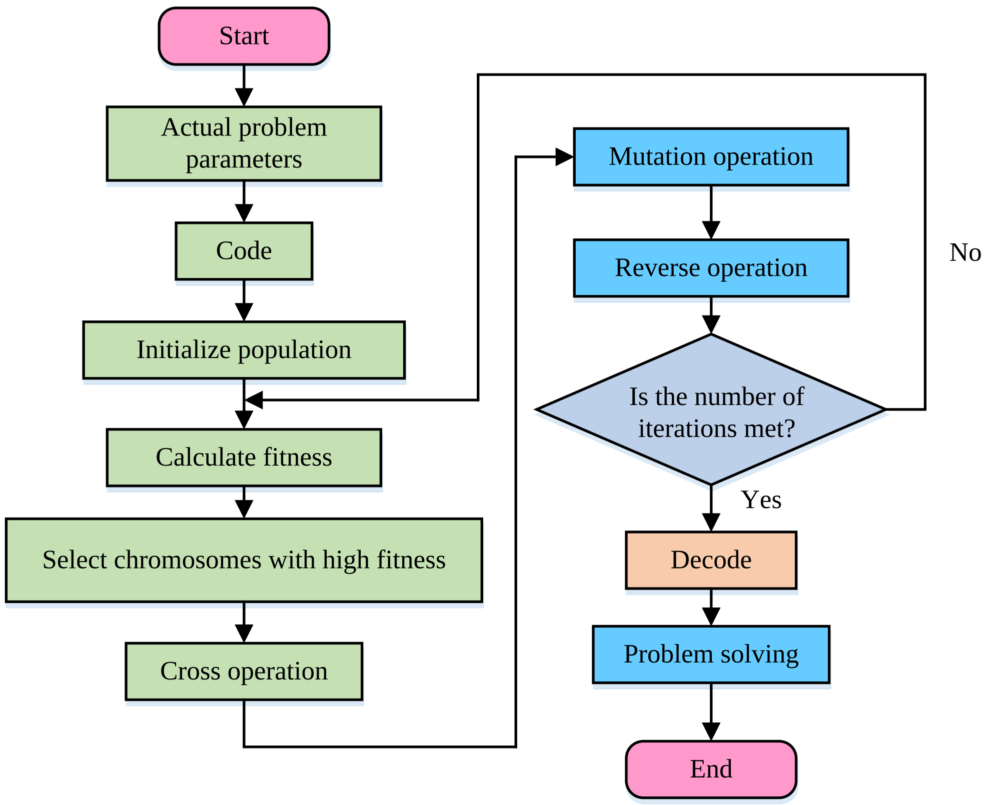


Figure 4: Operation process of GA

As shown in Figure 4, the general process of the traditional GA is to first encode the target point. The population is initialized, and then the fitness is calculated. Individuals with high fitness are selected. Crossover and mutation operations are carried out. All individuals are evaluated, and the fitness value of individuals is determined according to the fitness function. The above steps are repeated until the target problem result is selected. In terms of coding and decoding, the practical problem of logistics path optimization is not directly understood by traditional GA. Let a chromosome of length X replace X distribution points, i. e. each gene refers to one distribution point. When encoding, the solution to the distribution path problem is transformed into the genotype string structure data in the GA. When decoding, a route is constructed at each gene corresponding to as many distribution points as possible, and a new route is constructed if the route does not hold [20]. For example, the coding of 8 distribution points 51273684 is shown in equation (12).

$$\begin{cases} 0-5-1-2-0 \\ 0-7-3-6-0 \\ 0-8-4-0 \end{cases} \tag{12}$$

In equation (12), $0-5-1-2-0$ is path 1; $0-7-3-6-0$ is path 2; $0-8-4-0$ is path 3; and 0 represents the distribution centre. For the initial group, the study is improved using a savings algorithm. The algorithm takes the distance that can be saved by distribution points, inserts distribution points that are not on the distribution route into the route, and repeats the operation until all distribution points are on the distribution route. The calculation of the savings value for two nodes in the savings algorithm is shown in equation (13).

$$S(i,j) = d(i,0) + d(o,j) - d(i,j) \tag{13}$$

In equation (13), 0 is the distribution centre location; i and j are both distribution point locations. For the selection operator, perform cross operation This strategy not only ensures that the selected individual is optimal, but also guarantees its randomness and diversity. Equation (14) shows the specific calculation.

$$p_i = \frac{f_i}{\sum_{i=1}^{n} f_i} \tag{14}$$

In equation (14), f_i is the adaptation of distribution individuals i; n is the total number of distribution individuals; $\sum_{i=1}^{n} f_i$ the sum of all distribution individual's adaptation. In response to the problem that the traditional crossover operator affects the convergence of the algorithm, the study uses adaptive crossover probability. The method automatically selects the crossover probability depending on the actual situation. The specific calculation is shown in equation (15).

$$P = \begin{cases} P_c & f < \bar{f} \\ \dfrac{P_c\left(f - f_{\max}\right)\left(f - 2\bar{f} + f_{\max}\right)}{\left(f_{\max} - \bar{f}\right)\left(\bar{f} - f_{\max}\right)} & f > \bar{f} \end{cases} \quad (15)$$

In equation (15), P_c is the set crossover probability; $f_{\max}$ is the maximum fitness value; $\bar{f}$ is the average fitness value per individual per generation; and P is the desired adaptive crossover rate. The study uses the inversion method for the variation operator, i.e. two randomly selected variation points are inverted to obtain new individuals. The method is described in Figure 5.

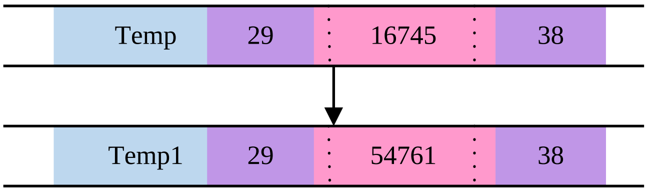
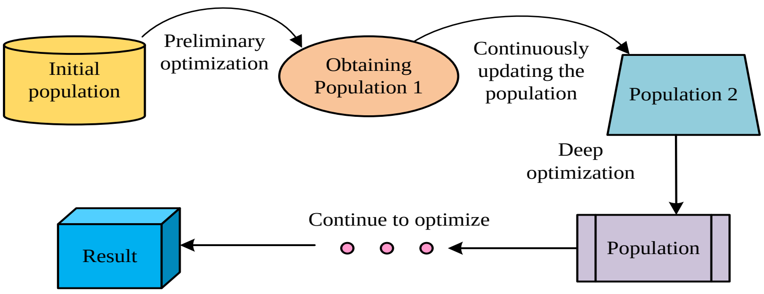


Figure 5: Mutation operation

For the individual assessment, the study introduces a new fitness function for the individual at the distribution point, which is used to evaluate the strengths and weaknesses of that individual. Equation (16) indicates the specific calculation.

$$f(x) = C_{\max} - \left(\alpha \sum_{k=1}^{k}\sum_{i=0}^{i}\sum_{j=0}^{j} C_{ij} X_{ijk} + \beta \sum_{k=1}^{k}\sum_{j=0}^{j} X_{ojk} + \gamma \sum_{i=1}^{k} P_i(S_i)\right) \quad (16)$$

In equation (16), $f(x)$ is the adaptive degree value; $C_{\max}$ is the maximum value of $g(x)$ in this generation; $g(x)$ is the objective function value. The specific optimization process and code flow are finally shown in Figure 6.



(a) Improvement of the optimization process of the GA for solving the problem

Pseudocode
1.BEGIN
2. Initialize the population
3. Evaluate the fitness of the initial population
4. WHILE number of iterations < T_max AND not converged DO
5. Select individuals with high fitness
6. Perform a crossover operation to generate new individuals, and then perform a mutation operation
7. Evaluate the fitness of new individuals
8. Update population
9. END WHILE
10. Output the optimal solution
END

(b) Pseudocode

Figure 6: Improvement of the optimization process of the GA for solving the problem and pseudocode

In Figure 6(a), GA is used to solve the transportation path of finished cigarette products. The parameters of the GA are set as follows: the population size is 100, the crossover rate is 0.8, the mutation rate is 0.01, and the number of iterations is 200. The initial population is initialized by randomly generating the path code of the distribution point.

4 Performance analysis and application of a storage path optimization model for finished cigarette logistics distribution

After completing the optimization of the storage path for the distribution of finished cigarette logistics, the performance of the algorithm under actual operation also needs to be analyzed to check its effectiveness. To test the performance of the improved GA-A* algorithm in the storage path planning for the distribution of finished cigarette logistics, the experimental environment and basic parameters were first set up. Table 2 demonstrated the basic hardware environment settings for the experiments.

Table 2: The experimental basic environmental parameters

Parameter variables	Parameter selection
The overall implementation platform of the system	Simulink
Operating system	Windows 10
Operating environment	MATLAB R2015b
System Memory Memory	8GB
CPU main frequency	2.62Hz
Global procurement unit	RTX-2070
Central Processing Unit	Intel Corei7-4590
Data storage	MySQL data bank
Data regression analysis platform	SPSS 26.0
Standard moving distance (meter)	400
Standard moving distance in overapping area(meter)	200

To ensure the ultimate fairness and reasonableness of the process of conducting the experiments, the parameters of the Improved Genetic Algorithm (IGA), the Improved A* Algorithm (IA*) and the reference [21], as well as the distance matrix for testing, were set identically to those of the research method, except for the experimental parameters specific to the research model. The Solomn dataset and the Gehring dataset were selected for the experiments and the three algorithms mentioned above were experimented with the research method and compared to record the optimal path distances. The convergence criterion of the algorithm was defined as: within 1010 consecutive generations, the objective function value of the optimal solution was improved to a minimum value of 4.2. The convergence of the different algorithms was first compared, as shown in Figure 7.

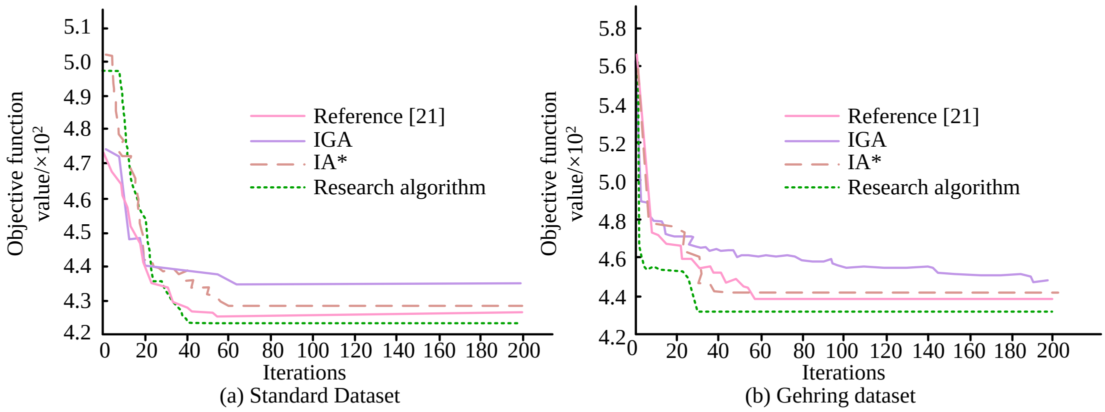


Figure 7: Comparison of convergence of the four algorithms

Figure 7(a) demonstrated the algorithm convergence results for the algorithms on the Solomn dataset. The objective function values of all the algorithms showed a decreasing trend under increasing changes in the number of iterations. Among them, IGA, IA* and reference [21] started to converge at the 62nd, 60th and 58th iteration, respectively, and appeared to have a minimum objective function value. In contrast, the research method startedto have a minimum function value at the 41st iteration, with better overall convergence accuracy. Figure 7(b) illustrated the method's results on the Gehring dataset. The IGA, IA* and the reference [21] all showed a rapid change in the objective function value at the beginning of the iteration. The change in IGA was not stable and kept fluctuating throughout the run. The research method reached a stable convergence at the 32nd iteration. It was verified that the research method could reach a stable convergence state relatively quickly throughout the run, with a fast convergence rate [22]. To extend the experiment, the study used the Solomn dataset as the main dataset and compares the AUC values of the four algorithms, corresponding to the ROC curves obtained in Figure 8.

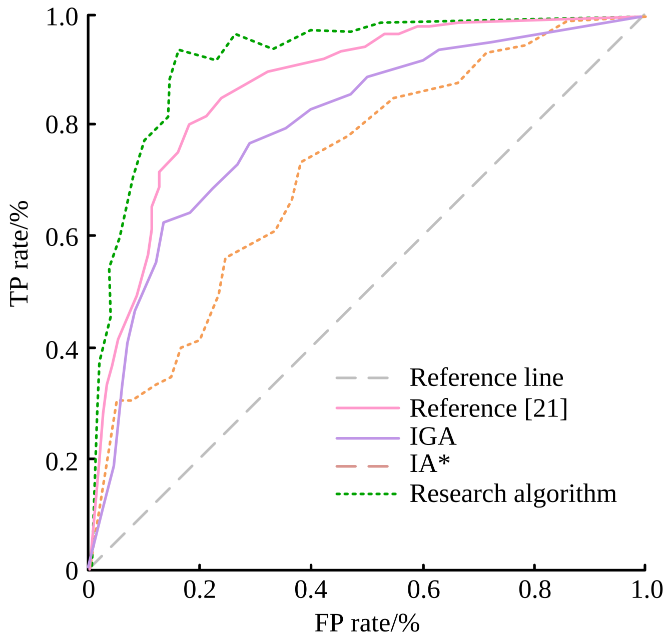


Figure 8: ROC curves for the four algorithms

In Figure 8, the area under the ROC curve for the research method was significantly larger than the other algorithms, while the AUC values for the method in the reference [21] were larger than the values for the other methods. The area under the ROC curve was calculated to be 0.985, 0.967, 0.941 and 0.875 for the research method, IGA, IA* and reference [21] respectively, and the AUC values for the research method were greater than the other methods by the significance results, which also indicated that the results obtained by the research method were more realistic. 80% and 60% of the Solomon and Gehring datasets were selected as training sets, respectively. The remaining data was used as a validation set for analyzing the path planning error results of the research method, as shown in Figure 9.

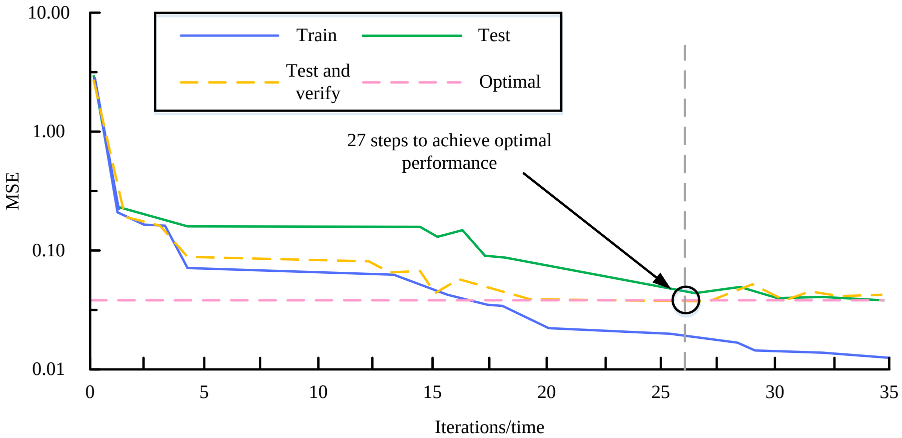
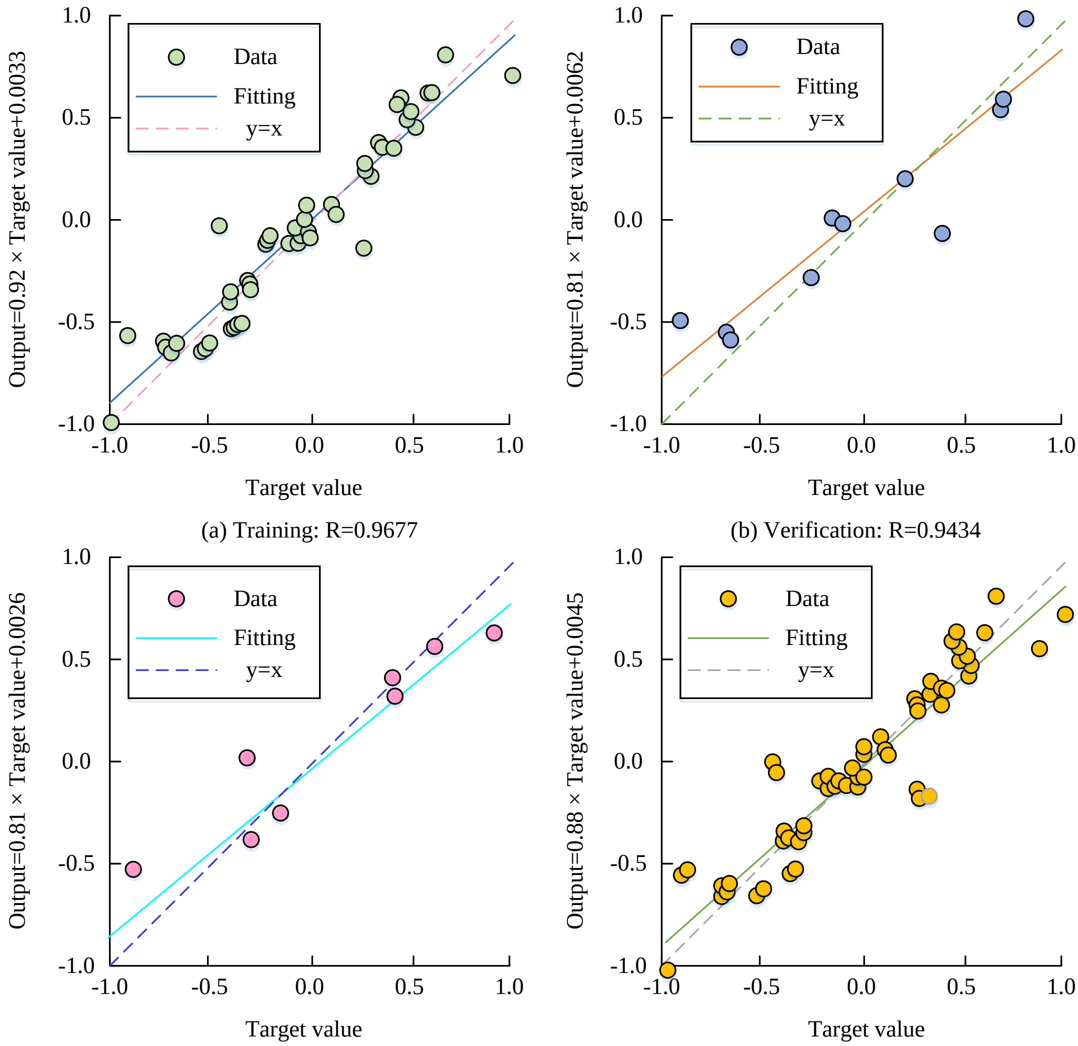


Figure 9: Error results for the study method path planning

In Figure 9, as the training steps increases, the error values under the study method gradually decreased and converged to the target value. By the 27th iteration, the model achieved an error value of 0.06 for the path planning results and the performance is optimal. This was a low value compared to the set error target and meets the system performance requirements. To further observe the path planning results of the system model under the research method, the experiments were statistically and graphically plotted for the research model training, validation, testing and all data in the dataset as output versus input, see Figure 10.



(a) Training: R=0.9677

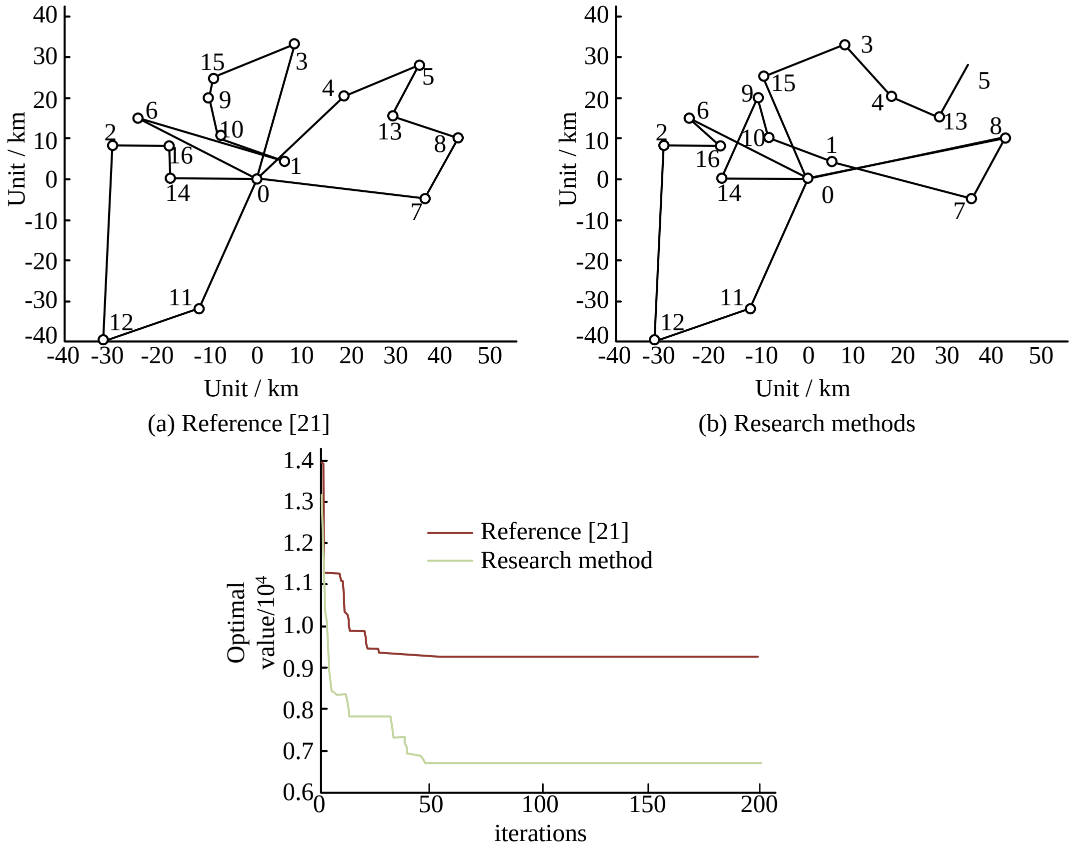
(b) Verification: R=0.9434

(c) Test: R=0.9605

(d) All: R=0.9625

Figure 10: Study model training, validation, testing and output and input correlation coefficients for all data

In Figure 10, the correlation coefficient R values of 0.9677, 0.9434, 0.9605, and 0.9625 for the training, validation, test, and all data of the research model were all greater than 0.9. This indicated that the research model was not overfitted during the run. In addition, the correlation coefficient value obtained from the training of the research model was larger than that obtained from the validation dataset, at 0.0243, indicating that the research model was also free from overfitting on the different datasets. In summary, the research model was able to map the relationship between input and output information, and thus made more accurate predictions for the route planning of finished cigarette logistics and distribution. To demonstrate the usefulness of the research method to find the optimal route planning for the distribution of finished cigarette logistics, the research method was validated by applying it to the data example Solomn dataset, and the results are shown in Figure 11.



(a) Reference [21]

(b) Research methods

(c) Comparison of Two Algorithms for Optimization

Figure 11: Path finding results of the study method on the Solomn dataset

Figure 11 showed the logistic optimization paths for the Solomn dataset obtained by solving using the study's improved GA-A* method. Figure 11(a) and Figure 11(b) showed the optimal path planning plots corresponding to the minimum total cost under the reference [21] and the research method respectively. Figure 11(c) then showedthe comparison of the optimal search curves for the two methods. From Figure 11(c), in the optimal path planning state, the research method started to stabilize and reach the optimal state when iterating for about 47 iterations under the same conditions, whereas the method in the reference [21] needed to iterate up to 69 times before it started to proceed to the steady state.In addition, based on the above results, the number of delivery points was increased to 20, and the number of iterations of different algorithms to find the optimal path was compared to analyze the scalability. Meanwhile, the percentage of system memory occupied by different algorithms was compared to analyze the computational complexity of different algorithms. The results were shown in Table 3.

Table 3: Comparison of scalability and computational complexity

Method/Algorithm	Number of iterations (scalability)	Memory usage/% (computational complexity)
Shen [7]	64	90.12
Imrane et al. [21]	67	89.17
IGA	62	84.21
IA*	58	80.30
Research algorithm	53	60.52

As shown in Table 3, compared with the other four algorithms, the proposed method had better scalability when it was iterated to the 53rd time, and the corresponding memory usage reached a stable state, with a value of only 60.52%. This showed that the proposed method could still maintain high performance when dealing with large-scale problems, and when dealing with problems of the same scale, the proposed method could obtain better solutions at a lower computational cost.

5 Discussion

The cigarette finished product logistics distribution and storage path optimization method proposed in the study based on the improved GA-A* algorithm showed significant advantages in multiple key performance indicators. In terms of convergence speed, the method proposed by the research reached a stable state when it iterated to the 41st iteration on the Solomn data set and the 32nd iteration on the Gehring data set; in addition, the method proposed by the research institute achieved path planning at the 27th iteration. The error was reduced to 0.06, and the performance reached the optimal level, which was not mentioned in the literature [21], indicating that the proposed method also had advantages in terms of accuracy. Although the logistics distribution path optimization model based on dynamic road network proposed by Scholar Shen could greatly improve the transportation efficiency of urban logistics, the cost was higher than the model proposed in the research. The model proposed by the study cost significantly less than other methods under the same path transportation conditions. The rapid convergence and high accuracy of the research method meant that it could provide logistics decision-makers with timely and reliable path planning solutions in practical applications. In addition, the high efficiency of this method also meant that in a dynamically changing logistics environment, it could quickly respond to demand changes and update distribution routes in real time, thereby improving the adaptability of the logistics system and customer satisfaction.

6 Conclusion

To reduce the loss of cigarette products during transportation and meet customer needs, the study proposes a storage path planning method based on an improved GA-A* algorithm for the logistics distribution of finished cigarette products. The results showed that the area under the ROC curves of the research method, IGA, IA* and reference [21] were 0.985, 0.967, 0.941 and 0.875 respectively in the comparison of AUC values. In the path planning error, when the 27th iteration was performed, the research method had an error value of 0.06 for path planning, which was the optimal performance; the correlation coefficient of the research method was greater than 0.9 on all data, and no over-fitting occurs. Convergence was also compared on the Solomn dataset, with the research method, IGA, IA* and the reference [21] starting to converge at the 41st, 62nd, 60th and 58th iterations respectively. In the path finding results, the research method was able to obtain the optimal path at around the 47th iteration. The above data show that the fusion of GA-A* algorithm can help in the sub-path planning of the distribution of finished cigarette logistics and can effectively identify the optimal path, which is of great significance to the development of the tobacco transportation industry in China. However, as the research method only considers for application testing on known datasets and does not extend the scope of the study, subsequent research is yet to be conducted using multiple datasets to ensure the universality of the method.

Acknowledgement

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (No. 2011-0009454).

References

[1] F. Xiong, P. S. Gong, Z. Q. Peng, and J. F. Fan, “Optimization of urban traffic distribution path under quick response demand,” Open House International, vol. 43, no. 1, pp. 108-112, 2018. http://dx.doi.org/10.1108/OHI-01-2018-B0022

[2] R. O'Connor, L. M. Schneller, N. J. Felicione, R. Talhout, M. L. Goniewicz, L. A. David, “Evolution of tobacco products: recent history and future directions,” Tobacco Control, vol. 31, no. 2, pp. 175-182, 2022. http://dx.doi.org/10.1136/tobaccocontrol-2021-056544

[3] V. S. Tida, and S. Hsu, “Universal spam detection using transfer learning of BERT model,” ArXiv, vol. 2022, 2202.03480, 2022. http://dx.doi.org/10.48550/arXiv.2202.03480

[4] H. Chen, P. Wang, C. Liu, S. Chang, J. Pan, Y. Chen, W. Wei, and D. Juan, “Complement objective training,” ArXiv, vol. 2019, 1903.01182, 2019. http://dx.doi.org/10.48550/arXiv.1903.01182

[5] T. Ya. Shevgunov, Z. A. Vavilova, O. A. Guschina, and E. N. Efimov, “Heuristic global optimization algorithm for estimating a radio source location by a passive radar system,” TEM Journal, vol. 9, no. 2, pp. 427-433, 2020. http://dx.doi.org/10.18421/TEM92-02

[6] Q Liu, “Automated logistics management and distribution based on RFID positioning technology,” Telecommunications and Radio Engineering, vol. 79, no. 1, pp. 17-27, 2020. http://dx.doi.org/10.1615/TelecomRadEng.v79.i1.20

[7] Y Shen, “Optimization of urban logistics distribution path under dynamic traffic network,” International Core Journal of Engineering, vol. 6, no. 1, pp. 243-248, 2020. http://dx.doi.org/P20190813001-202001-201912170001-201912170001-243-248

[8] M. Jardas, T. Krljan, A. P. Hadžić, and N. Grubišić, “Distribution center logistics optimization model-city of rijeka case study,” Pomorstvo, vol. 34, no. 1, pp. 185-194, 2020. http://dx.doi.org/10.31217/p.34.1.20

[9] X. Wang, H. Li, J. Yang, C. Yang, and H. Gui, “Optimal path selection for logistics transportation based on an improved ant colony algorithm,”

Journal of Embedded Systems, vol. 13, no. 2, pp. 200-208, 2020. http://dx.doi.org/10.1504/IJES.2020.10029455
[10] Z. Liu, H. Guo, Y. Zhao, B. Hu, L. Shi, L. Lang, and B. Huang, “Research on the optimized route of cold chain logistics transportation of fresh products in context of energy-saving and emission reduction,” Mathematical Biosciences and Engineering, vol. 18, no. 2, pp. 1926-1940, 2021. http://dx.doi.org/10.3934/mbe.20210100
[11] J. Peng, Y. Ding, G. Zhang, and H. Ding, “Smoothness-oriented path optimization for robotic milling processes,” Science China Technological Sciences, vol. 63, no. 9, pp. 1751-1763, 2020. http://dx.doi.org/10.1007/s11431-019-1529-x
[12] H. Min, X. Xiong, P. Wang, and Y. Yu, “Autonomous driving path planning algorithm based on improved A* algorithm in unstructured environment,” Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 235, no. 2-3, pp. 513-526, 2021. http://dx.doi.org/10.1177/0954407020959741
[13] R. S. Beed, S. Sarkar, A. Roy, S. D. Biswas, S. Biswas, “A hybrid multi-objective carpool route optimization technique using genetic algorithm and A* algorithm,” Computer Research and Modeling, vol. 13, no. 1, pp. 67-85, 2021. http://dx.doi.org/10.48550/arXiv.2007.05781
[14] G. Yue, M. Zhang, C. Shen, and X. Guan, “Bi-directional smooth A-star algorithm for navigation planning of mobile robots,” Scientia Sinica Technologica, vol. 51, no. 4, pp. 459-468, 2021. http://dx.doi.org/10.1360/SST-2020-0186
[15] T. Meng, T. Yang, J. Huang, W. Jin, W. Zhang, Y. Jia, K. Wan, G. Xiao, D. Yang, and Z. Zhong, “Improved hybrid a-star algorithm for path planning in autonomous parking system based on multi-stage dynamic optimization,” International Journal of Automotive Technology, vol. 24, no. 2, pp. 459-468, 2023. http://dx.doi.org/10.1007/s12239-023-0038-1
[16] Z. Zhang, J. Wu, J. Dai, and C. He, “Optimal path planning with modified A-Star algorithm for stealth unmanned aerial vehicles in 3D network radar environment,” Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, vol. 236, no. 1, pp. 72-81, 2022. http://dx.doi.org/10.1177/09544100211007381
[17] X. Shen, Y. Zhang, Y. Tang, Y. Qin, N. Liu, and Z. Yi, “A study on the impact of digital tobacco logistics on tobacco supply chain performance: taking the tobacco industry in Guangxi as an example,” Industrial Management and Data Systems, vol. 122, no. 6, pp. 1416-1452, 2022. http://dx.doi.org/10.1108/IMDS-05-2021-0270
[18] W. Shan, “Enterprise logistics management strategy under supply chain management mode,” 2019 International Conference on Management, Finance and Social Sciences Research, vol. 8, pp. 429-435, 2019. http://dx.doi.org/10.25236/mfssr.2019.088
[19] S. M. Farley, J. Sisti, J. Jasek, and K. R. J. Schroth, “Flavored tobacco sales prohibition (2009) and noncigarette tobacco products in retail stores (2017), New York City,” American Journal of Public Health, vol. 110, no. 5, pp. 725-730, 2020. http://dx.doi.org/10.2105/AJPH.2019.305561
[20] Y. Li, L. Xu, and S. Zhao, “Tobacco logistics retroactive system research based on RFID technology,” Internet of Things and Cloud Computing, vol. 4, no. 4, pp. 39, 2016. http://dx.doi.org/10.11648/j.iotcc.20160404.11
[21] M. L. Imrane, A. Melingui, J. J. B. M. Ahanda, F. B. Motto, and R. Merzouki, “Artificial potential field neuro-fuzzy controller for autonomous navigation of mobile robots,” Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, vol. 235, no. 7, pp. 1179-1192, 2021. http://dx.doi.org/10.1177/0959651820974831
[22] Y. Long, Z. Zuo, Y. Su, J. Li, and H. Zhang, “An A*-based bacterial foraging optimization algorithm for global path planning of unmanned surface vehicles,” Journal of Navigation, vol. 73, no. 6, pp. 1247-1262, 2020. http://dx.doi.org/10.1017/S0373463320000247