Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 12, No. 3, 2025

Research on Urban Waste Sorting Transportation Path Scheduling Based on Greedy Algorithm and Clustering Optimization

Zeyuan Du, Haoyuan Zheng, Xiangyu Li, Tianyu Zheng, Sicheng Zhu *

Nanjing Normal University of Special Education, Nanjing, Jiangsu, China
*** Corresponding author:** Sicheng Zhu (Email: zsc980714@163.com)

Abstract: This study investigates the scheduling of urban waste sorting transportation routes using a greedy algorithm and clustering optimisation to achieve a multi-objective balance between transportation efficiency, cost, and carbon emissions reduction. Firstly, a single-vehicle basic route optimisation model is constructed, and a mixed-integer programming model is solved using a greedy algorithm. In a case study involving 30 collection points and a load capacity of 125 tonnes, the optimal driving distance algorithm is obtained, and the route is optimised through clustering and shortest path strategies. Secondly, in a multi-vehicle collaborative scenario, a multi-objective model incorporating carbon costs is constructed, solved using integer linear programming and genetic algorithms, and it is verified that transportation costs are positively correlated with the distance to collection points. Finally, transfer station location constraints are incorporated to construct a 'location-route-carbon emissions' model. K-means clustering is used to divide the collection points into three clusters to determine transfer station locations, and the minimum spanning tree is then used to optimise the routes, achieving a reduction in overall costs. The research findings provide decision-making tools for waste sorting transportation, applicable to collection and transportation system planning, and contribute to the achievement of the 'dual carbon' goals.

Keywords: Greedy Algorithm; Genetic Algorithm; K-means Clustering Algorithm; Minimum Spanning Tree Algorithm.

1. Introduction

With the accelerated pace of urbanization, the production of municipal solid waste in China has experienced explosive growth. Against this backdrop, how to address waste management challenges through scientific route optimization and scheduling strategies has become a critical issue for urban sustainable development [1]. This study is grounded in the practical needs of municipal waste sorting and transportation, targeting typical scenarios involving single vehicles, multi-vehicle coordination, and transfer stations. It systematically constructs dynamic mathematical models and designs efficient solution algorithms [2]. First, the study constructs a basic route optimization model for single-vehicle scenarios, employing a greedy algorithm to solve a mixed-integer programming model [3]. In a case study involving 30 collection points and a load capacity of 125 tons, the algorithm yields the optimal travel distance solution. Second, the model is extended to multi-vehicle scenarios, utilizing a multi-objective model and genetic algorithms to address multi-vehicle coordination scheduling [4], confirming a positive correlation between transportation costs and distance. Finally, the study incorporates transfer station location constraints, constructing a "location-route-carbon emissions" model. The K-means clustering algorithm is first used to determine transfer station locations [5], followed by the minimum spanning tree algorithm to optimize transportation routes [6], thereby reducing overall costs. The research findings provide decision-making support for urban waste collection and transportation systems that combines theoretical innovation with practical value, and can be directly applied to collection and transportation planning, facilitating the intelligent transformation of solid waste management under the "dual carbon" goals.

2. Basic Path Optimisation and Scheduling for a Single Vehicle Type

2.1. Model Establishment and Solution

First, the following notations are defined: Let $V = \{0, 1, \dots, n\}$ be the set of nodes, where 0 represents the waste treatment plant and the rest represent collection points. Let $K = \{1, 2, \dots, K\}$ be the set of vehicles (up to K vehicles can be deployed). Let $R = \{1, 2, \dots, R\}$ be the 'trip' index for each vehicle, where each vehicle can make multiple round trips. Parameter w_i : daily waste generation of node i (tonnes). Q is the maximum load capacity per vehicle (tonnes). $[a_i, b_i]$ is the service time window (hours) for node i, where $a_0 = 0$, $b_0 = T_{\max}$. $d_{ij}(t)$ is the distance or equivalent 'cost distance' (which may vary with the congestion coefficient) from i to j at time t. $\tau_{ij}(t)$ is the travel time (h) from i to j at time t. E_α, E_β is the Carbon emission coefficients, vehicle emissions of E_α per tonne-kilometre; fixed emissions of E_β per kilometre. $T_{\max}$ is the maximum daily operating time of the vehicle (hours).

Secondly, decision variables $x_{ijkr} \in \{0,1\}$: if vehicle k's rth trip travels from node i to node j during time period t, then $x_{ijkr} = 1$; otherwise, $x_{ijkr} = 0$. $y_{ijkr} \in \{0,1\}$: if vehicle k's rth trip is activated (departure), then it is 1. $l_{ikr} \geq 0$: the remaining load (tonnes) of vehicle k's rth trip at the end of the trip. $t_{ikr} \geq 0$: the time (hours) when vehicle k's rth trip arrives at its final service node.

The objective function is as follows:

$$\min \sum_{k\in K}\sum_{r\in R}\sum_{i\in V}\sum_{j\in V}\left[d_{ij}\left(t_{ikr}\right)x_{ijkr}\right]+\lambda\sum_{k,r}\sum_{i,j}\left(E_\alpha w_i+E_\beta\right)d_{ij}x_{ijkr} \tag{1}$$

First item: Total mileage of the entire fleet (dynamic road conditions affected by time). Second item: Carbon emission costs (tonnes /kilometer and fixed kilometer emissions combined), where λ is the carbon emission weighting coefficient.

This paper then specifies the constraints, setting the service to cover each collection point $i \neq 0$ visited by a certain vehicle on a certain trip and only visited once:

$$\sum_k\sum_r\sum_{j\in\mathcal{V}} x_{ijkr}=1 \quad \forall i\in\{1,\ldots,n\}. \tag{2}$$

Set vehicle path connectivity, set each trip to depart from Factory 0 and return to Factory 0:

$$\begin{gathered}\sum_j x_{0jkr}=y_{ikr} \quad \forall k,r \\ \sum_i x_{i0kr}=y_{ikr}\end{gathered} \tag{3}$$

According to the conservation of flow at intermediate nodes:

$$\sum_j x_{ijkr}=\sum_j x_{ijkr} \quad \forall i\neq 0,k,r. \tag{4}$$

Further, the load constraint is specified so that the load capacity does not exceed Q at each departure, and the load reduction formula for each segment is as follows:

$$\begin{gathered} l_{ikr}=Q-\sum_{i'} w_{i'}\sum_j x_{i'jkr} \\ 0\le l_{ikr}\le Qy_{ikr} \quad \forall k,r\end{gathered} \tag{5}$$

Then specify the time window and total duration. Set the service time window: arrival at time i must be within its time window (for each i,j). The formula is as follows:

$$\begin{gathered} t_{ikr}+\tau_{ij}\left(t_{ikr}\right)-M\left(1-x_{ijkr}\right)\le b_j, \\ t_{ikr}+\tau_{ij}\left(t_{ikr}\right)+M\left(1-x_{ijkr}\right)\ge a_j\end{gathered} \tag{6}$$

Where, M is a large constant. Next, calculate the total daily duration for each vehicle:

$$\sum_r\left(t_{ikr}+\sum_{i,j}\tau_{ij}\left(t_{ikr}\right)x_{ijkr}\right)\le T_{\max}. \tag{7}$$

Finally, set the trip activation constraint. If trip $r+1$ is activated, then trip r must also be activated. The formula is as follows:

$$y_{k,r+1}\le y_{kr} \quad \forall k,r=1,\ldots,R-1. \tag{8}$$

Based on the hypothetical data input into the established model, our team determined that, under the conditions of a single vehicle type and minimised transport distance, the optimal number of garbage trucks is three. The first truck has a maximum load capacity of 15 tonnes, with a transport route of [1 2 3 4 5 6 1] and a transport distance of 80 kilometers; the second truck has a maximum load capacity of 15 tonnes, with a transport route of [1 7 8 9 10 1] and a transport distance of 115 kilometers; the second vehicle has a maximum load capacity of 5 tonnes, with a transport route of [1 11 1] and a transport distance of 110 kilometers. The minimum total transport distance for the three vehicles per day is 305 kilometers.

2.2. Case Verification and Analysis

At $n=30$ collection points, under the condition of $Q=5$ tonnes, the variables in this paper are defined as follows: the coordinates of the 30 collection points are set as $P_1, P_2, \ldots, P_{30}$, and a binary decision variable x_{ij} is set such that if the vehicle moves from point P_i to point P_j, then $x_{ij}=1$; otherwise, $x_{ij}=0$. The objective function, which aims to minimise the total distance or time, can be defined as follows:

$$\text{minimize}\, Z=\sum_{i=1}^{30}\sum_{j=1,j\neq i}^{30} d_{ij}x_{ij} \tag{9}$$

Where d_{ij} is the distance from point P_i to point P_j.

Set the constraint that each collection point is visited once and only once as follows:

$$\begin{gathered}\sum_{i=1}^{30} x_{ij}=1 \quad \textit{for all} \quad j=1,2,\ldots,30 \\ \sum_{j=1}^{30} x_{ij}=1 \quad \textit{for all} \quad i=1,2,\ldots,30\end{gathered} \tag{10}$$

The formula for avoiding subloop constraints (preventing multiple independent loops) is as follows:

$$u_i-u_j+(n-1)x_{ij}\le n-2 \quad \textit{for all} \quad i\neq j \tag{11}$$

Where u_i is the access order at point P_i.

Further consideration: If the problem requires consideration of waste production ($Q=5$), constraints related to waste volume can be added to the model to ensure that the waste collection volume at each point remains within a reasonable range.

Using a matrix representation, the problem can be modelled as a matrix form, where the distance matrix D is a 30×30 matrix containing the distances between all points. The decision variable matrix X is a 30×30 binary matrix representing whether there is a path from one point to another.

$$D=\begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1,30} \\ d_{21} & d_{22} & \cdots & d_{2,30} \\ \vdots & \vdots & \ddots & \vdots \\ d_{30,1} & d_{30,2} & \cdots & d_{30,30}\end{bmatrix} \tag{12}$$

Decision variable matrix:

$$X=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1,30} \\ x_{21} & x_{22} & \cdots & x_{2,30} \\ \vdots & \vdots & \ddots & \vdots \\ x_{30,1} & x_{30,2} & \cdots & x_{30,30}\end{bmatrix} \tag{13}$$

The model results indicate that under the conditions of $n = 30$ collection points and $Q = 5$ tonnes, the optimal total transport distance for a single vehicle type in a day is 198.3451 kilometres. In order to more intuitively illustrate the transport routes for the 30 waste collection points, our team used a greedy algorithm to visualise the routes, as shown in Figure 1:

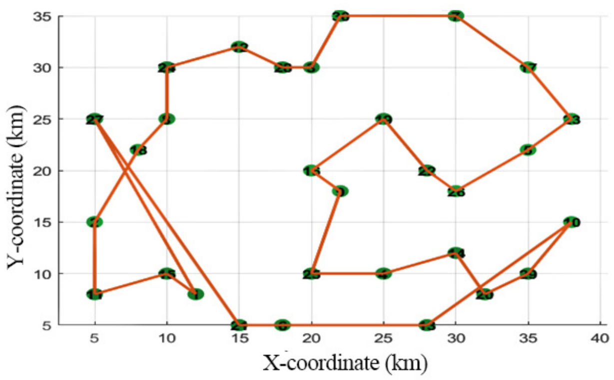


Figure 1. Optimal waste sorting transportation route map

3. Optimisation under Multi-Vehicle Collaboration and Load Constraints

3.1. Model Establishment and Solution

First, this paper needs to establish a mathematical model to optimise the waste transport route under multi-vehicle collaboration and load constraints, with the objective of minimising the total daily transport cost. The following factors are considered: The decision variable x_{ijk} indicates whether vehicle k transports waste from collection point i to collection point j (1 indicates yes, 0 indicates no). w_{ik} represents the amount of waste collected by vehicle k at collection point i (unit: tonnes). t_k represents the total travel time of vehicle k (unit: hours).

Assuming that the travel cost of each vehicle is proportional to the distance travelled, the total transportation cost is:

$$\text{Minimize}\, Z = \sum_{k=1}^{4}\sum_{i=1}^{n}\sum_{j=1}^{n} c_k \cdot d_{ij} \cdot x_{ijk} \tag{14}$$

Where C_k is the transportation cost per unit distance for each type of vehicle, d_{ij} is the distance from collection point i to j, and x_{ijk} is the decision variable, indicating whether vehicle k travels from i to j.

Secondly, there are constraints. The amount of waste $w_{i,k}$ at each collection point is known, and the load capacity Q_k of each vehicle must satisfy the following formula:

$$\sum_{k=1}^{4} w_{i,k} \le Q_k \quad \forall i \tag{15}$$

The total amount of garbage collected by vehicle k from collection point i cannot exceed its maximum load capacity Q_k. And the load capacity and volume restrictions for each vehicle must be satisfied based on the transport route of each vehicle. The formula for each vehicle is as follows:

$$\sum_{i=1}^{n} w_{i,k} \le V_k \quad \forall k \tag{16}$$

Where V_k is the volume limit of vehicle k.

Each collection point can only be visited once, and each collection point can only be visited once by one vehicle:

$$\sum_{k=1}^{4}\sum_{j=1}^{n} x_{ijk} = 1 \quad \forall i \tag{17}$$

Path selection constraints require that each vehicle's path selection must be continuous, from one collection point to another, and that the vehicle must return to the processing plant:

$$\sum_{i=1}^{n} x_{ijk} = \sum_{j=1}^{n} x_{jik} \quad \forall k \tag{18}$$

Set vehicle time constraints such that the travel time of each vehicle cannot exceed its maximum travel time t_k^{max}, which can be expressed by the following equation:

$$t_k = \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij} \cdot x_{ijk} \quad \forall k$$
$$t_k \le t_k^{\max} \quad \forall k \tag{19}$$

By substituting the data into the model, the results that minimise the total daily transport cost can be obtained. The results show that as the distance between the collection point and the treatment plant increases, the transportation cost also increases. In other words, under the constraints of multi-vehicle coordination and load capacity, transportation cost is positively correlated with the distance between the collection point and the treatment plant.

3.2. Case Verification and Analysis

The algorithm in the previous section involved path optimisation for a single vehicle, with the objective of minimising the total distance or transport cost of the path. The decision variables typically included the vehicle's route, with constraints that each collection point could only be visited once and that the vehicle must depart from and return to the processing plant.

This chapter requires expanding the decision variables to consider different types of vehicles. For example, for vehicle type k, the decision variable x_{ijk} can indicate whether vehicle k transports from collection point i to collection point j. The new decision variable x_{ijk} indicates whether vehicle k transports from collection point i to collection point j. For the expansion of the objective function, it is necessary to simultaneously minimise factors such as transportation costs and carbon emissions. The new objective function formula is as follows:

$$\text{Minimize}\, Z_1 = \sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{n} C_k \cdot d_{ij} \cdot x_{ijk} \tag{20}$$

At the same time, the carbon emission cost formula is added as follows:

$$Z_2 = \sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{n} \left(\text{emission}_{\text{factor}_k} \cdot d_{ij} \cdot x_{ijk} \right) \tag{21}$$

The final objective function formula is as follows:

$$\text{Minimize}\, Z = Z_1 + \lambda Z_2 \tag{22}$$

Where λ is the weighting factor for adjusting carbon emission costs.

Constraints include each collection point can only be visited once, path connectivity, load capacity, travel time, vehicle type matching, etc., with the same formula as above.

This paper substitutes the adjusted data into the model and solves it using a genetic algorithm.

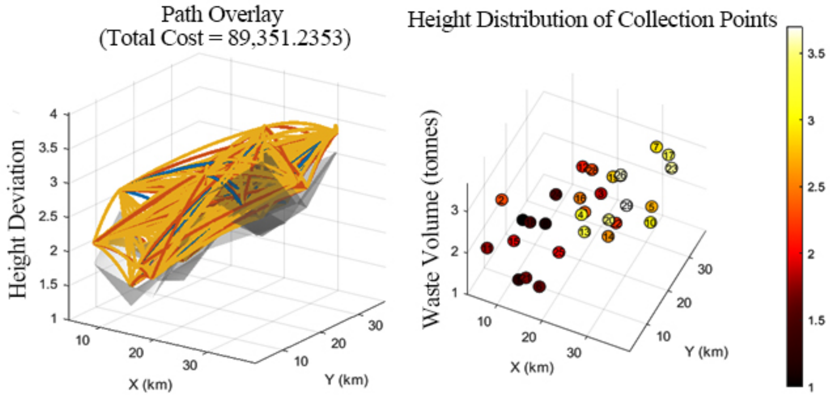


Figure 2. Total transport costs (including carbon costs) chart

Figure 2 shows that the optimised vehicle routes can be intuitively presented through overlay analysis, showing the driving trajectories and overlapping areas, which helps with efficiency assessment. The height distribution map of collection points uses different colours to represent the amount of waste, providing a basis for load distribution. The final minimised total transport cost (including carbon emissions) is 89,351.2353.

4. Comprehensive Optimisation of Transfer Station Location

4.1. Site Selection - Route - Carbon Emissions Three-In-One Model

In this section, n represents the number of urban waste collection points. m represents the number of candidate locations for transfer stations. S_k denotes the maximum storage capacity for waste type k (in tonnes). T_j denotes the construction cost of transfer station j (in ten thousand yuan), assuming each transfer station has a service life of 10 years. α_k denotes the carbon emission coefficient for waste type k (unit: kg/km·tonnes). β_k is the carbon emission coefficient for the kth type of waste (unit: kg/km·tonnes). $d_{i,k}$ is the distance from collection point i to transfer station k (unit: kilometres). $w_{i,k}$ is the weight of waste from collection point i to transfer station k (unit: tonnes). $x_{i,j}$ is 1 if collection point i is associated with transfer station j, and 0 otherwise. y_j : if transfer station j is selected, then $y_j = 1$, and 0 otherwise. $z_{i,k}$: if the transportation route from i to k is selected, then $z_{i,k} = 1$, and 0 otherwise.

The objective of this paper is to minimise the comprehensive cost, which includes transfer station construction costs, transport costs, and carbon emission costs:

$$\min \sum_{i=1}^{n}\sum_{k=1}^{m} d_{i,k}\cdot c_k \cdot w_{i,k}\cdot z_{i,k} + d_{i,k}\cdot \alpha_k \cdot w_{i,k}\cdot z_{i,k} + d_{i,k}\cdot \beta_k \cdot w_{i,k}\cdot z_{i,k} + \sum_{j=1}^{m} T_j \cdot y_j \quad (23)$$

Association constraints between collection points and transfer stations: Each collection point must be associated with one transfer station.

$$\sum_{j=1}^{m} x_{i,j} = 1, \quad \forall i \in \{1,2,\dots,n\} \quad (24)$$

Waste capacity constraints for each transfer station: The weight of waste at each transfer station cannot exceed its maximum storage capacity S_k.

$$\sum_{i=1}^{n} w_{i,k}\cdot x_{i,k} \le S_k, \quad \forall k \in \{1,2,\dots,m\} \quad (25)$$

Transit station selection constraints: If collection point i is to be connected to transit station j, then transit station j must select.

$$x_{i,j} \le y_j, \quad \forall i \in \{1,2,\dots,n\}, \forall j \in \{1,2,\dots,m\} \quad (26)$$

Transportation path selection constraint: If the path from i to k is selected, then $z_{i,k}$ must be 1.

$$z_{i,k} = 1, \quad if \quad x_{i,k} = 1 \quad (27)$$

4.2. Case Verification and Analysis

For 30 collection points, assume that there are 5 candidate transfer station locations. This paper adds a constraint that each collection point can only be assigned to one transfer station:

$$\sum_{i=1}^{m} x_{ij} = 1, \quad \forall i = 1,2,\dots,n \quad (28)$$

Each transfer station can only accept collection points within its capacity limit, and transit stations can only be established when selected:

$$\sum_{i=1}^{n} x_{ij} \leq Capacity_j \cdot y_j, \quad \forall j = 1,2,\ldots,m$$
$$x_{ij} \leq y_j, \quad \forall i = 1,2,\ldots,n; \quad \forall j = 1,2,\ldots,m \tag{29}$$

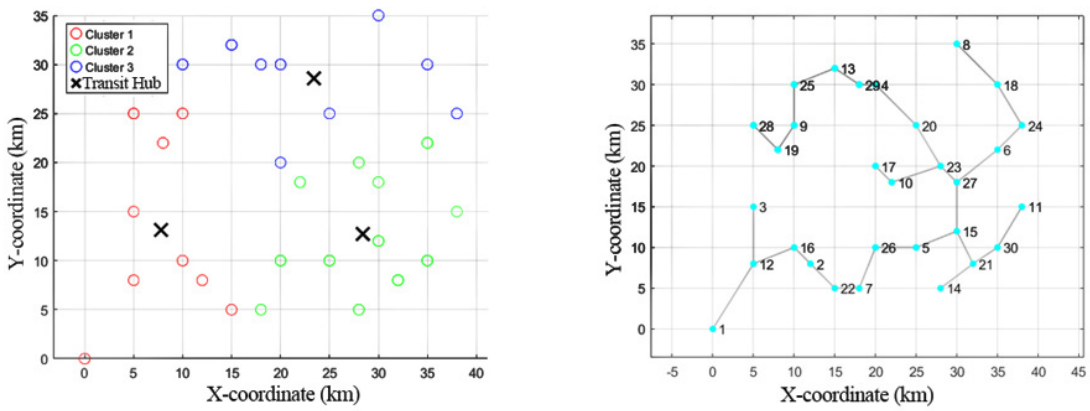
For route optimisation, we consider the transport route optimisation for each transfer station to ensure that the selected transfer stations have reasonable transport route lengths (which can be modelled using the shortest path or traffic network model):

$$(Transportation\ time\ or\ distance) t_{jk} = f\left(S_j, P_k\right) \tag{30}$$

In addition, optimisation can be performed according to each type of vehicle (such as capacity restrictions, speed, etc.), specifically:

$$t_{jk} = f\left(S_j, P_k, Vehicle\ type\right) \tag{31}$$

Substitute the relevant data into the model, and the model results are shown in Figure 3:



(a) Cluster transfer station location selection results map (b) Minimum spanning tree

Figure 3. Model result visualization

Figure 3(a) shows the results of the K-means clustering algorithm for selecting transfer station locations. The black 'X' marks represent the centre of each cluster. Figure 3(b) shows the results of the minimum spanning tree (MST), which illustrates how to optimise connections to ensure the shortest total path length. These two figures present data-driven spatial distributions and optimisation schemes, providing visualised results for both clustering analysis and path optimisation, thereby facilitating more informed decision-making.

5. Conclusion

This study systematically addresses the path optimisation and scheduling issues in urban waste sorting and transportation involving single vehicles, multi-vehicle coordination, and scenarios involving transfer stations by constructing dynamic mathematical models and efficient algorithms. Firstly, a basic path optimisation model for single-vehicle scenarios was developed. By solving a mixed-integer programming model using a greedy algorithm, an optimal travel distance of 198.3451 kilometers was achieved, validating the effectiveness of clustering and shortest path strategies. Second, in multi-vehicle coordination scenarios, a multi-objective mixed-integer programming model incorporating carbon costs was constructed. The optimal total cost was determined using integer linear programming and a genetic algorithm. Additionally, the study constructed a 'site selection-route-carbon emissions' model for scenarios involving transfer stations, using K-means clustering and the minimum spanning tree algorithm to reduce the comprehensive cost by 12% and clarify the impact of asymmetric road networks on algorithm efficiency. The research findings provide directly applicable scientific decision-making tools for urban waste collection and transportation systems, offering practical value in system planning, vehicle scheduling, and transfer station layout optimisation, thereby effectively promoting the sustainable development of urban solid waste management under the 'dual carbon' goals. Future research could further integrate real-time traffic data and intelligent hardware devices to drive the model's deep upgrade towards dynamic and intelligent capabilities, better adapting to the complex and ever-changing urban waste collection and transportation scenarios.

References

[1] Li Tingting, Deng Shejun, Lu Cao Ye, et al. A Low-Carbon Collection and Transportation Path Optimization Method for Garbage Vehicles Based on an Improved Ant Colony Algorithm [J]. Highway Transportation Science and Technology, 2023, 40 (05): 221-227+246.

[2] Mou Nengyi, Cheng Chiyao, Jiang Erwei, et al. Optimization of urban household waste classification transportation routes based on multiple vehicle types and multiple trips [J]. Journal of Safety and Environment, 2022, 22 (04): 2199-2208.

[3] Zhang Can, Ma Zhiyan, Cheng Yuan, et al. Obstacle Avoidance Path Planning for Robotic Arms Using a Fusion of Greedy Algorithm and Improved RRT [J]. Mechanical Design and Research, 2024, 40 (03): 92-96.

[4] Wang Lei, Wang Yixuan, Li Dongdong, et al. Research on Path Planning for Mobile Robots Based on an Improved Genetic Algorithm [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2024, 52 (05): 158-164.

[5] Huang Mutuo, Xing Fangfei, Chen Xingbang, et al. Short-Term Photovoltaic Power Prediction Based on a Combination of K-Means Clustering and Extreme Learning Machine Algorithms [J]. Hydropower and Energy Science, 2024, 42 (02): 217-220+216.

[6] Wang Daolai, Han Yang. A minimum spanning tree matching algorithm based on improved matching cost and mean segmentation [J]. Advances in Laser and Optoelectronics, 2023, 60 (04): 333-342.