https://doi.org/10.31449/inf.v48i14.6221 Informatica **48** (2024) 205–218

Lead Battery Reverse Logistics Center Location Model and Simulation Analysis Based on Genetic Algorithm and Greedy Algorithm

Yunqing Shi
School of Management, Fuzhou Technology and Business University, Fuzhou 350715, China
E-mail: s13338285676@126.com

Keywords: reverse logistics, center location selection, lead acid batteries, genetic algorithm, cost

Received: May 20, 2024

Lead acid batteries, as batteries with both cost and performance, are widely used in various fields such as transportation and communication. However, improper recycling can lead to increased environmental pollution. A hybrid lead acid battery reverse logistics center location model based on the genetic algorithm and greedy algorithm is proposed. Firstly, the basic mode of reverse logistics is introduced. A basic model of reverse logistics center location network for lead acid batteries is established based on relevant location principles such as non-zero constraints and cost control conditions. Then, genetic algorithm and greedy algorithm are introduced to solve and analyze the overall model. The performance of each algorithm is applied. Meanwhile, a hybrid algorithm is designed. Finally, the performance of the model is analyzed through experiments, comparing the performance of the individual genetic algorithm, greedy algorithm, and hybrid algorithm. Accuracy, recall, F1 value, and time complexity are selected as evaluation indicators. The total cost and sustainability scores of different models are compared. The experimental results showed that the accuracy of the hybrid algorithm model reached 98.82%, and the recall rate reached 97.39%. The average running time of the hybrid algorithm was 36.14% lower than that of the genetic algorithm. The average running time of the hybrid algorithm was 3.42 s. The average Gap value of the model used in the study was 51.02% lower than that of the comparison models based on dynamic adaptive particle swarm optimization algorithm, an optimized firefly algorithm, and a two-layer programming genetic algorithm-based center location model. The average total cost decreased by 39.96%. The sustainability score was 24.69% higher than the other models on average. The total construction cost of the hybrid algorithm model was lower than the other algorithms by 940000-yuan, 330000 yuan, and 850000 yuan, respectively, with an average cost reduction of 39.96%. Therefore, the proposed location selection model for lead acid battery reverse logistics centers based on the genetic-greedy hybrid algorithm can achieve low-cost and short transportation route center point calculation.

Povzetek: Študija predlaga model lokacije povratnega logističnega centra za svinčeve akumulatorske baterije, ki temelji na genetskem in požrešnem algoritmu.

1 Introduction

With the rapid development of electric bicycles, lead acid batteries have also become a hot selling product, with their total production accounting for nearly half of the global production. However, the improper recycling of waste lead acid batteries has a huge impact on the environment. However, it has not received the attention it deserves, and its organizational recovery rate is less than 30%. The recycling of waste lead acid batteries is not only beneficial for environmental reconstruction, but also greatly saves corresponding material resources and avoids waste [1]. Therefore, it is necessary to systematically organize the recycling of lead acid batteries. This requires further promotion of modern logistics, orderly connection of various nodes, and improvement of the logistics system. To address the above issues, further research has been conducted on the logistics transportation system, among which the center location strategy for logistics transportation is relatively popular and complex [2]. Due to the different distances between nodes in logistics and the uncertain occurrence time of waste lead acid batteries, it is difficult to achieve a satisfactory center location through manual calculation alone. Therefore, the study introduces heuristic algorithms for model calculations, namely algorithms inspired by natural organisms. A single algorithm often cannot achieve better optimization results, so multiple algorithms are usually used for mixed auxiliary operation [3]. Genetic algorithm (GA) is a commonly used global search algorithm with high optimization ability, but it still has certain drawbacks. Therefore, the study introduces the greedy algorithm to optimize GA. Through cost control and other constraints, the basic framework of a reverse logistics center location model is established. A hybrid algorithm is used for solution analysis to ultimately obtain the optimal solution for center location. This center location method, which solves through hybrid algorithms and layers reverse logistics with constraints such as cost control, can achieve better location results compared to traditional algorithm analysis. The study is divided into four parts. The first part introduces the current research status of reverse logistics, the second part designs a lead acid battery reverse logistics center location model using a genetic-greedy hybrid algorithm, the third part analyzes the center location model through experiments, and the fourth part summarizes the experimental results.

2 Related works

Reverse logistics is essentially an environmental protection concept that is widely used in research at present. Among them, the selection of distribution centers is very important and complex. Zhang et al. proposed a model for the location and scale of cold chain distribution centers with the goal of minimizing the total social cost of the entire logistics system. This method took into account carbon emission trading policies and the needs of logistics users. The cloud particle swarm optimization algorithm was used to determine the suitable locations of major cities. The results indicated that demand, scale, and carbon emission policies had an impact on the site selection scale, total carbon emissions, and total social costs of cold chain distribution centers [4]. Li et al. proposed a cuckoo search algorithm with balanced learning, which improved search ability by learning the beneficial behavior of two excellent individuals and exhibited good performance in solving optimization problems. The experimental results showed that the O-BLM-CS algorithm balanced resource utilization and exploration, which was competitive in both continuous and discrete optimization problems [5]. Uluta and others believed that the location of logistics center had an important impact on the cost and benefit of enterprises. Therefore, they put forward a mathematical model of comprehensive multi-objective decision-making variable weight coordination analysis based on gis. This model was combined with fuzzy strategy, simulated the location of logistics center in Sivas Province, Türkiye. The experimental results indicated that their method accurately determined the optimal location [6]. Liu et al. used two-dimensional language information to represent preference information and expert evaluation information for the location of comprehensive logistics distribution centers. They proposed improved operational rules and scoring functions. Based on two-dimensional language information, a clustering analysis method based on language similarity was proposed. Finally, a center location solution framework was constructed. The effectiveness of this method was verified through problem examples [7].

Heuristic algorithm is a type of method based on simulating natural bodies, which is constructed through experience or intuition, including ant colony algorithm and simulated annealing algorithm. The heuristic algorithm has been widely used due to its advantages of simple calculation, combined qualitative and quantitative analysis. Wang et al. designed a two-stage heuristic algorithm based on the cold chain logistics site selection for fresh agricultural products, using a multi-objective optimization model to minimize total cost and carbon emissions. Finally, a case study showed that this method effectively reduced costs and carbon emissions and promoted the sustainable development of logistics enterprises [8]. Cao et al. were concerned about the efficiency of biomass logistics systems. Therefore, they investigated the routing of biomass resources in two tiers and established a mixed integer programming model to determine the optimal biomass collection facility and vehicle center route. A mixed heuristic algorithm was introduced to address computational complexity. The effectiveness and efficiency of this method were verified through comprehensive calculation examples [9]. Yazdani et al. aimed to develop a two-stage decision model to achieve center site selection. Data envelopment analysis and rough set theory were used to determine five evaluation criteria for effective and ineffective alternatives. A combination compromise solution method was used to evaluate the performance of the effective community. Finally, sensitivity analysis was conducted to verify the robustness of the obtained results [10]. Geng et al. proposed a multi-criterion constrained site selection model to address the impact of diversion of shelters and pre stored materials on site selection. The model took into account various factors such as the needs of disaster victims and budget constraints. The model optimized the distance of shelters, allocation of shelter personnel, and pre-stored quantity of materials. The experimental results indicated that the model helped to influence the location of shelters and the allocation of disaster victims, obtaining multi-objective solutions [11].

Heuristic algorithms are often used in the establishment of center location models. Further research on hybrid optimization algorithms is needed. Therefore, a hybrid location selection model based on genetic-greedy algorithm is proposed. This model introduces the reverse logistics for the recycling of lead acid batteries, optimizes the initial model, and ultimately achieves global optimization.

Table 1: Summary of related works

Authors	Key methods	Results	Limitations
Zhang et al. [4]	This study proposes a location and scale model for a cold chain distribution center	Demand, scale, and carbon emission policies have an impact on site selection scale, carbon emissions, and costs.	This model requires collecting a large amount of data. The accuracy and completeness of the data directly affect the output results of the model.
Uluta et al. [6]	This study proposes a mathematical model for comprehensive multi-objective decision-making variable weight coordination analysis based on GIS, combined with fuzzy strategies.	This method can accurately determine the optimal position.	The introduction of the balanced learning mechanism improves the search capability and increases the complexity of the algorithm.
Liu and Li. [7]	Based on two-dimensional language information, this study proposes a clustering analysis method based on language similarity, and finally constructs a center location solution framework.	The effectiveness of this method is verified through problem examples.	Two-dimensional linguistic information can represent preference information and expert evaluation information. However, more dimensional information is needed to describe the various influences.
Li et al. [5]	The algorithm improves search by learning the beneficial behavior of two good individuals.	The O-BLM-CS algorithm balances resource utilization and exploration, which is competitive in both continuous and discrete optimization problems.	Although the two-stage heuristic algorithm can improve the solution efficiency to some extent, its computational complexity may be higher compared to simple optimization algorithms.
Cao et al. [9]	This study establishes a mixed integer programming model to determine the optimal biomass collection facility and vehicle center route. This study introduces a hybrid heuristic algorithm to address computational complexity.	The effectiveness and efficiency of this method are verified through comprehensive calculation examples.	The performance of the algorithm may be affected by the initial solution.
Yazdani et al. [10]	In this study, a two-stage decision-making model is developed for centre site selection.	The robustness of the results obtained from the proposed method is verified through sensitivity analysis.	Although sensitivity analyses are useful in identifying key factors that influence the outcome of a decision, their results may be limited by the factors selected and the magnitude of the changes.
Geng et al. [11]	In this study, a multi-criterion constrained site selection model is proposed, which considers multiple factors and is optimized.	The model helps to influence the location of shelters and the allocation of disaster victims, and obtain multi-objective solutions.	There may be conflicts between multiple target objectives, increasing the complexity of the problem.
Wang et al. [8]	This study designs a two-stage heuristic algorithm using a dual objective optimization model with the goal of minimizing total cost and carbon emissions.	Case analysis shows that this method can effectively reduce costs and carbon emissions, and promote the sustainable development of logistics enterprises.	When dealing with large-scale problems, longer computation times and more computational resources may be required.

3 Hybrid algorithm application in the location network of lead battery reverse logistics centers

Lead acid batteries have excellent performance. However, improper recycling can lead to serious environmental pollution, which requires reverse logistics operations to achieve the recycling of lead acid batteries. The study establishes a center location model through cost constraints and optimized the model through GA and the greedy algorithm, ultimately achieving the calculation of the optimal center location solution.

3.1 Reverse logistics mode and construction of basic center location network for lead acid batteries

Reverse logistics refers to the process of returning consumer goods to their production location after sale, including the logistics of remanufactured goods and waste materials. The former refers to defective products, returned goods, and processing scraps, while the latter refers to the flow and recycling of different types of products with declining economic value. The overall workflow of reverse logistics is shown in Figure 1.

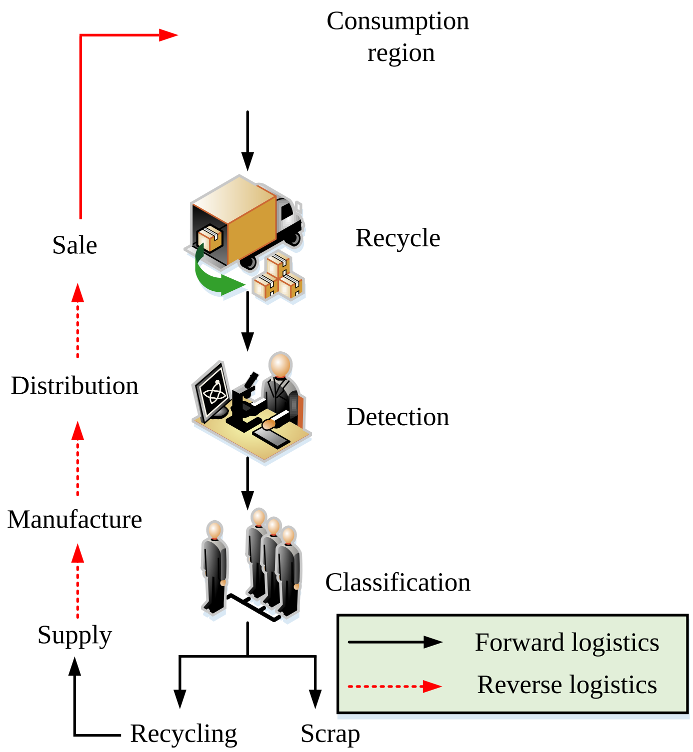


Figure 1: Overall workflow of reverse logistics

In Figure 1, the purpose of reverse logistics is to enhance economic or environmental benefits and reverse the return of remanufactured goods from downstream supply chains to upstream. Compared with the general forward logistics, the stability of reverse flow is worse. It not only operates slowly but also easily leads to a decline in value, and information such as product integrity, appearance time, and location varies among different products, which further exacerbates the difficulty of implementing reverse logistics [12]. According to the classification of waste products, reverse logistics networks can be divided into direct utilization, recycling mode, and commercial return mode networks. The reverse logistics for lead acid batteries aims to reduce environmental burden through recycling and reuse, including four layers of consumption, recycling, processing, and application, as shown in Figure 2.

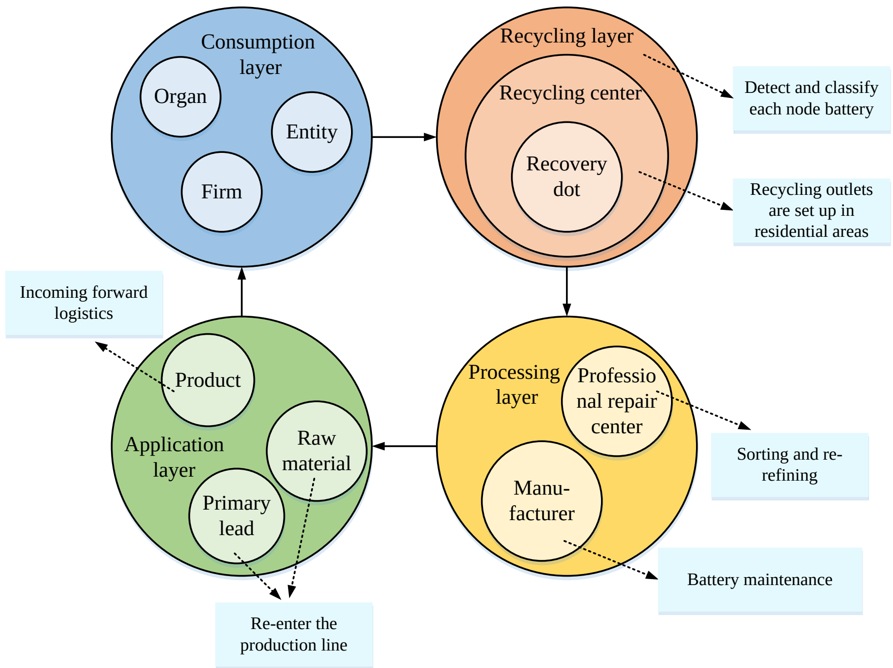


Figure 2: Reverse logistics layer of lead acid battery

In Figure 2, the consumption layer is the source of recyclable lead acid batteries and the key to the recycling process. The recycling center is the transportation point for waste lead acid batteries. The processing layer is the maintenance center, and its basic principle is harmless treatment. The application layer is to introduce updated products into the forward logistics line. The key to the entire reverse logistics network is the location of the center, which needs to simultaneously meet the convenience of logistics nodes such as processing and recycling centers [13]. Considering that logistics nodes are mainly influenced by economic, social, and environmental factors, the location should be located in a location with convenient transportation, while also considering their labor resources and environmental planning. The center of these points is to minimize economic costs as much as possible. Therefore, the center of the objective function of the center location model is to minimize costs, as shown in formula (1).

$$\min Z = \sum_{i=1}^{n} c_i \tag{1}$$

In formula (1), c_i represents the cost of each component. $\sum_{i=1}^{n} c_i$ represents the total cost. Of course, the purpose of selecting a center location is not only for geographical location, but also for reasonable consideration of node transportation volume. The entire reverse logistics network for lead acid batteries can be divided into three categories: direct utilization, recycle, and return of goods, as shown in Figure 3.

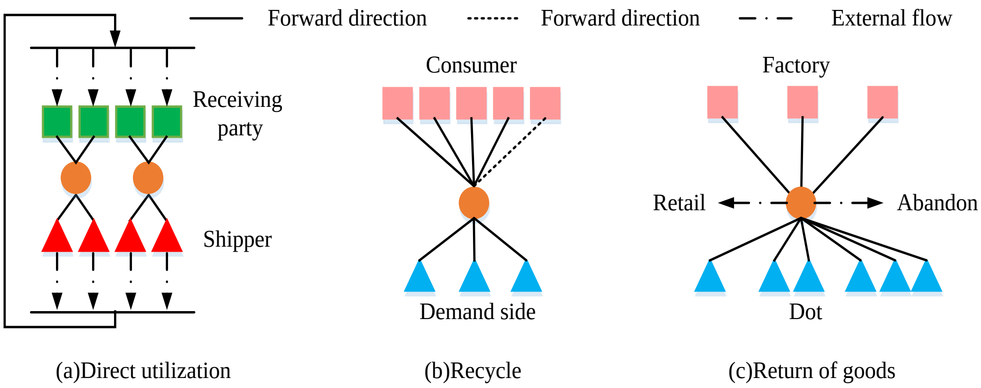


Figure 3: Reverse logistics network model for lead acid batteries

In Figure 3, the network types are divided into direct utilization, recycle, and return of goods reverse logistics structures. The first type includes recyclable materials such as product packaging and transportation containers, which can be directly used. Because these recyclable materials can be reused with simple or no processing, they can be recycled into the reverse logistics system after being used through forward logistics. The second type of recycle reverse logistics network recycles low value recycled products such as plastics, paper, and steel to achieve resource protection and utilization. The third type of return of goods reverse logistics network, with the rise of B2C e-commerce models, which has gradually become an indispensable part. The specific number of recycling outlets is shown in formula (2).

$$0 < \sum_{i \in I} y_i < Z_i \tag{2}$$

In formula (2), y_i is a 0-1 variable, representing whether a recycling outlet has been established in a certain city. If established, it is 1, and vice versa, it is 0, Z_i is the upper limit of the number of recycling outlets. The specific quantity of recycling centers and processing centers is shown in formula (2).

$$\begin{cases} 0 < \sum_{j \in J} y_j < Z_j \\ 0 < \sum_{h \in H} y_h < Z_h \end{cases} \tag{3}$$

In formula (3), y_j / y_h represent the 0-1 variables of the recycling center and processing center. Z_j / Z_h represent the upper limits of the construction quantity of the recycling center and processing center. $I / J / H$ represent the set of corresponding node locations. The transportation volume of goods needs to comply with the law of conservation of flow, as shown in formula (4).

$$\sum_{i \in I} \sum_{j \in J} f_{ij} = \sum_{j \in J} \sum_{h \in H} f_{jh} \tag{4}$$

In formula (4), f_{ij} is the amount of discarded lead batteries transported from the recycling network to the recycling center. f_{jh} is the amount of discarded lead batteries sent from the recycling center to the processing center. Secondly, the production volume of battery maintenance should not exceed the amount of lead batteries recovered at the recycling point, as shown in formula (5) [14].

$$\sum_{i \in I} \sum_{j \in J} f_{ij} \geq \sum_{i \in I} g_i \tag{5}$$

In formula (5), g_i is the upper limit of the recycling capacity of the recycling center. The product capacity and node construction cost of each network node also need to meet certain constraints, as shown in formula (6).

$$\begin{cases} f_{jh}, f_{jh} \geq 0 \\ F_i, F_j, F_h \in (0,1) \end{cases} \tag{6}$$

Formula (6) represents the non-negative constraint of recycling capacity and the 0-1 constraint of construction cost, respectively. Among them, $F_i / F_j / F_h$ represent the fixed facility construction costs of the corresponding nodes, respectively. To ensure the minimization of center location costs in reverse logistics networks, the model divides them into three categories: first, the fixed and unchanging costs in reverse logistics facilities, as shown in formula (7) [15].

$$C_f = \sum_{i \in I} F_i \cdot y_i + \sum_{j \in J} F_j \cdot y_j + \sum_{h \in H} F_h \cdot y_h \tag{7}$$

In formula (7), C_f is the fixed and unchanging cost in the reverse logistics facility of lead acid batteries. The second type is the logistics cost generated between node transportation, as shown in formula (8).

$$C_t = \sum_{i \in I} \sum_{j \in J} c_{ij} \cdot f_{ij} + \sum_{j \in J} \sum_{h \in H} c_{jh} \cdot f_{jh} \tag{8}$$

In formula (8), c_{ij} is the transportation cost from the recycling point to the recycling center. c_{jh} is the transportation cost from the recycling center to the processing center. Finally, the operating costs of each node facility are shown in formula (9).

$$C_w = \sum_{j \in j} w_j \cdot y_j + \sum_{h \in h} w_h \cdot y_h \tag{9}$$

In formula (9), w_j / w_h represent the operating costs of the recycling center and the processing center, respectively.

3.2 Solution of lead battery reverse logistics network model based on ga and greedy algorithm

The reverse logistics network model of lead acid batteries needs to be solved through corresponding algorithms, so GA and greedy algorithms are introduced to achieve it. GA is a global random search algorithm based on the principles of biological evolution. Each iteration selects the optimal individual from the candidate solutions and combines them through genetic operators to generate a new candidate solution group, continuously cycling until the algorithm finally converges. This is different from the principle of traditional algorithms that only use one initial value for optimal solution iteration. The GA uses a string set search strategy to greatly avoid the phenomenon of a single extreme value. Individual selection is carried out through fitness values. The requirements for searching spatial data and auxiliary data are not high, and its applicability is naturally larger [16]. The greedy algorithm abandons global optima and instead searches for local optima. Its continuous iteration generates two types of sets, namely the optimal candidate set and the cluster that includes the candidates that were not selected. The hybrid algorithm that combines the two compensates for each other's shortcomings, as shown in Figure 4.

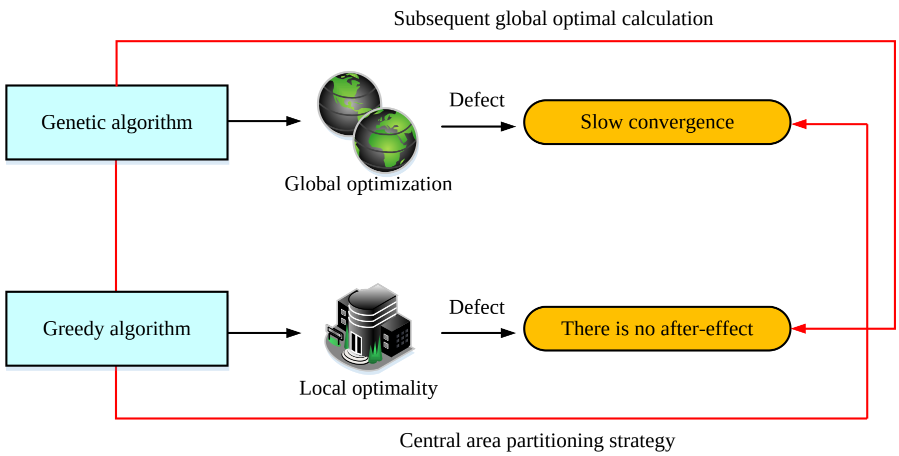


Figure 4: Optimization strategy of hybrid algorithm

From Figure 4, GA obtains the global optimal solution, which requires a large number of iterative calculations and is accompanied by cross-mutation operations. As a result, a huge amount of computation is generated, increasing the running time of the algorithm. The greedy algorithm can calculate the optimal recycling center area location strategy and transmit it to GA, ultimately improving the convergence speed of the model. However, greedy algorithms only make decisions on the current optimal solution and do not consider future solutions, which can result in greedy algorithms having no aftereffect and falling into local optima. GAs can precisely solve this defect, and the two complement each other to achieve algorithm optimization [17]. The running process of the overall hybrid algorithm is shown in Figure 5.

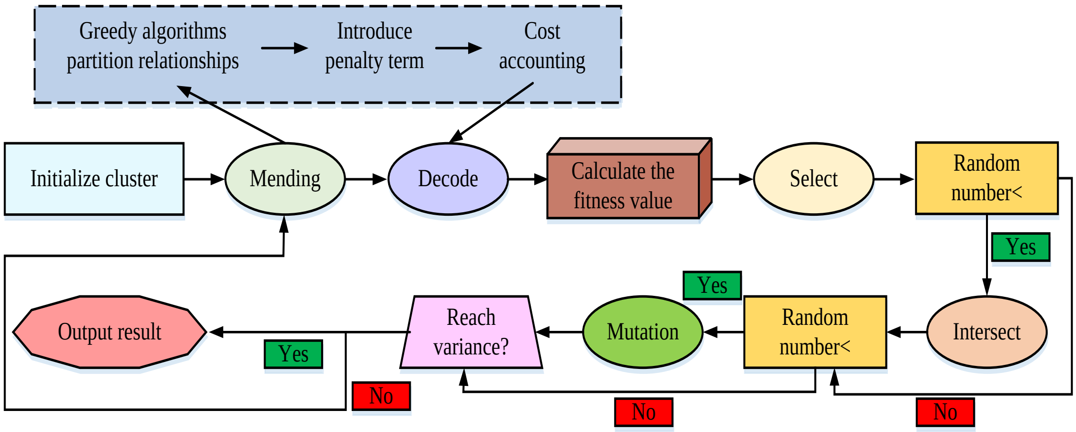


Figure 5

Figure 5: Operation flow of the hybrid genetic-greedy algorithm

The research sets the iteration number to 500, the population size to 220, the crossover probability to 0.92, and the mutation probability to 0.1. Before initializing the cluster, chromosome coding is required, and the basic principles are integrity, health, and simplicity. Firstly, it is necessary to ensure that all data points or candidate sets in the initial problem space can be encoded accordingly. That is, the points in the encoding space need to correspond one-to-one with the points in the initial problem space. Common chromosome encoding strategies include binary encoding, integer encoding, and symbolic encoding [18]. The study selects a binary encoding form, with the encoding symbol set represented as $I = \{0,1\}$. The coding formula should include three nodes: the recycling center, recycling node, and processing center. The composition of chromosomes is shown in Figure 6.

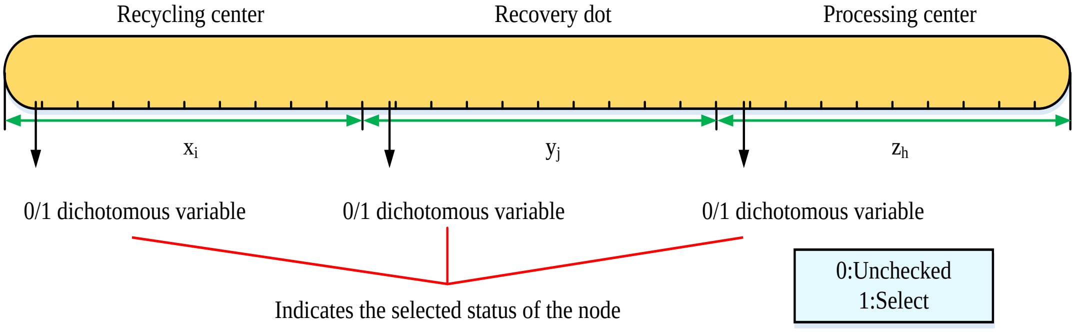


Figure 6: Chromosome coding structure

Each chromosome is composed of three fragments, xi/yj/zh, representing the selection status data of the recycling center, recycling node, and processing center. The first digit of each fragment is a 0/1 binary variable used to represent the selected status of the node. If the value is 1, the corresponding node is selected. Conversely, if the value is 0, the corresponding node is not selected. Then, cluster initialization is performed, with two initialization strategies: random and utilizing prior knowledge, depending on the specific model situation. The size of the initial cluster is related to the efficiency of the algorithm. If it is too large, it will increase the complexity of the algorithm, while if it is too small, it will lead to a decrease in the efficiency of the algorithm [19]. To optimize the diversity of the initial cluster, it is established according to randomization, as shown in formula (10).

$$\begin{cases} pop_{p,g} = (p_{p,g}); & n = 0,1,...,N; \\ p_{n,g} = (p_{m,n,g}); & m = 0,1,...,I \times J \times H; \end{cases} \quad (10)$$

In formula (10), n/m are the population number and individual number, g is the number of iterations, $pop_{p,g}$ is the g-th generation population, $p_{n,g}$ is the n-th population in the population, and $p_{m,n,g}$ is the m-th individual in the population. The random initialization function is shown in formula (11).

$$P_{m,n,0} = f(rand(1,1)) \quad (11)$$

In formula (11), $P_{m,n,0}$ represents the initial population without iteration, and $rand(1,1)$ represents any uniformly distributed values in the interval (1,1). The specific representation of this function is shown in formula (12).

$$f(rand(1,1)) = \begin{cases} 0, & (rand(1,1)) < 0 \\ 1, & (rand(1,1)) \geq 0 \end{cases} \quad (12)$$

The main function of the greedy algorithm is to calculate the minimum economic cost under the current iteration, which is divided into node minimization cost and optimization route. Taking the recycling route s_{ij} from the recycling outlet to the recycling center as an example, the calculation is shown in formula (13).

$$s_{ij} = \begin{cases} 1, & if \quad T_{ij} = \min(T_{i \times j}) \quad j \in J \\ 0, & other \end{cases} \quad (13)$$

In formula (13), T_{ij} is the time taken for transportation from the recycling point to the recycling center, as shown in formula (14).

$$T_{ij} = y_i F_{ij} c_{ij} \quad (14)$$

In formula (14), F_{ij} represents the fixed cost between two nodes. The recycling route requires freight calculation for the restriction scheme and recycling interval. Finally, greedy algorithms are used to minimize operating costs. Afterwards, it is necessary to calculate the individual fitness value, which represents the genetic performance of the individual. GA is used to calculate the fitness value. The larger the fitness value, the stronger the individual's survival ability, inheriting this gene to the next generation. Generally, the objective function is chosen as the fitness calculation function. The main purpose of selecting a reverse logistics center location is to minimize the total cost of construction work, which requires consideration of both constraint conditions and objective functions. Therefore, the study selects the reciprocal of the total cost and calculates the fitness of individuals. The record of any individual being selected is shown in formula (15) [20].

$$P = \frac{1}{C_f + C_t + C_w} \quad (15)$$

After calculating the crossover mutation operator, the algorithm can be terminated if the conditions are met. The setting of termination conditions takes into account the random search nature of GAs. If there are no restrictions, the algorithm will not stop operations. The termination conditions can be set based on a threshold of iteration times or by using a fitness value range.

4 Network simulation experiment analysis of lead battery reverse logistics center location

Simulation analysis experiments were conducted to verify the reliability of the lead battery reverse logistics center location model. Firstly, the optimization performance of the genetic-greedy hybrid algorithm was verified to ensure its effective application. Subsequently, simulation was conducted on the reverse logistics center location model to ensure that it could successfully complete the selection of the optimal center address.

4.1 Performance verification of optimization model based on the genetic-greedy hybrid algorithm

The study first conducted experimental analysis on the hybrid algorithm for solving center location and compared the GA and greedy algorithm with the hybrid algorithm. The experimental environment and parameter settings for each algorithm are shown in Table 2.

Table 2: Experimental environment and parameter settings

Name	Settings
Operating system	Windows 10
Running memory	4GB
Processor	Intel (R) Core (TM) i5-3337U CPU @ 1.80 GHz.
Population size N	100
Maximum iteration	150
Crossover probability Pc	0.7
Variation probability Pm	0.08

The study first selected the recycling situation of lead batteries in a certain city as the dataset and simulated the genetic-greedy hybrid algorithm. The fitness function values of the algorithm and the precision recall (PR) curve results are shown in Figure 7.

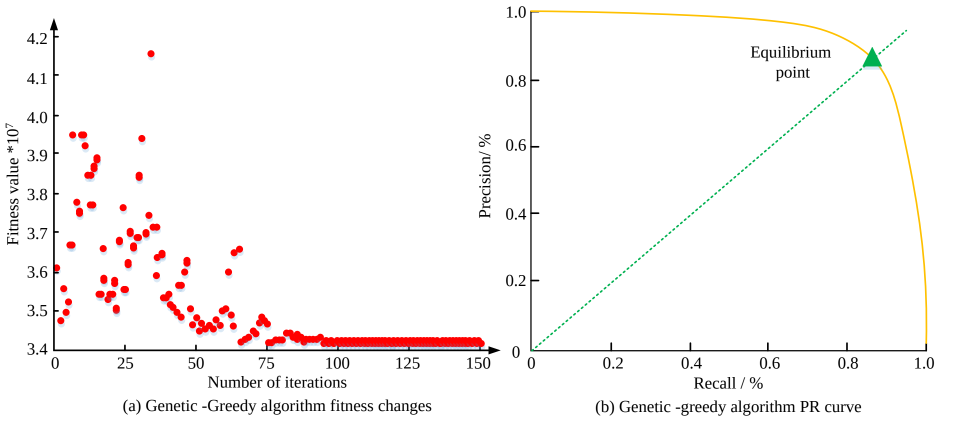
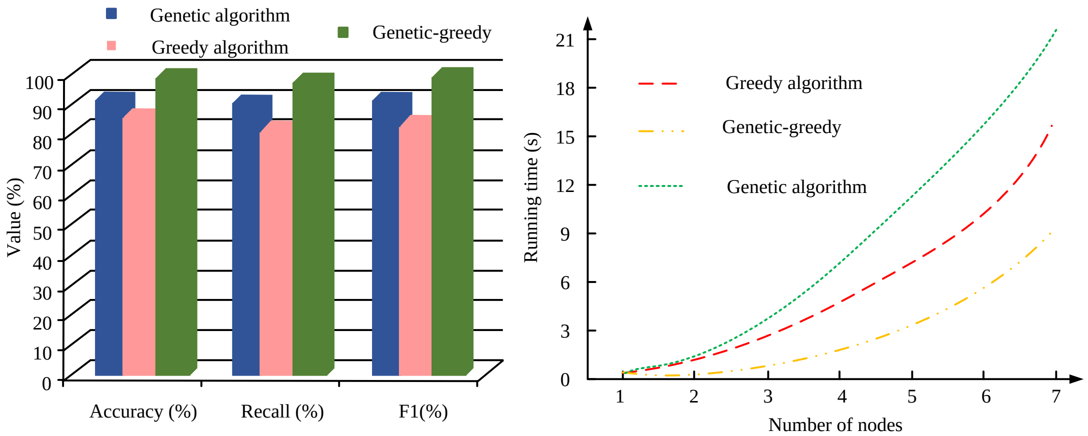


Figure 7: Simulation results of the hybrid genetic-greedy algorithm

Figure 7 (a) shows the scatter plot of the fitness function value changes of the genetic-greedy hybrid algorithm in optimizing the location selection of lead acid battery recycling centers. From Figure 7 (a), the genetic-greedy hybrid algorithm converged near 100 iterations. Before that, the distribution of fitness values was relatively dispersed. The final convergence result was 3.414*107, indicating that the minimum cost obtained by the model through center location optimization was 3.41*107 yuan. Figure 7 (b) shows the PR curve of the genetic-greedy hybrid algorithm for the classification of lead acid battery recycling, showing the relationship between precision and recall. Accuracy and recall represent the precision and recall of the model, respectively. The two could not obtain the optimal value at the same time, and the optimal solution needed to be obtained at the equilibrium point. The green triangle symbol in the figure represents the equilibrium point of the two. At this point, the precision of the genetic-greedy hybrid algorithm model reached 98.82%, and the recall reached 97.39%. In statistical analysis, the research method was compared with other algorithms (such as individual GA, individual greedy algorithm, etc.) to verify the advantages of the genetic-greedy hybrid algorithm in terms of convergence and optimal solution quality. Through multiple experiments and calculations of the average number of convergence iterations and the average cost of the optimal solution, the genetic-greedy hybrid algorithm had a faster convergence speed and a lower cost of finding the optimal solution. In addition, confidence intervals were used to evaluate the stability and reliability of the results. By calculating the confidence interval of multiple experimental results, the genetic-greedy hybrid algorithm had significant advantages in optimizing the location of lead acid battery recycling centers. The study further compared the performance of the individual GA, greedy algorithm, and the hybrid algorithm. Accuracy, recall, F1 value, and time complexity were selected as evaluation indicators. The experimental results are shown in Figure 8.



(a) Comparison of query rates of different algorithms
(b) Comparison of time complexity of each algorithm

Figure 8: Comparison of performance indexes between independent algorithms and the hybrid algorithm

From Figure 8 (a), the greedy algorithm performed the worst in terms of accuracy, recall, and F1 value, with 85.47%, 81.62%, and 83.59%, respectively. This is because the algorithm was prone to falling into local optima. Although GA improved, its various indicators were below 95%, with 91.24%, 90.75%, and 90.98%, respectively. The accuracy of the genetic-greedy hybrid algorithm reached 98.82%, which was higher than the greedy algorithm and GA by 13.35% and 7.59%, respectively. The recall of the hybrid algorithm was 97.39%, which was higher than the greedy algorithm and GA by 15.77% and 6.64%, respectively. The F1 value of the hybrid algorithm reached 98.09%, which was higher than the greedy algorithm and GA by 14.5% and 7.11%, respectively. Figure 8 (b) shows the time complexity of each algorithm, with GA, greedy algorithm, and hybrid algorithm taking 20.47s, 11.38s, and 7.93s, respectively. The proposed genetic-greedy hybrid algorithm was 61.26% and 30.32% lower than the other two algorithms, respectively. In summary, the hybrid algorithm significantly improved in various indicators compared to the initial algorithms.

4.2 The practical application of genetic-greedy hybrid algorithm in the location model of lead battery reverse logistics center

To verify the practical application effect of the hybrid algorithm in reverse logistics center location selection, five different cases were selected in the study, with different distribution points and coordinates. To simplify model operations and improve efficiency, the transportation costs and rates of the five cases were adjusted to a value of 1, and center location simulation experiments were conducted on them. The visualization results of each case are shown in Figure 9.

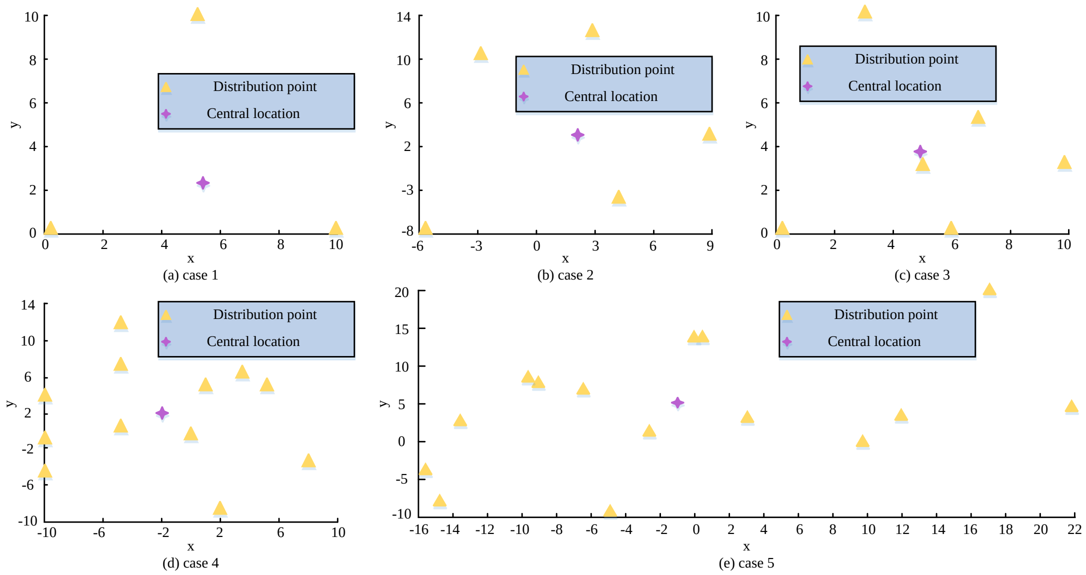


Figure 9: Visual diagram of distribution point and center location for each case

From Figure 9, the number of nodes in Cases 1-5 was 3, 5, 6, 12, and 16, respectively. Among them, the center location coordinate of Case 1 was (5.01, 2.89), Case 2 was (2.59, 3.05), Case 3 was (4.89, 3.93), Case 4 was (-2.01, 2.04), and Case 5 was (-3.82, 6.02). According to the visualization image, the center point position in each case was roughly the same as the distance from the other nodes. In cases where there were more nodes, it was generally distributed in the locations with more nodes, making the transportation distance of most nodes closer while also minimizing the transportation distance of slightly distant nodes. This was in line with the principle of the lowest cost for center location construction. Therefore, the proposed hybrid lead acid battery reverse logistics center location model based on the GA and greedy algorithm was effective and reliable. Similarly, the GA was tested in Cases 1-5, and the total transportation cost and running time of the two algorithms were compared, as shown in Figure 10.

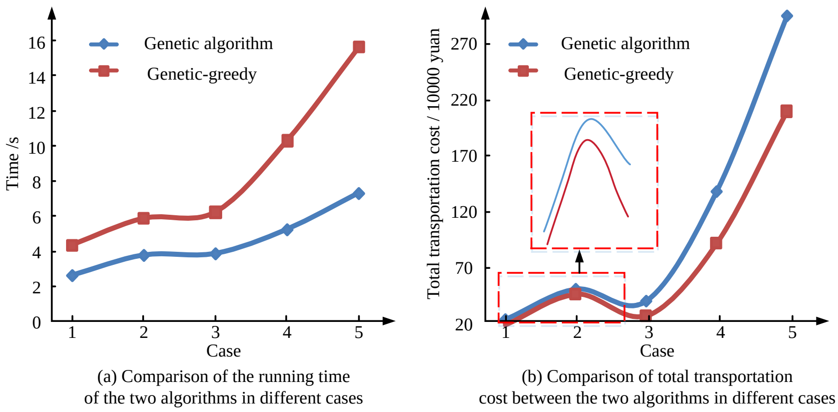


Figure 10: Comparison of location selection results between the GA and hybrid algorithm

From Figure 10 (a), as the number of nodes increased, the difference in runtime between the independent GA and the hybrid algorithm model became larger. In the first three cases, the average runtime of the hybrid algorithm was 36.14% lower than that of the GA, while the average runtime of the hybrid algorithm was 3.42s. When the number of nodes exceeded 10, the gap between the two algorithms gradually increased. In Case 4, the runtime of the hybrid algorithm was 5.24s, which was lower than that of the GA 49.27%. In Case 5, the hybrid algorithm took 7.32s, which was 53.29% lower than the GA. In Figure 10 (b), in the comparison of transportation costs among the first three cases, the GA averaged 74600 yuan higher than the hybrid algorithm. In Case 4, the total transportation costs of the two algorithms were 1.3694 million yuan and 907900 yuan, respectively, with a decrease of 33.70% for the hybrid algorithm. In Case 5, the total transportation costs of the two algorithms were 2.9372 million yuan and 2.0809 million yuan, respectively, with a decrease of 41.15% for the hybrid algorithm. The research further introduced a center location model based on the Dynamic Adaptive

Particle Swarm Optimization (DAPSO) algorithm, an optimized firefly (OF) algorithm, and a two-layer programming GA (TPG) based center location model. Comparative experiments were conducted with the hybrid algorithm model. The error measurement indicators including Gap value, sustainability, and total construction cost were selected to evaluate the accuracy of the algorithm. The experimental results are shown in Table 3.

Table 3: Comparative analysis of the performance of each center location model

Model	Index		
	Gap	Sustainability	Total construction cost (10000 yuan)
DAPSO	1.96	7.24	196
OF	3.44	6.83	135
TPG	2.98	7.56	187
Genetic-greedy	1.37	8.99	102

In Table 3, the Gap value was an evaluation of the accuracy of model site selection, and the smaller the value, the higher the accuracy of the model. The hybrid algorithm location model had an average Gap value lower than the other models by 51.02%. Sustainability represents the spatial sustainability of the center's location, which is closely related to the human resources and development level of the location. Therefore, the sustainability of the location is a relatively comprehensive evaluation indicator. The higher the score, the better the sustainability of the model. The sustainability score of the hybrid algorithm location model was on average 24.69% higher than that of other models. The total construction cost of the hybrid algorithm model was lower than the other algorithms by 94000 yuan, 330000 yuan, and 850000 yuan, respectively, with an average cost reduction of 39.96%. In summary, the center location model based on the genetic-greedy hybrid algorithm had good performance in multiple indicators.

5 Discussion

In the research of reverse logistics center location problem, a variety of algorithmic models for optimal location decisions have emerged. The hybrid model based on GA and greedy algorithm significantly outperformed GA and greedy algorithm alone, reflecting its unique advantages in several aspects. The greedy algorithm performed the worst in terms of accuracy, recall, and F1 value, which were 85.47%, 81.62%, and 83.59%, respectively. This is because the algorithm is prone to fall into local optimum. Although GA was improved, all of its indexes were lower than 95%, which were 91.24%, 90.75%, and 90.98%, respectively. The genetic-greedy hybrid algorithm achieved an accuracy of 98.82%, which was 13.35% and 7.59% higher than the greedy algorithm and GA, respectively. In previous studies, genetic and greedy algorithms have been widely used in the logistics centre location problem. GA also has problems such as slow convergence speed and easy to fall into local optimum [21]. In contrast, the recall of the hybrid algorithm was 97.39%, which was 15.77% and 6.64% higher than that of the individual greedy algorithm and GA, respectively. The F1 value of the hybrid algorithm was 98.09%, which was 14.5% and 7.11% higher than the greedy algorithm and GA, respectively. The time taken by the GA, greedy algorithm, and hybrid algorithm was 20.47s, 11.38s, and 7.93s, respectively. The proposed genetic-greedy hybrid algorithm reduced 61.26% and 30.32% than the other two algorithms, respectively. Overall, the hybrid algorithm improved significantly in all the indicators compared to the initial algorithm.

The greedy algorithm is popular with its high computational efficiency and simple implementation. However, it is often difficult to find the global optimal solution because the greedy algorithm only considers the current optimal choice. In the transport cost comparison of the cases, GA was on average $74,600 higher than the hybrid algorithm. In the two cases, the total transport costs of the two algorithms were $1,369,400 and $907,900 as well as $2,937,200 and $2,080,900, respectively. The hybrid algorithm reduced the total transport costs by 33.70% and 41.15%, respectively. Compared with previous clustering analysis methods based on two-dimensional language information which addresses selection optimization and has higher efficiency [7], the genetic-greedy hybrid algorithm combines the advantages of both algorithms. Through the global search capability of GA, the hybrid algorithm can find potential optimal solutions in a wider search space. The local optimization capability of the greedy algorithm can fine-tune these potential solutions to find more accurate solutions. This combination makes the hybrid algorithm outperform the genetic and greedy algorithms alone in terms of accuracy, recall, and F1 value. It is superior to the two-stage heuristic algorithm in reference [8] and the multi-criteria constrained location model in reference [11], which can select addresses faster and more accurately. In addition, the hybrid algorithm shows high efficiency in terms of time complexity. With the fast screening of the greedy algorithm and the parallel processing capability of GA, the hybrid algorithm can complete the computation and find a high-quality solution in a short period of time. This high efficiency makes hybrid algorithms more competitive in practical applications.

The reason for the difference is that, in terms of algorithm design, the hybrid algorithm combines the advantages of GA and the greedy algorithm to achieve the combination of global search and local optimization. In terms of experimental setup, suitable evaluation indexes and comparison algorithms are selected and fully verified through experiments [22-23]. The genetic-greedy hybrid algorithm shows unique novelty in the problem of reverse logistics center location. This combination approach has rarely been reported in previous studies. Secondly, the algorithm shows excellent performance in several aspects. Finally, the study also demonstrates the effectiveness and superiority of the hybrid algorithm in the reverse logistics centre location problem through sufficient experimental validation and comparative analysis.

6 Conclusion

A reverse logistics center location model based on genetic-greedy hybrid algorithm is proposed to address environmental pollution caused by improper recycling of lead batteries. The model is constructed based on various constraints. Then, a hybrid algorithm is introduced for simulation solution. To verify the reliability of the location selection model, a training simulation experiment is conducted on the hybrid algorithm. The experimental results showed that the hybrid algorithm began to converge at nearly 100 iterations, resulting in a minimum cost of 3.414*107 yuan. In comparison with independent algorithms, the accuracy of the hybrid algorithm was generally higher than that of the greedy algorithm and GA by 10.47%, while the recall was on average higher than the other two algorithms by 11.20%. Furthermore, actual condition simulations were conducted on the model. In all five cases, the model was able to calculate the optimal solution that met the location principle, ensuring cost minimization. In the first three cases, the average running time of the hybrid algorithm was 36.14% lower than that of the GA. As the number of nodes increased, the difference between the two increased to about 50%. The transportation cost of the hybrid algorithm was reduced by an average of 38.82% compared to the GA. Finally, three center location models, DAPSO, OF, and TPG, were introduced and compared with the research model. The results showed that the Gap value of the hybrid model was 51.02% lower on average than the other models, and the sustainability score was 24.69% higher on average than the other models. The total construction cost was reduced by 39.96% on average. Therefore, the hybrid algorithm model has better center location performance. In real life, this model uses a hybrid algorithm combining the GA and greedy algorithm to solve the reverse logistics center location model for the recycling of waste lead acid batteries, making the results more accurate and efficient. However, the assumption of center location in the study is static and does not take into account the occurrence of dynamic periodic changes. The recycling systems in various cities have not been involved. Establishing a complete reverse logistics recycling system for lead acid batteries can be refined to each city. Considering the location selection of lead acid battery reverse logistics in multi-cycle situations and the changes in location results over time, further research should be conducted in this area in the future.

7 Acknowledgements

The research is supported by 2021 Fujian Provincial Social Science Fund Project: “Research on the Coordinated Development of Regional Logistics and Regional Economy in Fujian Province under the Background of High Quality Development” (Project No.: FJ2021X018) and 2021 Fujian Provincial Education and Research Project for Middle and Young Teachers (Science and Technology) “Research on Optimization of Agricultural Product Supply Chain Information Based on Blockchain Technology” (Project No.: JAT210628).

Conflict of interest statement

There’s no conflict of interest.

References

[1] A. Garai, and B. Sarkar, “Economically independent reverse logistics of customer-centric closed-loop supply chain for herbal medicines and biofuel,” Journal of Cleaner Production, vol. 334, no. 1, pp. 129977-130000, 2022. https://doi.org/10.1016/j.jclepro.2021.129977

[2] R. Rostamzadeh, A. Esmaeili, H. Sivileviius, H. Bodaghi, K. Nobard, “A fuzzy decision-making approach for evaluation and selection of third party reverse logistics provider using fuzzy ARAS,” Transport, vol. 35, no. 6, pp. 635-657, 2021. https://doi.org/10.3846/transport.2020.14226

[3] S. Sharma, S. R. Singh, and M. Kumar, “A reverse logistics inventory model with multiple production and remanufacturing batches under fuzzy environment,” RAIRO-Operations Research, vol. 55, no. 2, pp. 571-588, 2021. https://doi.org/10.1051/ro/2021021

[4] S. Zhang, N. Chen, N. She, and K. Li, “Location optimization of a competitive distribution center for urban cold chain logistics in terms of low-carbon emissions,” Computers & Industrial Engineering,

vol. 154, no. 1, pp. 107120-107133, 2021. https://doi.org/10.1016/j.cie.2021.107120

[5] J. Li, H. Lei, and G. G. Wang, "Solving logistics distribution center location with improved cuckoo search algorithm," International Journal of Computational Intelligence Systems, vol. 14, no. 1, 2020. https://doi.org/10.2991/ijcis.d.201216.002

[6] A. Uluta, Can Bülent Karaku, A. Topal, "Location selection for logistics center with fuzzy SWARA and CoCoSo methods," Journal of Intelligent and Fuzzy Systems, vol. 38, no. 1, pp. 191400-191417, 2020. https://doi.org/10.3233/JIFS-191400

[7] P. Liu, and Y. Li, "Multiattribute decision method for comprehensive logistics distribution center location selection based on 2-dimensional linguistic information," Information Sciences, vol. 538, no. 1, pp. 209-244, 2020. https://doi.org/10.1016/j.ins.2020.05.131

[8] H. Wang, H. Ran, and X. Dang, "Location optimization of fresh agricultural products cold chain distribution Ccenter under carbon emission constraints," Sustainability, vol. 14, no. 11, pp. 4693-4709, 2022. https://doi.org/10.3390/su14116726

[9] X. G. J. Cao, "A two-echelon location-routing problem for biomass logistics systems," Biosystems Engineering, vol. 202, no. 1, pp. 106-118, 2021. https://doi.org/10.1016/j.biosystemseng.2020.12.007

[10] M. Yazdani, P. Chatterjee, D. Pamucar, and S. Chakraborty, "Development of an integrated decision-making model for location selection of logistics centers in the Spanish autonomous communities," Expert Systems with Applications, vol. 148, no. 1, pp. 113208-113229, 2020. https://doi.org/10.1016/j.eswa.2020.113208

[11] S. Geng, H. Hou, and S. Zhang, "Multi-criteria location model of emergency shelters in humanitarian logistics," Sustainability, vol. 12, no. 5, pp. 1759-1780, 2020. https://doi.org/10.3390/su12051759

[12] O. Lingaitien, A. Burinskien, and V. Davidaviien, "Case Study of Municipal Waste and Its Reliance on Reverse Logistics in European Countries," Sustainability, vol. 14, no. 3, pp. 1809-1809, 2022. https://doi.org/10.3390/su14031809

[13] D. Prajapati, S. Pratap, M. Zhang, Lakshay, G. Q. Huang, "Sustainable forward-reverse logistics for multi-product delivery and pickup in B2C E-commerce towards the circular economy," International Journal of Production Economics, vol. 253, no. 3, pp. 108606-108606, 2022. https://doi.org/10.1016/j.ijpe.2022.108606

[14] M. Barma, and U. M. "Modibbo, Multiobjective mathematical optimization model for municipal solid waste management with economic analysis of reuse/recycling recovered waste materials," Journal of Computational and Cognitive Engineering, vol. 1, no. 3, pp. 122-137, 2022. https://doi.org/10.47852/bonviewJCCE149145

[15] N. Koshta, S. Patra, and S. P. Singh, "Sharing economic responsibility: Assessing end user's willingness to support E-waste reverse logistics for circular economy," Journal of Cleaner Production, (1): 130057-130068, 2022. https://doi.org/10.1016/j.jclepro.2021.130057

[16] E. Hasanlou, N. Ali Shams, H. Izadan, "Instrumental shade sorting of coloured fabrics using genetic algorithm and particle swarm optimisation," Coloration Technology, vol. 139, no. 4, pp. 454-463, 2023. https://doi.org/10.1111/cote.12663

[17] J. Cao, H. Yan, W. Li, D. Li, and Y. Wang, "Optimization of stator ventilation structure of high-speed railway traction motor based on the genetic algorithm," IET electric power applications, vol. 17, no. 3, pp. 281-292, 2023. https://doi.org/10.1049/elp2.12263

[18] Z. H. Ahmed, "Solving the traveling salesman problem using greedy sequential constructive crossover in a genetic algorithm," IJCSNS, vol. 20, no. 2, pp. 99-105, 2020. https://doi.org/10.1109/ICCISci.2019.8716483

[19] H. X. Qin, Y. Y. Han, Y. P. Liu, J. Q Li, , Q. K. Pan, and X. Han, "A collaborative iterative greedy algorithm for the scheduling of distributed heterogeneous hybrid flow shop with blocking constraints," Expert Systems with Application, vol. 201, no. 9, pp. 117256- 117271, 2022. https://doi.org/10.1016/j.eswa.2022.117256

[20] M. Karimi-Mamaghan, M. Mohammadi, B. Pasdeloup, P. Meyer, "Learning to select operators in meta-heuristics: An integration of Q-learning into the iterated greedy algorithm for the permutation flowshop scheduling problem," European Journal of Operational Research, vol. 304, no. 3, pp. 1296-1330, 2023. https://doi.org/10.1016/j.ejor.2022.03.054

[21] G. Tasoglu, and M. A. Ilgin, "A simulation-based genetic algorithm approach for the simultaneous consideration of reverse logistics network design and disassembly line balancing with sequencing," Computers & Industrial Engineering, vol. 187, no. 1, pp. 1-14, 2024. https://doi.org/10.1016/j.cie.2023.109794

[22] H. Zhao, and A. Sharma, "Logistics distribution route optimization based on improved particle swarm optimization," Informatica, vol. 47, no. 2, pp. 243-251, 2023. https://doi.org/10.31449/inf.v47i2.4011

[23] N. Alkan, and C. Kahraman, "Prioritization of supply chain digital transformation strategies using multi-expert fermatean fuzzy analytic hierarchy process," Informatica, vol. 34, No. 1, pp. 1-33, 2023. https://doi.org/10.15388/22-INFOR493