International Journal of Mechatronics and Applied Mechanics
2025, Issue 20, Vol. 1

THE APPLICATION OF BIG DATA TECHNOLOGY COMBINED WITH IMPROVED GENETIC SIMULATED ANNEALING ALGORITHM IN THE CONSTRUCTION OF INTELLIGENT DELIVERY MODELS

Yu Zhou[1] and Pei Yan[1]*
[1]School of Humanities and Management, Xi'an Traffic Engineering Institute, Xi'an, 710300, China

Abstract - With the increasing demand for takeout, traditional path planning cannot achieve ideal results in the face of complex and ever-changing delivery environments. This study proposes an intelligent food delivery path planning model that combines K-means clustering algorithm and improved genetic simulated annealing algorithm to solve the problem of path planning. Firstly, it uses K-means to perform clustering analysis on the location of delivery customers. Then, based on this analysis, an improved genetic simulated annealing algorithm is used to plan the delivery path and construct a path planning model. The results indicated that the average order clustering value of the delivery model was 501 orders, and compared to the comparison method, the order clustering effect was more efficient. The accuracy and recall of the delivery model in customer clustering data analysis were 90.87% and 92.18%, respectively. In the path planning of the intelligent food delivery model, the ground truth refers to the shortest path, and the F1 value and area under the precision-recall curve of the proposed model in the data processing process are 83% and 82%, respectively. This indicates that the constructed model can effectively reduce the distance and time traveled by delivery personnel, improve delivery efficiency, and reduce delivery costs. The aim of this study is to improve the efficiency of food delivery and provide strong data support for the intelligence of food delivery.

Keywords: K-means; Genetic algorithm; Simulated annealing algorithm; Delivery of food delivery; Big data technology.

1. Introduction

Against the backdrop of the rapid expansion of online shopping and fast-paced lifestyle, takeout has become a daily choice for many people. With the rapid growth of food delivery business, food delivery has also become an important component of urban life. With the rapid growth of order volume and the increasing diversification of consumer demand, how to efficiently and timely complete the delivery of takeout has become a focus of attention both inside and outside the industry [1-2]. In food delivery, there is a direct correlation between path planning and delivery efficiency. Although traditional path planning methods can to some extent solve the optimization problem of distribution paths [3]. However, in the face of complex and ever-changing urban transportation environments and scattered customers, the overall path planning results cannot meet the increasing demand for takeout delivery [4]. Therefore, this study constructs an intelligent path planning model on the basis of existing Intelligent Takeaway Delivery (ITD) to handle path planning in ITD. K-means Clustering Algorithm (K-means), as an unsupervised learning clustering algorithm, has the characteristics of simplicity and efficiency, and has been widely used in data processing and pattern recognition [5-6]. Through the K-means, this study can cluster delivery orders based on factors such as geographical location and order volume, achieving reasonable division of delivery areas and providing a foundation for subsequent path planning. Simple clustering cannot completely solve path planning problems. Therefore, this study constructs an improved GA-SA algorithm based on clustering using Genetic Algorithm (GA) and Simulated Annealing (SA), and applied this algorithm to the path planning research of ITD models. This study innovatively combines K-means with GA-SA, ensuring the global search ability of the algorithm while improving its robustness and stability based on clustering analysis. It aims to provide new path planning methods for

https://doi.org/10.17683/ijomam/issue20.7

ITD and improve delivery efficiency.

Part 1 of the study analyzes and summarizes the delivery and logistics path planning for delivery. Part 2 first uses K-means to perform clustering analysis on the location of delivery customers, and then based on this analysis, GA-SA algorithm is applied to ITD to effectively plan the delivery order and path according to the different locations of customers. Part 3 is to set validation metrics and use simulation experiments to verify the performance of ITD model path planning. Part 2 summarizes the experimental results, highlighting the strengths and weaknesses of the research.

2. Related Work

With the rapid development of intelligent delivery, intelligent delivery has also been widely applied in food delivery. At present, the order volume of food delivery business is growing rapidly, and consumer demand is also becoming increasingly diversified. How to efficiently and accurately cluster customer locations and provide reliable delivery routes for food delivery has become a focus of attention both inside and outside the industry. Therefore, many scholars have conducted extensive research on distribution path planning. Lang K et al. applied cloud computing systems to predict food delivery orders in order to improve the accuracy of prediction during the delivery process. They combined artificial neural networks to propose a cloud computing resource scheduling method based on food delivery order volume prediction. This method could significantly improve the accuracy of order prediction and contribute to the rational utilization of resources [7]. Punhani R et al. used K-means to effectively analyze customer information data to prevent e-commerce customers from experiencing product substitution purchases. By identifying the characteristic attributes of different customers, identifying their demand points, and formulating corresponding sales strategies, this method could effectively complete customer clustering analysis and improve the company's sales performance [8]. To enhance the evaluation of potential customer value by enterprises, scholars such as Fang C have applied K-means to cluster analysis of the customer value system. This study was able to develop corresponding relationship maintenance strategies based on the value of different customer classifications, significantly improving the clustering of different customer values and enhancing the experiential capabilities of enterprises [9].

Mesa J P et al. used a greedy random adaptive search process to collect historical data in order to improve the path planning ability in the actual distribution process, and based on this, carried out route planning. This method achieved a similarity score of 0.0328 in the challenge race, which can effectively solve the route planning problem of the last mile delivery sequence [10]. Li XB et al. proposed a logistics distribution path calculation method that combines Ant Colony Optimization (ACO) and GA to improve the optimization of multi-point logistics distribution paths in online shopping. It significantly reduced the total transportation cost of logistics distribution and saved delivery time by integrating two algorithms [11]. Shang Y et al. proposed a dynamic step-by-step path planning method based on Fast Exploring Random Tree algorithm to address the obstacle avoidance capability in path planning. This algorithm could significantly reduce the distribution density of obstacles and dynamically adjust the step size to quickly generate initial paths and reduce planning time [12]. Anndrei A R D et al. investigated the performance of GA algorithm in vehicle path planning for food delivery companies based on literature review methods. The survey results have demonstrated the positive role of GA algorithm in food delivery [13]. Ait Saadi et al. outlined five types of drone path planning methods, including metaheuristics, machine learning, and hybrid algorithms, and proposed future research directions for drone path planning, which will help promote the development of path planning technology [14]. Yu Z et al. proposed a hybrid particle swarm optimization algorithm for optimal path planning of unmanned aerial vehicles. The results indicate that the proposed method can quickly search for high-quality paths and has good robustness [15]. From the above research, it can be seen that intelligent optimization algorithms such as GA have been widely applied in the field of path planning. However, traditional GA algorithms are prone to getting stuck in local optima and have a long search process. Therefore, the study further introduces the SA algorithm to improve the GA algorithm, in order to enhance its global search capability and convergence speed.

In summary, the combination of K-means and improved GA-SA in this study plays a crucial role in the path planning of ITD models. These methods not only enhance the clustering analysis ability of customer location, but also increase the stability of ITD model path planning. Therefore, this study utilizes K-means and improved GA-SA to effectively plan the path of the ITD model, aiming to improve the timeliness of delivery and enhance the quality of delivery services.

3. Construction of a Model for Delivery Path Planning

With digitization and intelligence, delivery, as an important component of urban services, the efficiency and accuracy of its path planning directly affect delivery costs, customer satisfaction, and enterprise competitiveness. This study combines K-means and improved GA-SA, aiming to optimize the path delivery process and improve delivery efficiency for the ITD model by clustering customer locations and planning delivery paths.

3.1 Cluster Analysis of Delivery Customers based on K-means

In the ITD model, how to improve the delivery efficiency of takeout is the basis for building the model. During the delivery process by delivery personnel, customers may be distributed in different areas and there may be multiple orders delivered simultaneously. In actual delivery, the timeliness of delivery will also be affected by different time periods and weather conditions [16-17]. To ensure timely delivery of food, the delivery platform will conduct real-time clustering analysis on the distribution of customers based on actual situations, and improve actual delivery efficiency based on the optimization of delivery goals for orders. The platform usually takes the shortest time required to complete order delivery, the shortest distance for order delivery, and the order delivery timeout rate as the goals for optimizing delivery, which is also the basis for customer clustering analysis. The shortest distance for order delivery can be calculated using equation (1).

$$\min g_1(\Omega) = \frac{\sum_{j=1}^{m} \sum_{k1,k2 \in seqj} l_{k1,k2}}{n} \tag{1}$$

In equation (1), Ω represents the set of all delivery tasks. m represents the number of delivery personnel. $k1$ and $k2$ respectively represent the number of orders delivered by the delivery person. n represents the quantity of takeout orders. $seqj$ represents the order in which delivery person j delivers takeout. $l_{k1,k2}$ represents the shortest distance value between two orders delivered by the delivery person. The shortest time required to complete order delivery can be calculated using equation (2).

$$\min g_2(\Omega) = \frac{\sum_{j=1}^{m} \max_{(i,x) \in seqj} T(i,x)}{n} \tag{2}$$

In equation (2), $T(i,x)$ represents the task value of order i reaching merchant x. The order delivery timeout rate is equation (3).

$$\min g_3(\Omega) = \frac{\sum_{i=1}^{n} 1(fr_{(i,c)} \geq d_i)}{n} \tag{3}$$

In equation (3), $fr_{(i,c)}$ represents the time when order i arrived at the customer's location. d_i represents the expected delivery time of order i to the customer. Through the calculation of the shortest delivery time, shortest delivery distance, and delivery timeout rate, it is found that these three indicators are crucial in the clustering analysis of takeaway customers. This has a significant impact on understanding and optimizing the delivery service process and improving customer satisfaction. The next step is to use K-means to perform clustering analysis on takeout customers. K-means is an iterative clustering analysis algorithm aimed at dividing data into K clusters, each represented by a centroid (mean of samples within the cluster). The algorithm selects the initial centroid, allocates samples to the nearest cluster of centroids, updates the centroids, and iterates repeatedly until the termination condition is met. Figure 1 shows the K-means clustering flowchart.

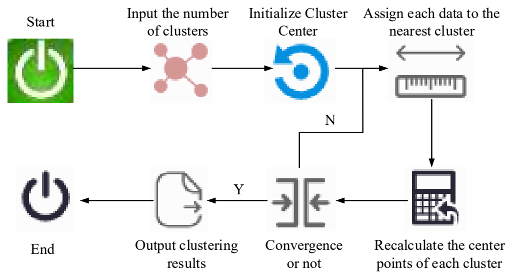


Figure 1: K-means clustering flowchart

This study first needs to define the K value. The K value used for clustering analysis of takeout delivery is defined as equation (4).

$$K = \left(\frac{T_Q}{Q * load}\right) \tag{4}$$

In equation (4), T_Q represents the predicted daily delivery volume of takeout. Q represents the daily average delivery volume of each food delivery site. $load$ represents the daily load value of the distribution station, with a load value of [0,1]. Its function is to achieve unified management, integration, and distribution of takeaway orders at the station, and better cope with fluctuations in platform business volume.

The next step is to select the initial cluster center. The density of takeout customers in cities is very high, while the density in relatively remote areas becomes very sparse [18-19]. If the position meets the criteria, it will be automatically added to the cluster center set. If it does not meet the criteria, a new random position will be selected. Finally, when the number of distance objects in the clustering set is K, the improvement iteration ends, and the Euclidean distance between the two positions is greater than the set value. After completing the optimization of the clustering center, clustering operations can be carried out, and further iterative optimization of customer clustering analysis can be carried out during the clustering process. The clustering function at this point is equation (5).

$$Tc = \sum_{j=1}^{k}\sum_{i=1}^{n}\delta_{ij} OD(p_i, Cen_j) \tag{5}$$

In equation (5), p_i represents the order position of a single customer. δ_{ij} represents whether the position of a single customer order belongs to the clustering vector. O represents the sample distance value. D represents the Euclidean distance value. Cen_j represents the point value of the customer's distance from the center of the collection. In the delivery business, the workload of delivery stations can usually be measured by the number of orders, delivery distance, delivery time, number of delivery personnel, and other relevant factors during the delivery process. This study defines the workload of food delivery stations to complete distance analysis of different customer locations. The definition of workload is equation (6).

$$W = S * q \tag{6}$$

In equation (6), S represents the coverage area of customer distribution at delivery stations. q represents the number of orders with similar distances in customer clustering. By analyzing the similarity between coverage area and customer distance, it is possible to effectively complete the clustering analysis of takeaway customer distance. Figure 2 shows the clustering analysis process of takeout customers based on K-means.

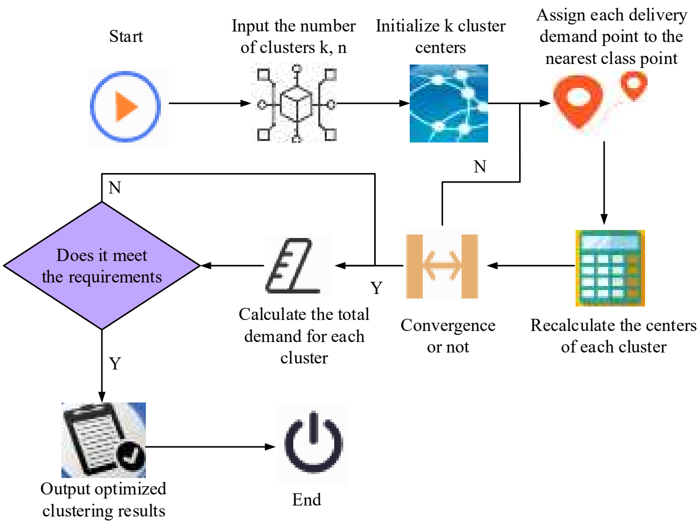


Figure 2: Flow chart of K-means based clustering analysis for takeout customers

3.2 Construction of a Delivery Model Combining Clustering and Genetic Simulated Annealing

This study is based on customer clustering results, assigning corresponding customer clusters to each delivery site and generating initial delivery paths. These paths can be random or generated based on some heuristic rules. Using these initial delivery paths as the initial population of GA-SA, each individual (i.e., each delivery path scheme) is composed of a series of customer point access sequences [20-21]. Both GA and SA algorithms have significant characteristics. GA has good global optimization ability, while SA algorithm has the ability to effectively jump out of the optimal solution. However, both have the drawbacks of premature puberty and prolonged optimization time. Therefore, this study fuses the two algorithms to construct the GA-SA algorithm, which embeds the SA algorithm into GA for the construction of a delivery model. When constructing a model using the GA-SA algorithm, it is necessary to first set the relevant parameters of GA and perform genetic operations and embedding the SA algorithm.

In this study, to avoid the occurrence of invalid solutions and reduce the search efficiency of GA, natural number coding method is used for chromosome coding.

Figure 3 shows a simulation diagram of chromosome encoding using natural number encoding method.

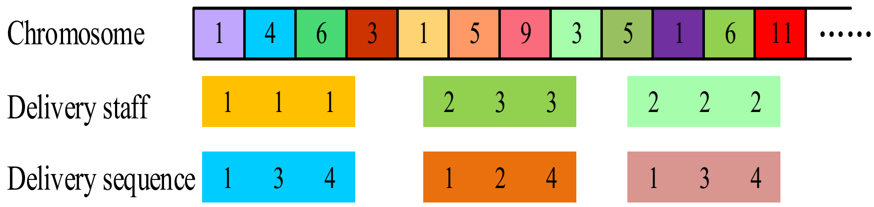


Figure 3: Chromosome coding simulation diagram of natural number coding method

In Figure 3, 1 represents the location of the food delivery merchant, and the number between two 1s represents the delivery path of the food delivery person once. This path is arranged in the order of ordering takeout. After completing chromosome coding, it is necessary to design a fitness function to evaluate the advantages and disadvantages of each delivery path scheme. The design of this function needs to consider multiple factors, such as total delivery distance, total delivery time, delivery cost, etc. The goal of designing a fitness function is to minimize the function and find a delivery path scheme that minimizes the total delivery cost (including distance, time, and other factors) as much as possible. Due to the fact that the cost of delivery and customer satisfaction cannot be directly considered in the actual delivery process, this study normalized the two. The normalized objective function at this point is equation (7).

$$Z = \partial_1 \frac{z_1}{z_1 \max} + \partial_2 \frac{z_2}{z_2 \max} \tag{7}$$

In equation (7), z_1 represents the satisfaction value of delivery customers. ∂_1 represents the weight value of customer satisfaction for takeout. ∂_2 represents the weight value of food delivery cost. z_2 represents the cost value of food delivery. The fitness function constructed on the basis of this objective function is equation (8).

$$f_r = \frac{1}{z_r} \quad r \in (1, 2, \cdots, popsize) \tag{8}$$

In equation (8), z_r represents the normalized objective function value corresponding to chromosome r. After completing the design of the fitness function, the next step is to perform selection, crossover, and mutation in genetic operations. The purpose of selection operation is to retain individuals with strong adaptability, so that the population can evolve towards the direction of optimization. This can effectively improve convergence ability. The probability of each individual inheriting to the next generation during the selection process is equation (9).

$$p(x_r) = \frac{f(x_r)}{\sum_{j=1}^{N} f(x_r)} \tag{9}$$

In equation (9), $f(x_r)$ represents the fitness function value. N represents the size of the population. At this point, the cumulative probability corresponding to each individual is equation (10).

$$q_i = \sum_{r=1}^{i} P(x_r) \tag{10}$$

After completing the selection operation, the crossover operation can be performed. The crossover operation in GA can ensure the inheritance of excellent genes to the next generation, and through the exchange of new individuals, the diversity of genes in the next generation is also improved. During the crossover process, it is necessary to consider the similarity of crossover operations, which can be represented by equation (11).

$$\Delta x = \sum_{i=1}^{l} (x_1^i - x_2^i)^2 \tag{11}$$

In equation (11), x_1^i and x_2^i represent the l-th position of the two individuals to be manipulated during the genetic process. l represents the length value of a chromosome. If the same gene exists after crossing is completed, gene elimination will occur until a new pair of offspring genes are formed. After obtaining new offspring genes, mutation operations are required. The purpose of mutation operation is to randomly change certain genes of an individual, thereby increasing the diversity of the population and preventing the algorithm from converging to local optimal solutions too early. The mutation operation conducted in this study is divided into two parts. The first is to generate mutation points in the population according to a certain mutation probability. The second is to reverse the genes at the mutation point to obtain new chromosomes. After completing genetic operations on individuals in the population, a completely new population is formed [22-23]. Due to the fact that genetic operations can lead to a decrease in population diversity and the emergence of local optima, the SA algorithm is embedded into genetic operations. Figure 4 shows the search graph of the improved GA-SA.

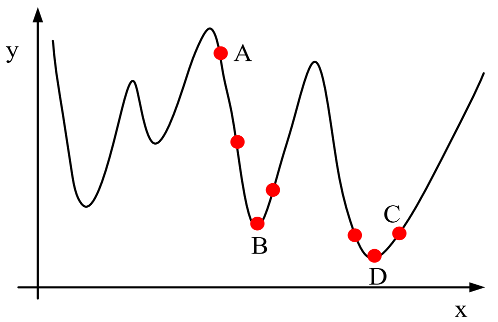


Figure 4: Search schematic diagram of improved genetic simulated annealing algorithm

In this study, the SA algorithm uses acceptance criteria to perform annealing operations on each individual in the new population obtained through genetic operations. The acceptance criteria are shown in equation (12).

$$p = \exp((f_{new} - f_{old}) / T) \tag{12}$$

In equation (12), f_{new} represents the offspring of the new population. f_{old} represents the parent. T represents the current annealing temperature. After completing the annealing operation, the GA-SA algorithm can obtain the optimal path during the delivery process, and through relevant decoding operations, the delivery path recommendation of the delivery personnel can be completed.

Figure 5 shows the process of the best delivery path model for food delivery based on customer clustering and GA-SA algorithm.

This study used accuracy, recall, F1 score, and Precision Recall (PR) curve as evaluation indicators. Among them, accuracy represents the proportion of the model that correctly finds the shortest path. The PR curve is plotted by calculating the accuracy and recall at different thresholds to evaluate the performance of the algorithm under different conditions. The recall rate represents the proportion of the algorithm correctly identifying all actual shortest paths. Accuracy measures the proportion of samples predicted by the model as the shortest path that are actually the shortest path. The calculation of recall and accuracy is shown in formula (13).

$$\begin{cases} R = \dfrac{TP}{TP + FN} \\ P = \dfrac{TP}{TP + FP} \end{cases} \tag{13}$$

In equation (13), TP represents the number of correctly predicted shortest paths. TN represents the number of correctly predicted non shortest paths. FN represents the number of incorrectly predicted non shortest paths. FP represents the number of incorrectly predicted shortest paths. The F1 value is the harmonic average of precision and recall.

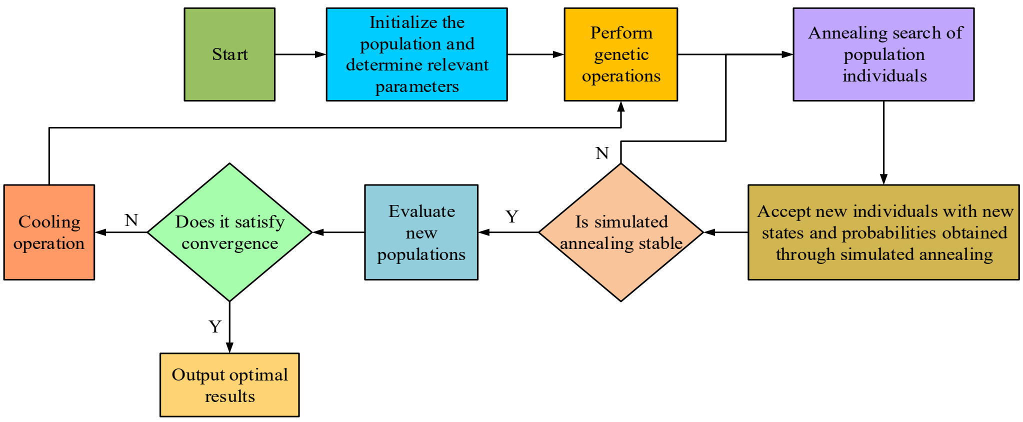


Figure 5: The best delivery path model for takeaway delivery based on customer clustering and GA-SA

4. Performance Analysis of ITD Model Path Planning

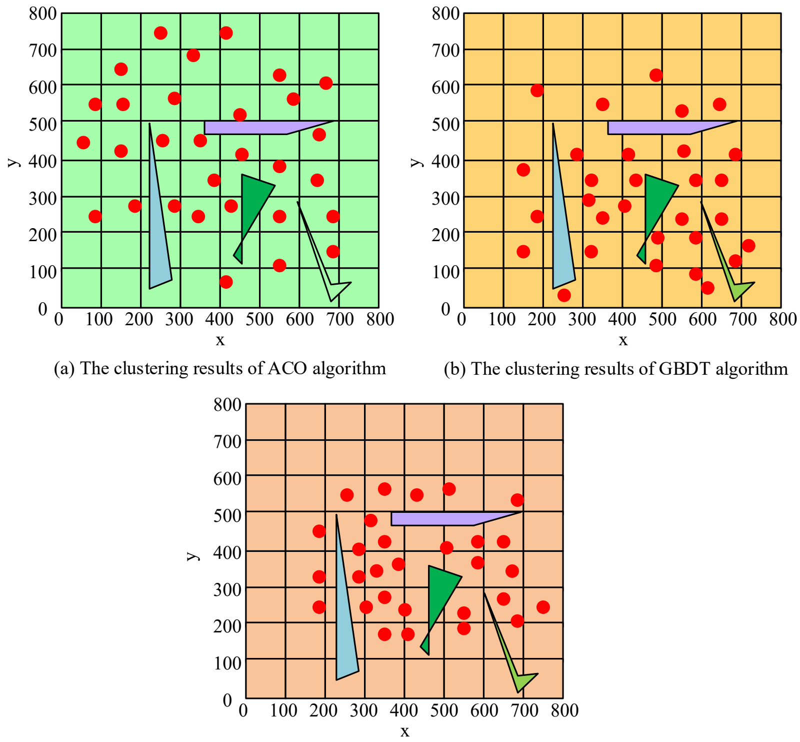
To verify the performance of the ITD model based on customer clustering analysis and GA-SA construction, this study uses clustering results of takeout customer data, algorithm convergence performance, accuracy of clustering data, recall rate, F1 value, and PR curve as validation indicators. Simultaneously, the ACO, Gradient Boosting Decision Tree (GBDT) algorithm are compared with the proposed method to verify the performance of the research model.

4.1 Analysis of Data Processing Performance of ITD model

To verify the performance of the takeaway delivery model, this study uses binary competition as the selection operator and acceptance criteria as the screening criteria for the crossover and mutation operators.

The overall parameter settings are: population size is 100, entry frequency is 500, crossover rate and mutation rate are 0.8 and 0.1, respectively. The initial temperature for simulating cooling is 1000 ℃, and the temperature decay value is 0.8. The shutdown condition is a time stop strategy.

The downtime is the same as the service time at the starting point, which is 60 seconds. Comparison is made between ACO, GBDT, and GA-SA algorithms. Figure 6 shows the comparison results of three methods in clustering takeaway customer data.



(a) The clustering results of ACO algorithm (b) The clustering results of GBDT algorithm

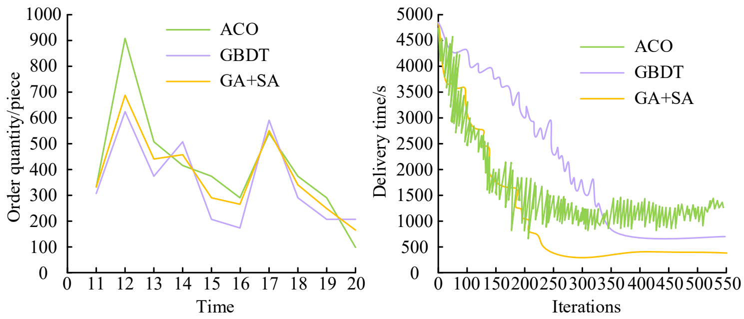
(c) The clustering results of GA+SA algorithm

Figure 6: Comparison results of three methods in clustering takeaway customer data

As shown in Figure 6 (a), in the clustering of the distribution area of takeaway customers, the clustering results of the ACO algorithm are relatively scattered and the cluster boundaries are unclear. As shown in Figure 6 (b), compared to the ACO algorithm, the clustering results of the GBDT algorithm are more concentrated, which can reduce the occurrence of detours to a certain extent. As shown in Figure 6 (c), compared with the previous two algorithms, the clustering results of GA+SA algorithm are most concentrated in distribution, and the boundaries between clusters are more obvious. This indicates a higher degree of similarity within the cluster, thereby improving the accuracy of clustering and greatly enhancing the delivery speed of food delivery drivers. This indicates that the constructed delivery model has excellent performance in customer location clustering. To verify the clustering performance of the model in processing orders, this study used the clustering results of order volume at the same time and the convergence results of three methods as indicators.

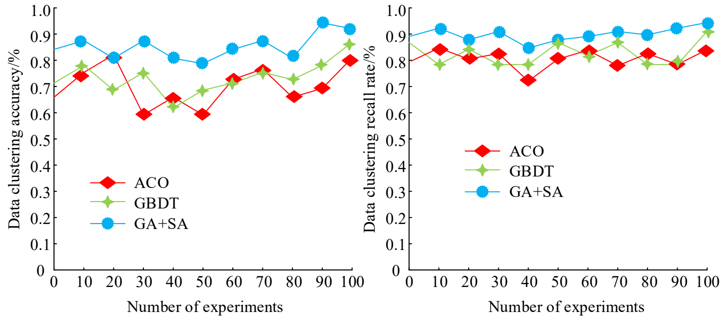
Figure 7 shows the clustering and convergence comparison results of three methods for order quantity at the same time.

In Figure 7 (a), among the clustering results of orders at the same time, the average clustering value of GA+SA orders is 501, and the average clustering values of ACO and GBDT orders are 613 and 562. In Figure 7 (b), there are significant differences in the convergence results of the three methods when simulating the delivery efficiency of the algorithm. GA+SA and GBDT both showed algorithm convergence, but ACO did not show any convergence. This indicates that the delivery model constructed using GA+SA algorithm has more stable performance in order clustering and algorithm convergence. To verify the data processing capability of the delivery delivery model, this study used the accuracy and recall of customer location data clustering for delivery as validation indicators. Figure 8 shows the comparison of data clustering accuracy and recall for three methods.



(a) Clustering results of three methods for order quantity at the same time (b) Comparison of convergence results of three methods

Figure 7: Comparison of clustering and convergence results of three methods for order quantity at the same time

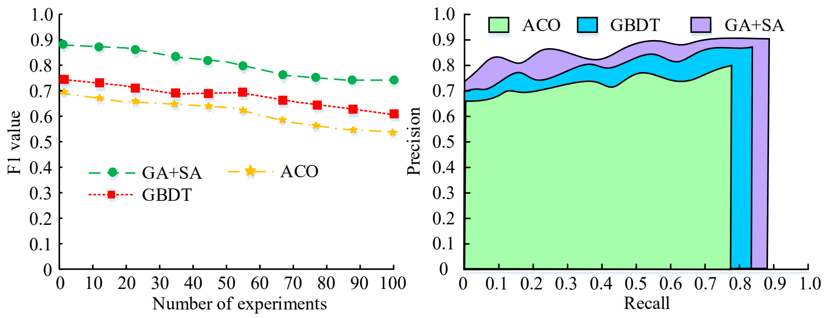


(a) Comparison results of data clustering accuracy using three methods (b) Comparison of recall rates for data clustering using three methods

Figure 8: Comparison results of data clustering accuracy and recall of three methods

In Figure 8 (a), the accuracy of the GA+SA algorithm is 90.87%, while the accuracy of GBDT and ACO are 82.69% and 78.74%, respectively. In Figure 8 (b), the recall rates of GA+SA, GBDT, and ACO in data clustering are 92.18%, 88.79%, and 84.62%, respectively. The comparison between accuracy and recall indicates that the delivery model constructed using GA+SA has higher robustness in customer data clustering. In order to further validate the data processing performance of the model, this study used F1 value and PR curve as comparative indicators. Figure 9 shows the comparison results of F1 values and ROC curves for three methods.

In Figure 9 (a), the F1 values of GA+SA, GBDT, and ACO are 0.83, 0.76, and 0.69, respectively, and the F1 values all show a decreasing trend with increasing number of experiments. In Figure 9 (b), in the comparison of PR curves, the PR curve values of GA+SA, GBDT, and ACO are 0.82, 0.76, and 0.71, respectively. This indicates that the delivery system constructed using GA+SA has more stable performance in data processing.

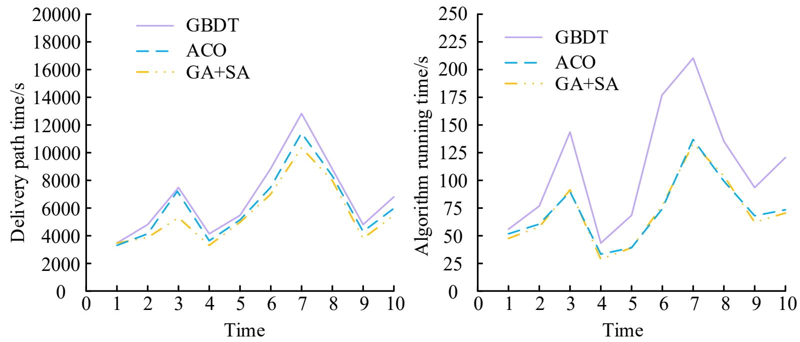


(a) Comparison results of F1 values for three methods (b) Comparison results of PR curves for three methods

Figure 9: Comparison results of F1 values and PR curves of three methods

4.2 Performance Analysis of Distribution Model Path Planning Simulation Experiment

To verify the application performance of the ITD model in road planning, this study uses simulation to verify the performance of the distribution model. The time required for delivery personnel to plan delivery routes and the time required for method operation during the same time period are used as validation indicators. Figure 10 shows the comparison between the delivery path time and the algorithm running time.



(a) Comparison of Three Methods for Planning Delivery Paths in the Same Time Period (b) Comparison of Three Methods for Running Time

Figure 10: Comparison Results of Delivery Path Time and Algorithm Running Time

In Figure 10 (a), the delivery routes for GA+SA, GBDT, and ACO take 4956.7s, 6598.5s, and 7615.2s, respectively. In Figure 10 (b), the operating times of GA+SA, GBDT, and ACO are 74.6s, 80.1s, and 123.5s. This indicates that the constructed ITD model can not only reduce delivery time in path planning, but also shorten its own running time. To further validate the performance of the intelligent delivery model, this study uses the number of takeout delivery personnel, probability of delayed delivery, and average single delivery revenue as evaluation indicators for the performance evaluation of the ITD model under different delivery capabilities. Figure 11 shows the changes in the results of three evaluation indicators under different delivery capabilities.

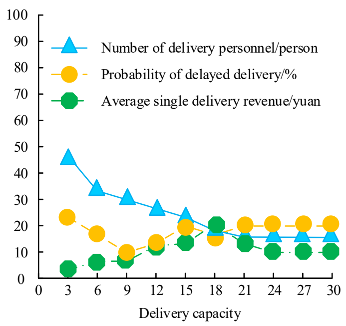


Figure 11: Changes in results of three evaluation indicators under different delivery capabilities

In Figure 11, when the delivery capacity of the delivery personnel reaches 21, the number of delivery personnel no longer changes. When the delivery capacity reaches 21, the probability of delayed delivery tends to stabilize. When the delivery capacity is 18, the average single delivery revenue is the highest. Through the analysis of intelligent delivery models, the delivery capacity of delivery personnel can be set to 18, which can effectively solve the problems of low power supply efficiency and delayed delivery, and maximize the delivery revenue of delivery personnel. To verify the performance of the ITD model, this study compares manual delivery planning with system model delivery planning. Table 1 compares manual delivery planning with system model delivery planning.

In Table 1, in the comparison of total delivery distance, maximum delivery time, and average order waiting time, the values of manual delivery planning and system model delivery planning within three days are 106.43km, 102.4km, 1308s and 903s, 521.33s and 425s, respectively. This indicates that the system model not only improves delivery efficiency, reduces transportation costs, but also reduces customer waiting time, which helps to improve customer satisfaction and service quality.

Table 1. Comparison results between manual delivery planning and system model delivery planning

Project	Manual delivery planning			System Model Distribution Planning		
	Total delivery distance/km	Maximum delivery time/s	Average waiting time for orders/s	Total delivery distance/km	Maximum delivery time/s	Average waiting time for orders/s
first day	103.5	1324	509	99.6	901	416
Second Day	105.6	1295	518	101.8	895	431
Third day	110.2	1305	537	105.8	913	428
Average value	106.43	1308	521.33	102.4	903	425

5. Conclusions

To solve the problem of traditional path planning not being able to obtain effective path planning in the face of complex and changing delivery environments, this study combined K-means with improved GA-SA to construct a food delivery path planning model. The results showed that the delivery path planning process of the distribution model took 4956.7s during the same time period, and the running time in the algorithm path planning was 74.6s. In the comparison between manual delivery planning and system model delivery planning, the average total delivery distance, maximum delivery time, and average order waiting time of the system model delivery planning were shortened by 4.03km, 405s, and 96.33s, respectively. In summary, this study innovatively constructed a new delivery path planning model by combining K-means with improved GA-SA. This model has shown significant advantages in practical applications, as it can significantly shorten the planning time and running time of delivery paths, while also optimizing delivery distance, duration, and order waiting time, thereby improving delivery efficiency and customer satisfaction. Although this study has achieved certain results, there are still certain limitations in the model framework. This is because the path planning used in the study is based on order allocation strategies, which may have some discrepancies with actual delivery. In the future, further exploration of the application effectiveness of this model in more complex delivery environments and larger datasets can be conducted to consolidate the reliability of the ITD path planning model.

Reference

[1] Sha W. Literature review on waste management of online food delivery industry in China. Chinese Journal of Population, Resources and Environment, 2023, 21(3): 197-202.

[2] Li J, Noto M, Zhang Y. Research on Delivery Network Optimization Based on Crowdsourcing Theory. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2022, 26(2): 147-159.

[3] G Mehdi, H Hooman, Y Liu, S Peyman and R. Arif Data Mining Techniques for Web Mining: A Survey. Artificial Intelligence and Applications, 2022, 1(1):3-10.

[4] Falciano A, Puleo S, Colonna F. Novel high-quality takeaway Neapolitan pizza from unused dough balls: hedonic and textural properties, and carbon footprinting assessment. International Journal of Food Science & Technology, 2023, 58(8): 4332-4343.

[5] Colbert S, Wilkinson C, Thornton L. Cross-sectional survey of a convenience sample of Australians who use alcohol home delivery services. Drug and alcohol review, 2023, 42(5): 986-995.

[6] Mohammadi N, Rasoolzadegan A. A two-stage location-sensitive and user preference-aware recommendation system. Expert Systems with Applications, 2022, 191(11): 1-25.

[7] Lang K, Zhao Y. Cloud computing resource scheduling based on improved ANN model takeaway order volume forecast. Journal of Intelligent & Fuzzy Systems, 2021, 40(4): 5905-5915.

[8] Punhani R, Arora V P S, Sabitha A S, Shukla V K. Segmenting E-commerce customer through data mining techniques//Journal of Physics: Conference Series. IOP Publishing, 2021, 1714(1): 1-11.

[9] Fang C, Liu H. Research and application of improved clustering algorithm in retail customer classification. Symmetry, 2021, 13(10): 1-15.

[10] Mesa J P, Montoya A, Ramos-Pollán R. A two-stage data-driven metaheuristic to predict last-mile delivery route sequences. Engineering Applications of Artificial Intelligence, 2023, 125(2): 1-14.

[11] Li X B. A comparative study of different algorithms for optimizing the path of multi-point logistics distribution. Advances in Transportation Studies, 2022, 58(6): 153-162.

[12] Shang Y, Liu F, Qin P, Guo Z, Li Z. Research on path planning of autonomous vehicle based on RRT algorithm of Q-learning and obstacle distribution. Engineering Computations, 2023, 40(5): 1266-1286.

[13] Anndrei A R D, Jesus G L Y. Application of Genetic Algorithms to Optimize Distribution in Food Transport Companies: A Systematic. 2023.

[14] Ait Saadi A, Soukane A, Meraihi Y, Benmessaoud Gabis A, Mirjalili S, Ramdane-Cherif A. UAV path planning using optimization approaches: A survey. Archives of Computational Methods in Engineering, 2022, 29(6): 4233-4284.

[15] Yu Z, Si Z, Li X, Wang D, Song, H. A novel hybrid particle swarm optimization algorithm for path planning of UAVs. IEEE Internet of Things Journal, 2022, 9(22): 22547-22558.

[16] Aktaş A A, Tunalı O, Bayrak A T. Comparative unsupervised clustering approaches for customer segmentation//2021 2nd International Conference on Computing and Data Science (CDS). IEEE, 2021, 15(3): 530-535.

[17] Bellini P, Palesi L A I, Nesi P, Pantaleo G. Multi clustering recommendation system for fashion retail. Multimedia Tools and Applications, 2023, 82(7): 9989-10016.

[18] Othayoth S P, Muthalagu R. Customer segmentation using various machine learning techniques. International Journal of Business Intelligence and Data Mining, 2022, 20(4): 480-496.

[19] Chattoraj M, Vinayakamurthy U R. A self adaptive new crossover operator to improve the efficiency of the genetic algorithm to find the shortest path. Indonesian Journal of Electrical Engineering and Computer Science, 2021, 23(2): 1011-1017.

[20] Wen T, Sun L. Research on optimization algorithm of agv path planning//2021 4th International Conference on Information Systems and Computer Aided Education. 2021,6(9): 1773-1776.

[21] Euchi J, Sadok A. Optimising the travel of home health carers using a hybrid ant colony algorithm//Proceedings of the Institution of Civil Engineers-Transport. Emerald Publishing Limited, 2021, 176(6): 325-336.

[22] Yuan M, Kan X, Chi C, Cao L, Shu H, Fan Y. An adaptive simulated annealing and artificial fish swarm algorithm for the optimization of multi-depot express delivery vehicle routing. Intelligent data analysis, 2022, 26(1): 239-256.

[23] Zhai L, Feng S. A novel evacuation path planning method based on improved genetic algorithm. Journal of Intelligent & Fuzzy Systems, 2022, 42(3): 1813-1823.