IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received August 14, 2020, accepted August 28, 2020, date of publication September 15, 2020, date of current version December 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3024281

Two-Stage Optimization Model of Agricultural Product Distribution in Remote Rural Areas

HAO ZHANG[1,3], HUIXIA FENG[1], AND HONGMEI WANG[2,3]

[1]School of E-Business and Logistic, Beijing Technology and Business University, Beijing 100048, China
[2]Beijing Jiaotong University, Beijing 100044, China
[3]Beijing Food Safety Research Base, Beijing 100044, China

Corresponding author: Huixia Feng (fenghuixia2019@163.com)

This work was supported in part by the Project of Beijing Philosophy and Social Science under Grant 17GLB013, in part by the Project of 2018 Beijing Talents Foundation of Organization Department of Beijing Municipal Committee of the CPC under Grant 2018000026833ZS09, in part by the Project of Science and Technology Innovation Service Capacity Provincial-Ministerial Scientific Research Platform Construction Social Science Provincial-Ministerial Scientific Research Platform Construction Project under Grant 19008020111, in part by the Project of the Planning Subject of the 12th Five Year Plan in National Science and Technology for Rural Development in China Demonstration of Key Technology and Equipment for the Safe Distribution of Agricultural Logistics under Grant 2015BAD18B01, and in part by the Project of 2020 Postgraduate Research Capacity Improvement.

ABSTRACT Due to the lack of systematically optimized logistics networks in many remote rural areas, the online sales of agricultural products in these areas have the disadvantages of high cost, high damage and slow speed. To address these logistic problems, this paper proposes a two-stage layout optimization model of agricultural product joint distribution centres based on the geographical features of remote rural areas. The number and location of distribution centres are selected in two stages to optimize the logistics network. Chengkou County, located in Chongqing city, China, has been selected as the study area. In the first stage, AP clustering is carried out using the distance between the logistics nodes of villages and towns, and the transit station of the source of agricultural products is obtained using the correlation between the nodes, which are regarded as alternative locations for the second-level logistics nodes. In the second stage, a joint distribution centre optimization model with the lowest cost is constructed. A fruit fly optimization algorithm is used to select the optimal locations from the alternative locations as the second-level logistics nodes and obtain the specific delivery path. Optimizing the logistics network of remote rural areas can reduce the logistics costs of agricultural products in these areas, promote online sales of agricultural products, and provide the government and logistics companies with theoretical references to open up new markets in remote areas.

INDEX TERMS AP clustering, fruit fly optimization algorithm, joint distribution centres, two-stage logistics layout optimization.

NOTATIONS

One can build the joint distribution model by setting the following parameters and variables:

VARIABLES AND FUNCTIONS

n	– total number of township logistics nodes;
N	– total number of source transfer stations;
M	– total number of joint distribution centres;
num	– set of township logistics nodes, num = {1,2,..., n};
A	– set of source transfer stations, A = {1,2,..., N};
B	– set of joint distribution centres, B = {1,2,..., M};
C	–set of demand points
P_{BL}	–unit land price of distribution centre B;
L	– required area for a single distribution centre;
C_{conB}	–annual construction cost of distribution centre B;
C_w	–individual distribution centre facility and operating costs per year;
D	– maximum capacity of the agricultural product distribution centre;
Dis_{AB}	– spatial distance from A to B;
Dis_{BC}	–spatial distance from B to C;

The associate editor coordinating the review of this manuscript and approving it for publication was Zhiwu Li.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

q_{AB} – annual average agricultural product traffic from A to B;
q_{BC} –annual average agricultural product traffic from B to C;
t – fixed life of each agricultural product distribution centre;

AS – average annual total supply in the region;
θ –agricultural product corruption rate;
p – average selling price of agricultural products;
h – unit cost per delivery vehicle;
h_{AB} –unit cost per delivery vehicle from A to B;
h_{BC} –unit cost per delivery vehicle from B to C;
x_B –one Boolean variable to mark whether to build a distribution centre in B;

$$x_B = \begin{cases} 1, \textit{setting up a distribution center in candidate B} \\ 0, \textit{otherwise} \end{cases}$$

y_{AB} one Boolean variable to mark whether to transport agricultural products from A to B;

$$y_{AB} = \begin{cases} 1, \ \textit{transport agricultural products from A to B} \\ 0, \ \textit{otherwise} \end{cases}$$

y_{BC} one Boolean variable to mark whether to transport agricultural products from B to C;

$$y_{BC} = \begin{cases} 1, \ \text{transport agricultural products from B to C} \\ 0, \ \text{otherwise} \end{cases}$$

ABBREVIATIONS
AP – affinity propagation
FOA – fruit fly optimization algorithm

I. INTRODUCTION
The logistics industry integrates transportation, warehousing, freight forwarding, information and other industries, thereby playing an important role in supporting the development of the national economy. Operations in logistics are significant economic activities for competitive businesses, but with the increasing use of e-commerce and cargo transportation, the current problem regarding logistics has attracted the attention of several production companies and scholars. Taking China as an example, according to the "China Cold Chain Logistics Development Report (2018)", with the application and popularization of the Internet, the transaction scale of China's fresh agricultural products e-commerce market has maintained a growth rate of more than 50% for five consecutive years. However, there are many problems in China's agricultural product logistics, especially in remote areas. These problems are mainly reflected in the high cost and loss rate of agricultural product logistics. According to the data, the logistics costs of rural areas with convenient transportation are 2-3 times those of urban areas, and the logistics costs of remote villages are 4-5 times the urban costs. The circulation costs of agricultural products in some regions generally account for 30%-40% of the total cost, and fresh products account for more than 60%. Logistics cost of e-commerce for fresh agricultural products accounts for 25%-40% of the sales price, which is much higher than the logistics cost of other products, especially the distribution cost accounts for 35% of the total logistics cost [1]. Agricultural product logistics costs account for 70% of the cost of perishable goods, and according to international standards, the perishable goods logistics cost does not exceed 50% of its total cost [2]. According to statistics, the average loss rates of fresh agricultural products, such as fruits and vegetables, can reach 25%-30% because of picking, transportation and storage logistics, while the loss rates of fruits and vegetables in developed countries are less than 5% and only 1% in the United States [3].

Based on the characteristics of the Internet era and the remote rural agricultural product logistics in the context of rural revitalization, this paper selects the remote rural agricultural product logistics model, considers the optimization method of logistics nodes layout of remote areas, and constructs its optimization model. This paper establishes a two-stage multi-logistical node location model. In the first stage, the first-level source transfer stations are obtained by clustering method and used as alternative locations for the second-level logistics nodes. In the second stage, an optimization model of the joint distribution centers for online sales of agricultural products in remote rural areas was constructed. The locations are screened and the final joint distribution centers are calculated using Fruit Fly Optimization Algorithm as the second level logistics nodes. In the process of layout optimization, the specific geographical situation of the scattered agricultural products in remote villages is considered, which makes the joint distribution center established through two stages more tally with the actual situation. In the site selection, not only the loss cost, transportation cost, operation cost and other factors are fully considered in the established location model, but also the route optimization is carried out when the distribution center is selected.

With the development of location theory, numerous methods have been proposed for the location of logistics distribution centres. Due to the different algorithms and factors considered, many types of logistics distribution centre location problems exist. Based on the economic benefits of the logistics system and the specific case, it is easy to find a problem of two-stage multi-logistics distribution centre locations based on cost minimization. Therefore, this paper aimed to determine the number and location of the two levels of logistics nodes and transportation route planning.

The logistics network optimization problem in this paper minimizes the cost of the whole logistics system by solving the following sub-problems:

-How to determine the number and location of the source transfer stations;

-How to determine the location of the joint distribution centres when the number of distribution centres needed to be built is known;

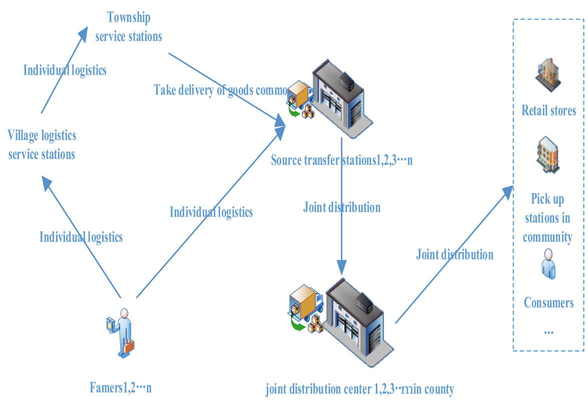

FIGURE 1. **Flow chart of online sales of agricultural products in remote rural products.**

-How to optimize the distribution routing after determining the number and location of the joint distribution centres.

For the first sub-problem, because the construction cost of the source transfer stations is not high, the focus is on how to solve the location to minimize the distribution cost, which depends mainly on the distance between logistics nodes. Thus, solving the first sub-problem requires minimizing the total distance between logistics nodes by site selection.

For the second and third problems, it is necessary to consider the characteristics of traffic conditions and infrastructure in remote villages and a series of constraints such as customer demand and vehicle load restrictions. Finally, a distribution centre optimization model with the lowest total cost can be built using a suitable heuristic algorithm.

Because the distribution centres in the county are far from the demand points outside the county, so the demand points outside the county are only served by the distribution centres outside the county. Therefore, the main service receivers of the distribution centre in the county are the demand points inside the county and the distribution centres outside the county. Thus, the overall distribution flow chart is constructed as shown in Fig1.

II. LITERATURE REVIEW

Because of increasing logistics costs, many scholars have realized the importance of joint distribution. Kazemi *et al.* [4] noted that joint distribution not only enables enterprises to enjoy long-term economic benefits but also improves the operational efficiency of the entire supply chain, serving as an effective solution to distribution planning. Sheng [5] constructed a four-in-one rural e-commerce alliance operation mode based on a long tail theory of ''third-party logistics + postal logistics + passenger transport logistic + grassroots logistics''.

In logistics studies, many papers have been published on the location problem of distribution centres and the distribution routing problem. In 1929, Harold Hotelling proposed the location of two competing suppliers on a straight line. Kuehn and Hamburger [6] proposed a heuristic algorithm to solve the warehouse location problem. Many scholars have carried out in-depth research on the location of distribution centres since then, including establishing various mathematical models [7] and solving algorithms [8], [9]. Various forms of location models have been developed, which can be categorized into analytic, continuous and network models [10]. The continuous location model is calculated with the centre of gravity method [11], [12]. Aiken [13] proposed several discrete location models, such as dynamic programming models, linear programming models [14]–[16],and 0-1 integer programming models [17], when studying the location model of a distribution centre. Through in-depth analysis and comparison of several models, it can be seen that these models are aimed at achieving the minimum total cost of selecting facilities, but the difference lies in that different programming models have different forms of objective functions and constraints. The Delphi method, which is based on qualitative analysis, is commonly used in the fuzzy evaluation method [18], [19]. Chen [20] made location decisions for fuzzy multi-objective decision-making data and put forward a multi-objective optimization decision-making method. Klapita and Švecová [21] dealt with a possible method of finding the optimal location of logistics centres at uncertain costs represented by fuzzy numbers to minimize the complete costs of a system. Turskis and Zavadskas [22] presented a newly developed ARAS–F method to select the most suitable site for a logistics centre among a set of alternatives. With an illustrative example—the selection of a logistics centre location—the proposed methodology was validated. Xu *et al.* [23] solved the single-plant and the multiplant location decision problem by establishing the conceptual and mathematical model.

For the computation of a location optimization model and distribution routing optimization, the gravity centre location model is mainly calibrated by the iterative method. Various heuristic algorithms are used to solve the NP hard problems, mainly including the ant colony algorithm [24], the genetic algorithm [4], the particle swarm optimization algorithm [25], [26] and corresponding improvements [27], and the fruit fly optimization algorithm [28]. Hiassat *et al.* [29] considered the characteristics of perishable goods, established an optimization model of the distribution centre inventory problem, and used a genetic algorithm and local search heuristic method to solve it. Song and Ko [30] considered the characteristics of cold chain cars and established a nonlinear mathematical model. Bo [31] established the mathematical model of logistics distribution center location, and optimized the solution to realize the optimal allocation of distribution path. A fruit fly optimization algorithm based on Logistic chaotic system was proposed. Mulloorakam *et al.* [32] considered the combined objective capacitated vehicle routing problem (CVRP) based on the genetic algorithm. Simsir and Ekmekci [33]used the artificial bee colony (ABC) algorithm to produce low-cost solutions based on few parameters. Habibi *et al.* [34] presented a location inventory routing problem (LIRP) optimization model to reduce the total cost. Xu and Yin [35] introduced ant colony algorithm and particle swarm optimization algorithm and it provided a reference for the algorithm in this paper.

Tuzun and Burke [36] and others used a two-stage tabu search algorithm to solve the LRP problem, which separated the location problem from the routing problem and used a two-stage idea to solve it. Wu [37] and others used a simulated annealing algorithm to solve the LRP problem of multiple distribution centres and multiple vehicles. Albareda-Sambola *et al.* [38] used a simulated annealing algorithm to solve the LRP problem. Albareda-Sambola *et al.* [38] used tabu search to study a deterministic LRP with a single time window. Qureshi *et al.* [39] used the tabu search algorithm to solve the LRP location allocation problem and a simulation system to solve the vehicle routing problem. Sadjady and Davoudpour [40] formulated a two-echelon supply chain network design problem as a mixed-integer programming model and solved it using a Lagrangian-based heuristic algorithm. In the literature on logistics location, there are many studies on agricultural product location logistics. The study [41] of logistics can be considered as the location of non-linear transportation costs in the distribution system of agricultural products, and the problem was solved by the branch and bound method. Hwang [42] considered that the decay rate of goods presents an exponential function, and a two-stage random coverage location model for perishable goods was proposed. Gong *et al.* [43] proposed a location model of a perishable goods distribution centre based on particle swarm optimization (PSO) considering the security inventory and the capacity limitation of a distribution centre.

Although the above scholars conducted research on logistics networks, including distribution centre locations, common distribution models, optimized distribution paths, etc., there is room for improvement. First, most of the location optimization models established at present are for single-level distribution centres, but there are few studies on multi-level distribution centre location models, and the algorithm process is complex and computationally intensive. The multi-stage distribution center can better complete the distribution task and is more convenient and efficient for remote areas. The multi-level distribution center serves as a bridge between the single-level distribution center and the delivery point. Second, these studies solved the layout problem or the distribution path problem separately but lack the research that considers these two issues comprehensively. Third, there are few studies on the logistics network optimization of remote rural areas as a specific case. Therefore, it is necessary, important, and innovative to study the layout of agricultural product logistics and the distribution routes of fresh agricultural products in remote areas.

III. AP CLUSTERING ALGORITHM

This study is mainly designed to optimize the location of distribution centres in remote areas that have relatively scattered residential areas and wide geographical areas. Therefore, in the first stage, one clustering algorithm must be used to cluster these isolated villages into several regions and designate a source transfer station in each region.

The clustering algorithms include the k-means algorithm, SOM neural network, FCM algorithm, hierarchical clustering algorithm, etc. However, in the clustering process, the clustering number of most clustering methods is difficult to determine. The application of the AP clustering algorithm in the location selection of the agricultural product logistics centre does not need to specify the number of clusters in advance, which can simplify the selection process of rural logistics centre nodes and is more suitable for the complex situation of the rural areas of China. Therefore, the AP clustering algorithm is used to optimize the location of the source transfer stations.

The basic idea of the AP clustering algorithm is to select the location of the agricultural product logistics centres as follows. According to the correlation between the rural logistics nodes (the AP clustering algorithm in this chapter measures the correlation between each node based on the spatial distance between the logistics nodes), first cluster the logistics nodes in rural logistics systems, then regard the obtained exemplars as the logistics centres, and finally regard the other nodes within the cluster as the service objects of the exemplar.

A. PROCESS DESCRIPTION

First, the distance between the nodes is derived from the position information of the logistics nodes to form the node distance matrix. Second, the distance matrix is preprocessed so that the distance matrix can better reflect the relationship between the logistics nodes. Finally, based on the preprocessed data used to cluster logistics nodes, regard the exemplars obtained as the locations of the logistics centres and the number of clusters as the number of logistics centres; the flow chart is shown in Fig. 2.

B. DESCRIPTION OF AP CLUSTERING ALGORITHMS

1) DISTANCE MATRIX ACQUISITION

First, we obtain the spatial distance between the nodes according to the position information between the rural nodes and construct the distance matrix. Due to the inadequacy of rural transportation facilities, the geometric distance between nodes does not necessarily reflect the cost of logistics transportation. Therefore, this paper uses the length of the road between nodes to represent the ''space distance'' between nodes. $Node^i$ and $Node^j$ represent any two logistics nodes, and $Dis\,(i, j)$ represents the distance between the nodes. The distance matrix of n rural logistics nodes reflecting the spatial distance between the logistics nodes is $DisD$, and the matrix element is $DisD\,(i, j)$.

2) DISTANCE MATRIX PREPROCESSING

The AP clustering algorithm clusters all nodes by measuring the correlation between nodes according to the value of the matrix elements. The larger the value of the matrix elements is, the larger the correlation between the nodes. However,

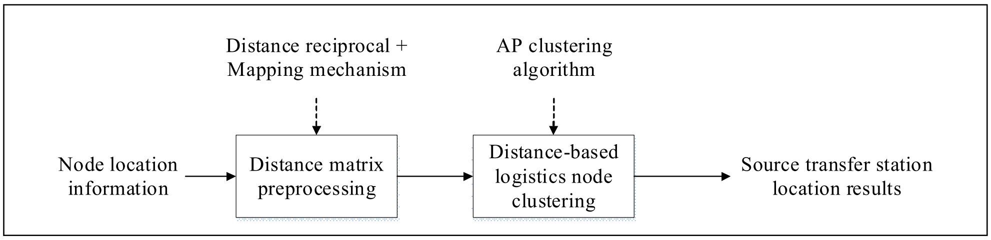

FIGURE 2. Flow chart of site selection of agricultural product source transfer station based on AP clustering.

the distance matrix reflects the opposite: the larger the original value of the matrix is (the larger spatial distance between the nodes), the smaller the correlation between the nodes, which is different from the input data required by the AP clustering algorithm. Therefore, the distance matrix must be preprocessed first.

The element values of the distance matrix are processed as shown in formula (1).

$$Dis\,(i,j) = \begin{cases} 1/Dis\,(i,j), & i \neq j \\ minDis\,(i,j)/0.9, & i = j \end{cases} \tag{1}$$

After preprocessing, the values of the elements of the distance matrix, which range between 0-1, can reflect the correlation between the nodes.

Because rural areas are generally broad, the distance between the logistics nodes in the rural logistics system may vary widely. The clustering results obtained by directly inputting the distance matrix into the AP clustering algorithm may not be ideal [45]. Therefore, the distance matrix after preprocessing is processed again using the "mapping mechanism" to adjust the results to some extent so that the distance between the nodes is mapped in a relatively small range.

The mapping rule is as follows: use the matrix element plus adjustment coefficient b (the value is smaller), and then take the logarithm of a as the base to generate a new value $f\,(i,j)$ of the matrix element and the distance matrix $[f\,(i,j)]$ after mapping. During the test, this paper assumes $a = 2$, $b = 1$, as shown in formula (2).

$$f\,(i,j) = \{log_a(Dis\,(i,j) + b)\} \tag{2}$$

3) CLUSTERING OF LOGISTICS NODES

The preprocessed distance matrix $[f\,(i,j)]$ is used as the input data of the AP clustering algorithm, and the output result is the collection of cluster centre points.

The clustering process of the AP clustering algorithm on the input matrix is described in detail below:

The AP clustering algorithm regards n logistics nodes as potential cluster centres based on two types of information passed between nodes: responsibility and availability. If node k represents the candidate node, $S\,(k,k)$ is the correlation between nodes. $R\,(i,k)$ represents the degree of responsibility which is the numerical information sent from the logistics node i to the candidate k and is used to describe the responsibility of node k regarding as the logistics distribution center of node i; $A\,(i,k)$ represents the degree of availability, which is the numerical information sent from candidate k to logistics node i and is used to describe the possibility of node i to select node k as its logistics distribution centre. $[R\,(i,j)]$ is a responsibility degree matrix, and $[A\,(i,j)]$ is a candidate degree matrix.

The calculation process of the node responsibility value and the availability value is shown in formulas (3) (4) and (5).

$$R\,(i,k) = Dis\,(i,k) - max_{j\neq k}A\,(i,j) + Dis\,(i,j) \tag{3}$$

$$A\,(i,k) = min\{0,R\,(k,k)\sum\nolimits_{j\neq i,k} max\,\{0,R\,(j,k)\}\} \tag{4}$$

$$A\,(k,k) = \sum\nolimits_{j\neq k} max\,\{0,R\,(j,k)\} \tag{5}$$

The iterative process continuously updates the r and a values of each point until $R\,(i,k)$ *and* $A\,(i,k)$ remain unchanged, generating multiple high-quality exemplars (multiple final logistics centres).

Since the execution process of the AP clustering algorithm is prone to oscillation, the damping factor γ is introduced to adjust the calculation speed and avoid oscillation. In the current iteration process, the update result of responsibility value $R\,(i,k)$ and the availability value $A\,(i,k)$ of the logistics node i is obtained by weighting the result of the previous iteration. The value of γ is generally larger than or equal to 0.5 and less than 1. In this calculation, it is set to 0.5. If the number of iterations reaches the set number of times, the calculation process is terminated; otherwise, it continues to iterate. The calculation process is shown in formulas (6) and (7).

$$R_{i+1}\,(i,j) = \gamma * R_i\,(i,j) + (1-\gamma) * R^{old}_{i+1}\,(i,j) \tag{6}$$

$$A_{i+1}\,(i,j) = \gamma * A_i\,(i,j) + (1-\gamma) * A^{old}_{i+1}\,(i,j) \tag{7}$$

After the distance matrix of the logistics nodes is clustered, the logistics nodes are divided into N clusters according to the correlation between them. Each cluster is recorded as $C_N(\mathrm{N} < \mathrm{n})$ such that:

$$C_N = \{C^1_N, C^2_N, \ldots\ldots C^i_N\} \tag{8}$$

C^i_N is the node in the logistics system, $C^c_N\,(i,j)$ is the location of the logistics centre, which is a suitable logistics centre for C^1_N *and* C^2_N and other logistics nodes.

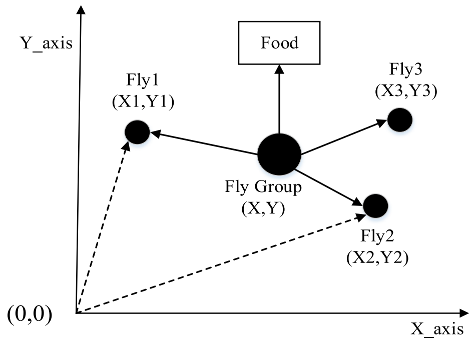

FIGURE 3. Illustration of the fruit fly group and fruit fly group iterative food search.

IV. FOA

The FOA is a new group intelligence optimization algorithm based on the food finding behaviour of the fruit fly, proposed by Pan [44]. The fruit fly optimization algorithm is an abstraction of powerful smell and visual searches of food finding behaviour. It has the advantages of group collaboration, information sharing, easy programming, and faster search speed.

The second stage selects several source transfer stations from alternative distribution centres as the joint distribution centres. This is a non-deterministic polynomial (NP) problem. Therefore, when solving an NP problem of multi-distribution centre location selection, a heuristic algorithm is more convenient.

A. SOLUTION ALGORITHM

First, initialize the population size, Sizepop, the maximum number of iterations, Maxgen, and randomly initialize the position of the fruit fly group ($X_{\_axis}$, $Y_{\_axis}$). The random initial fruit fly swarm location is shown in Fig. 3.

Taking the initial position as the starting point, the flying direction (*RandoM*) and the optimization step (*Value*) of the individual flies are randomly given, and the RandomValue is used as the search distance. The formula is as follows:

$$\begin{cases} X_i = X_{\_axis} + Random\ Value \\ Y_i = Y_{\_axis} + Random\ Value \end{cases} \tag{9}$$

The distance to the origin (Dist) is estimated first, and then the smell concentration judgement value (S) is calculated, which is the reciprocal of distance.

$$Dist_i = \sqrt{X_i^2 + Y_i^2} \tag{10}$$

$$S_i = \frac{1}{Dist_i} \tag{11}$$

Substitute smell concentration judgement value (S) into the smell concentration judgement function (fitness function) to find the smell concentration ($Smell_i$) of the individual location of the fruit fly; the formula is as follows:

$$Smell_i = Function\,(S_i) \tag{12}$$

Determine the fruit fly with optimal smell concentration (the maximal value) among the fruit fly swarm and record

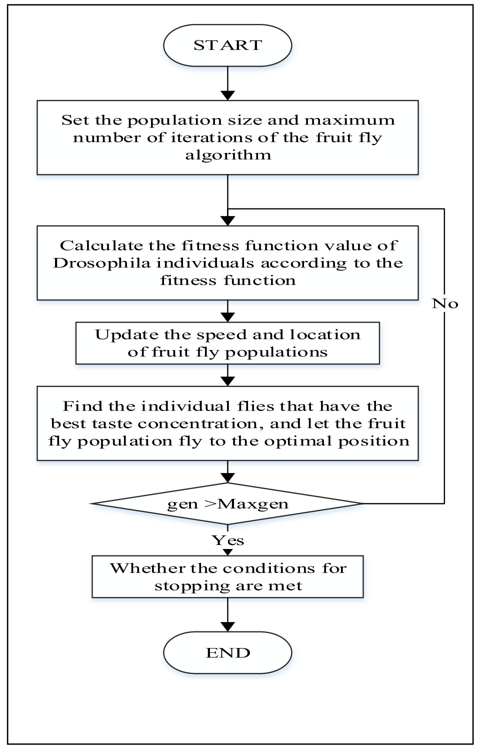

FIGURE 4. Flow chart for optimizing the location of the distribution centre by FOA.

the optimal individual smell concentration (*bestSmell*) and the current optimal individual number (*bestIndex*).

$$[bestSmell, bestIndex] = max(Smell) \tag{13}$$

Keep the best smell concentration value ($bestSmell_i$) and X, Y coordinate, and at that moment, other individuals of the fruit fly swarm will fly to the optimal fruit fly through visual observation.

$$\begin{cases} smellBest = bestSmell \\ X_{axis} = X(bestIndex) \\ Y_{axis} = Y(bestIndex) \end{cases} \tag{14}$$

Start the iteration, repeat the implementation of Steps 2-5 when the number of iterations is less than the minimum number of iterations, then judge if the smell concentration is superior to the previous iterative smell concentration; if so, implement Step 6; otherwise, keep the previous minimum taste concentration and end the algorithm.

B. DISTRIBUTION CENTRE LOCATION OPTIMIZATION BASED ON FOA

The FOA optimizes the location of the logistics distribution centres as follows. Fig. 4 shows the flow chart of optimization.

Step 1: Set the population size of the IFOA (improved fruit fly optimization algorithm) algorithm Sizepop, the maximum number of iterations Maxgen;

Step 2: According to the fitness function formula (12), calculate the fitness function value of the individual fruit fly and find the position and optimal value of the individual and global optimal individual of the fruit fly;

Step 3: Update the speed and position of the fruit fly population;

Step 4: Calculate the fitness and update the position and speed at the same time;

Step 5: If gen > Maxgen, save the optimal solution; otherwise gen = gen +1, and go to Step 2;

Step 6: Select the best location and optimal distribution range of the corresponding logistics distribution centre according to the optimal location.

V. MATHEMATICAL FORMULATION

A. HYPOTHESES

The specific location of each supply location and the number of suppliers is known;

There are transportation costs that are proportional to the transportation volume and transportation distance;

The construction cost of each source transfer station is the same;

The construction area of each distribution centre is the same;

The damage costs of agricultural products in warehouse storage and the storage costs of distribution centres are not considered;

Demand outweighs supply;

The maximum storage per distribution centre is the same and known;

It is assumed that goods of each township logistics node are only transported to one source transfer station, goods of each source transfer station are only transported to one distribution centre, and demand points can only be served by one distribution centre;

The specific steps for calculating the number of distribution centres are as follows: Use linear regression to predict the annual average agricultural products and the average annual agricultural commodity rate in the region in the next decade to calculate the total annual agricultural product supply. Then, the annual average total supply of agricultural products is divided by the maximum storage capacity of the distribution centre to obtain the number of distribution centres.

B. JOINT DISTRIBUTION CENTRE LOCATION MODEL

The objective of the joint distribution centre model is to minimize the total cost. The objective function and constraints are written as follows:

The fixed cost of the joint distribution centre could be calculated according to the following equation:

$$\sum_{B=1}^{M} P_{BL} * (L/t) + C_w L * (AS/D) \tag{15}$$

The transportation cost from the source transfer stations to the joint distribution centres can be expressed as follows:

$$\sum_{B=1}^{M}\sum_{A=1}^{N} q_{AB} h_{AB} Dis_{AB} \tag{16}$$

The transportation cost from the joint distribution centres to the demand points can be expressed as follows:

$$\sum_{B=1}^{M}\left\{\sum_{C}^{K} q_{BC} h_{BC} * Dis_{BC}\right\} \tag{17}$$

The damage cost of agriculture products transported from the source transfer stations to the joint distribution centres and to the demand points can be represented as follows, respectively.

$$\sum_{B=1}^{M}\left\{\sum_{A=1}^{N} q_{AB} * p * \left[\frac{1}{(1-\theta)^{Dis_{AB}}} - 1\right]\right\} \tag{18}$$

$$\sum_{B=1}^{M}\left\{\sum_{C}^{K} q_{BC} * p * \left[\frac{1}{(1-\theta)^{Dis_{BC}}} - 1\right]\right\} \tag{19}$$

Then, the objective function is used to make the construction and operation costs of the joint distribution centres and the distribution and damage costs from the source transfer stations to the joint distribution centres the lowest, which is calculated as follows:

The number of joint distribution centres selected from the alternative centres is expressed as follows:

$$\sum_{B=1}^{M} x_B = AS/D \tag{21}$$

The quantity of goods transported to the joint distribution centres does not exceed the maximum storage capacity, which is written as follows:

$$\sum_{A=1}^{N} q_{AB} \le D_B, \quad B = 1, 2 \ldots M \tag{22}$$

It is assumed that the goods of each township logistics node are only transported to one source transfer station, and the goods of each source transfer station are only transported to one distribution centre as follows:

$$\sum_{B=1}^{M} y_{AB} = 1, \quad A = 1, 2 \ldots N \tag{23}$$

$$\sum_{C=1}^{K} y_{BC} = 1, \quad B = 1, 2 \ldots M \tag{24}$$

The total amount of agricultural products transported from the source transfer stations to the joint distribution centres is equal to the total agricultural product supply in the area, which is as follows:

$$AS = \sum_{A=1}^{N} q_{AB}, \quad B = 1, 2 \ldots M \tag{25}$$

TABLE 1. All 31 township rural logistics nodes in Chengkou County.

Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node8
Gaonan Township	Zuolan Township	Houyu Township	Bashan Township	Yanhe Township	Zhongxi Township	Shuanghe Township	Pingba Township
Node 9	Node 10	Node 11	Node 12	Node 13	Node14	Node 15	Node 16
Miaoba Township	Zhouxi Township	Mingtong Township	Jiming Township	Xianyi Township	Liaozi Township	Heyu Township	Beiping Township
Node 17	Node 18	Node 19	Node 20	Node 21	Node 22	Node 23	Node 24
Xiuqi Townsip	Shanmu Township	Shifang Township	Zhipinj Township	Mingzhong Township	Houping Township	Gaoguan Town	Dongan Town
Node 25	Node 26	Node 27	Node 28	Node 29	Node 30	Node 31	
Longtian Township	Chengkou County	Lantian Township	Huangan Township	Huangxi Township	Mingyue Township	Jianling Township	

VI. CASE STUDY

A. BACKGROUND DESCRIPTION

Chengkou County is located in the northernmost part of Chongqing, which is at the junction of the three provinces (cities) of Chongqing, Shanxi and Sichuan in China. The county covers an area of 3,292 square kilometres. There are two subdistrict offices in Gecheng and Fuxing and 31 townships. There are 184 administrative villages in 22 communities. In 2017, the registered population reached 253,000, of which the agricultural population was 216,500.

Chengkou County has an excellent ecological environment and outstanding agricultural resources. The agricultural land area is very large, and the output is high. There are many free-range chickens, potatoes, freshwater fishes and other agricultural products. Chengkou free-range chickens, honey, artichokes, etc. However, the economic development of Chengkou County is low. In 2016, the GDP in Chengkou County was only 4.512 billion yuan, twenty times less than Wan County, with the highest GDP of 89.739 billion yuan. At the same time, the agricultural commodity rate in Chengkou County is low, far lower than the city's average level, which is far from the Dadukou District, which has the highest agricultural commodity rate in the city. Chengkou County has the lowest agricultural commodity rate except for Wuxi County and Wushan County. The specific data are shown in Fig. 5 [46].

Generally, Chengkou County has the following agricultural logistics problems:

Insufficient investment in agricultural products logistics, and there is a lack of large-scale specialized logistics enterprises.

Insufficient logistics infrastructure: there are no railways or national highways in Chengkou County, which seriously hinders the external distribution of agricultural products.

Remote geographical location leads to long-term transportation of agricultural products, and decay and deterioration easily occur.

Farmers' logistics awareness is weak, and they are reluctant to hand over the logistics business to professional logistics enterprises, resulting in product dispersion, low transportation efficiency and high logistics costs.

Lack of distribution mode for fresh agricultural products. The status quo of agricultural products logistics still cannot adapt to agricultural development, which becomes the bottleneck of agricultural development in Chengkou County.

B. OPTIMIZATION OF SOURCE TRANSFER STATION LOCATION

To solve the related problems of Chengkou County logistics, the established model and method are used. Therefore, it is necessary to address the first-stage problem, that is, to determine the location of source transfer stations in towns and villages.

1) SPATIAL DISTANCE ACQUISITION OF TOWNSHIP LOGISTICS NODES

By consulting relevant data, the relevant information of rural logistics nodes in 31 townships in Chengkou County is obtained as shown in Table 1 and Fig. 6.

$$minCost = \sum_{B=1}^{M} P_{BL}*(L/t) + C_w L*(AS/D) + \sum_{B=1}^{M} \left\{ \begin{array}{l} \left\{ \sum_{A=1}^{N} q_{AB} h_{AB} Dis_{AB} + \sum_{A=1}^{N} q_{AB} * p * \left[\frac{1}{(1-\theta)^{Dis_{AB}}} - 1 \right] \right\} * y_{AB} \\ + \left\{ \sum_{C}^{K} q_{BC} h_{BC} Dis_{BC} + \sum_{C}^{K} q_{BC} * p * \left[\frac{1}{(1-\theta)^{Dis_{BC}}} - 1 \right] \right\} y_{BC} \end{array} \right\} * x_B \quad (20)$$

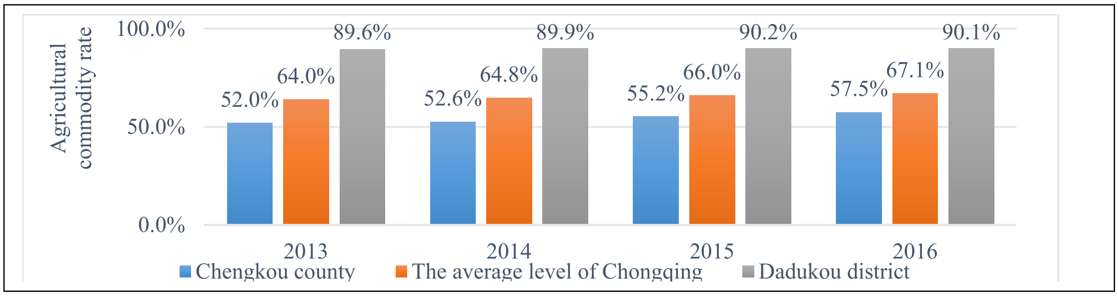

FIGURE 5. Annual Agricultural Commodity Rate of Chengkou County, Chongqing, 2013-2016.

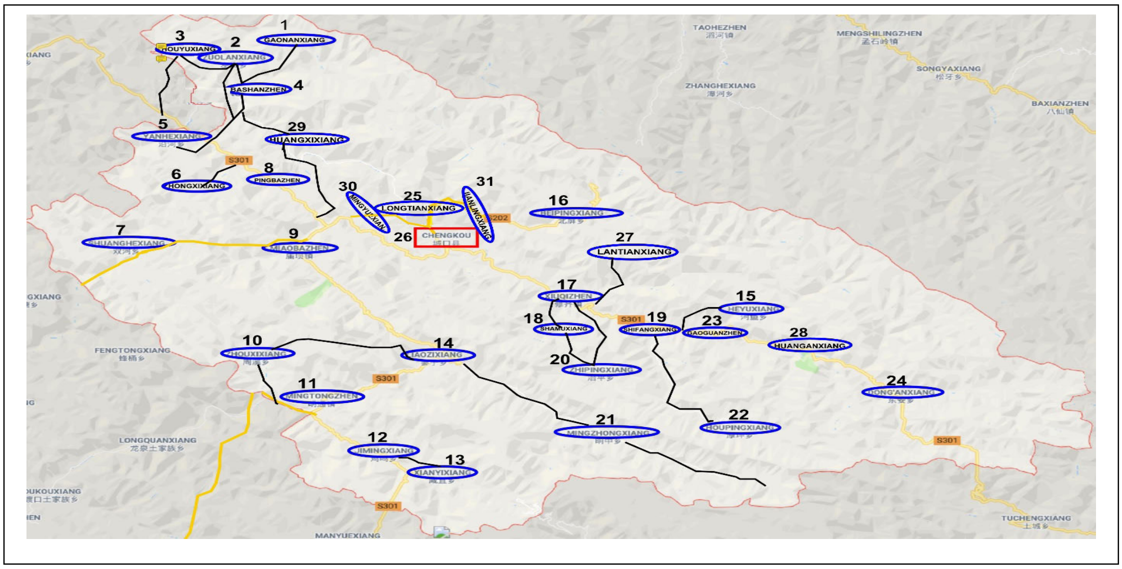

FIGURE 6. Spatial location of rural logistics nodes in 31 townships in Chengkou County.

Table 2 shows part of the spatial distance (unit: km) between the 31 township logistics nodes shown in Fig. 6. The details are shown in Appendix A. To make the experimental results more realistic, the actual distance is used instead of the straight-line distance between the nodes.

To make the matrix better reflect the correlation between nodes, it is necessary to preprocess the spatial distance between the rural logistics nodes shown in Table 2 and Appendix A and obtain the matrix as shown in Table 3 and Appendix A.

2) ACQUISITION OF THE SIMILARITY MATRIX

By preprocessing the distance matrix, the following similarity matrix is obtained in Table 3.

The matrix data shown in Table 3 and Appendix are the input matrix data of AP clustering. The matrix element values reflect the "correlation" of the logistics nodes: the larger the value, the higher the correlation.

3) ANALYSIS AND DISCUSSION OF CLUSTERING RESULTS

In this paper, Python 3.7.3 is used to calculate the results. The AP clustering algorithm flow is used to process the data, setting the maximum number of iterations of AP clustering 'max_iter'= 500, the damping factor 'damping'= 0.5, and the convergence coefficient 'convergence_iner'= 15. A CPU with an Intel (R) Core(TM) i5-5200U is used with a Windows system of 2.20 G. The results show that the original 31 township logistics nodes are divided into 8 clusters, and the clustering results are shown in Fig. 6, which results in the location of 8 logistics centres in the rural logistics system shown in Fig. 7.

The location of the eight logistics centres in the rural logistics system can be seen in Fig. 8, which are Bashan town (Node 4), Miaoba town (Node 9), Mingtong town (Node 11), Xianyi town (Node 13), Beiping township (Node 16), Xiuqi town (Node 17), Longtian township (Node 25), and Huang'an township (Node 28).

C. OPTIMIZATION OF JOINT DISTRIBUTION CENTRE LOCATION

In the first stage, 8 source transfer stations were selected by the AP clustering algorithm in Chengkou County, Chongqing. The second stage uses the fruit fly algorithm to obtain joint distribution centres.

TABLE 2. Spatial distance between township logistics nodes 1.

Nodes	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8
Node 1	0	22.9	32.2	9.3	27.3	37.5	66.7	37.1
Node 2	22.9	0	9.9	13.7	19	29.2	58.5	28.9
Node 3	32.2	9.9	0	23	28.4	36.5	65.7	36.1
Node 4	9.3	13.7	23	0	18.1	28.3	56.7	27.9
Node 5	27.3	19	28.4	18.1	0	18.5	47.7	18.1
Node 6	37.5	29.2	36.5	28.3	18.5	0	44.8	15.1
Node 7	66.7	58.5	65.7	56.7	47.7	44.8	0	33.2
Node 8	37.1	28.9	36.1	27.9	18.1	15.1	33.2	0
Node 9	47.6	39.4	46.6	38.5	28.6	25.7	20.6	12
Node 10	103.1	95.6	104.2	94.6	84.2	81.8	76.1	67.6
Node 11	91	82.9	90.1	82	72.1	69.2	64	55.5
Node 12	102.9	94.6	101.8	93.7	83.8	80.9	75.8	69.4
Node 13	109.7	101.6	108.8	100.6	90.8	87.8	82.7	76.3
Node 14	78.7	70.6	77.8	69.7	59.8	56.9	51.7	63.2
Node 15	94.2	85.8	95.3	84.9	75.2	72	76.5	60.6
Node 16	75.2	66.9	74.2	66	56.3	53.2	57.5	61.6
Node 17	71.6	64.5	72.7	63.6	52.7	50.8	53.9	38
Node 18	75.2	67	76.4	66.1	56.3	53.3	57.5	41.7
Node 19	81	72.8	82.1	71.9	62.1	59.1	63.3	47.4
Node 20	84.1	75.9	85.2	74.9	65.1	62.1	66.4	50.5
Node 21	92.8	84.7	91.9	83.6	73.9	70.8	65.9	59.5
Node 22	102.9	94.8	104.1	93.9	84	81.1	85.2	69.4
Node 23	89.3	81.1	90.4	80.1	70.4	67.3	71.6	55.7
Node 24	113.2	104.9	114.3	104	94.2	91.2	95.4	79.6
Node 25	58.3	50.2	57.4	49.2	39.4	36.3	40.6	24.7
Node 26	57.2	49	58.3	48	38.3	35	39.2	21.7
Node 27	88.5	80.2	89.7	79.3	69.5	66.5	70.7	54.8
Node 28	108.7	100.5	110	99.5	89.7	86.6	90.8	74.9
Node 29	24.4	26.3	35.6	15.3	30.6	40.9	48.9	30.2
Node 30	47.1	38.9	46.2	38	28.2	25.2	29.4	11.6
Node 31	67.1	58.7	66	57.8	48	45	49.2	33.4

TABLE 3. Preprocessed distance data between township logistics nodes.

Nodes	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8
Node 1	0.3701	0.0617	0.0441	0.1473	0.0519	0.0380	0.0215	0.0384
Node 2	0.0617	0.3701	0.1388	0.1016	0.0740	0.0486	0.0245	0.0491
Node 3	0.0441	0.1388	0.3701	0.0614	0.0499	0.0390	0.0218	0.0394
Node 4	0.1473	0.1016	0.0614	0.3701	0.0776	0.0501	0.0252	0.0508
Node 5	0.0519	0.0740	0.0499	0.0776	0.3701	0.0759	0.0299	0.0776
Node 6	0.0380	0.0486	0.0390	0.0501	0.0759	0.3701	0.0318	0.0925
Node 7	0.0215	0.0245	0.0218	0.0252	0.0299	0.0318	0.3701	0.0428
Node 8	0.0384	0.0491	0.0394	0.0508	0.0776	0.0925	0.0428	0.3701
Node 9	0.0300	0.0362	0.0306	0.0370	0.0496	0.0551	0.0684	0.1155
Node 10	0.0139	0.0150	0.0138	0.0152	0.0170	0.0175	0.0188	0.0212
Node 11	0.0158	0.0173	0.0159	0.0175	0.0199	0.0207	0.0224	0.0258
Node 12	0.0140	0.0152	0.0141	0.0153	0.0171	0.0177	0.0189	0.0206
Node 13	0.0131	0.0141	0.0132	0.0143	0.0158	0.0163	0.0173	0.0188
Node 14	0.0182	0.0203	0.0184	0.0206	0.0239	0.0251	0.0276	0.0226
Node 15	0.0152	0.0167	0.0151	0.0169	0.0191	0.0199	0.0187	0.0236
Node 16	0.0191	0.0214	0.0193	0.0217	0.0254	0.0269	0.0249	0.0232
Node 17	0.0200	0.0222	0.0197	0.0225	0.0271	0.0281	0.0265	0.0375
Node 18	0.0191	0.0214	0.0188	0.0217	0.0254	0.0268	0.0249	0.0342
Node 19	0.0177	0.0197	0.0175	0.0199	0.0230	0.0242	0.0226	0.0301
Node 20	0.0171	0.0189	0.0168	0.0191	0.0220	0.0230	0.0216	0.0283
Node 21	0.0155	0.0169	0.0156	0.0172	0.0194	0.0202	0.0217	0.0240
Node 22	0.0140	0.0151	0.0138	0.0153	0.0171	0.0177	0.0168	0.0206
Node 23	0.0161	0.0177	0.0159	0.0179	0.0203	0.0213	0.0200	0.0257
Node 24	0.0127	0.0137	0.0126	0.0138	0.0152	0.0157	0.0150	0.0180
Node 25	0.0245	0.0285	0.0249	0.0290	0.0362	0.0392	0.0351	0.0573
Node 26	0.0250	0.0291	0.0245	0.0297	0.0372	0.0406	0.0363	0.0650
Node 27	0.0162	0.0179	0.0160	0.0181	0.0206	0.0215	0.0203	0.0261
Node 28	0.0132	0.0143	0.0131	0.0144	0.0160	0.0166	0.0158	0.0191
Node 29	0.0579	0.0538	0.0400	0.0913	0.0464	0.0348	0.0292	0.0470
Node 30	0.0303	0.0366	0.0309	0.0375	0.0503	0.0561	0.0483	0.1193
Node31	0.0213	0.0244	0.0217	0.0247	0.0297	0.0317	0.0290	0.0426

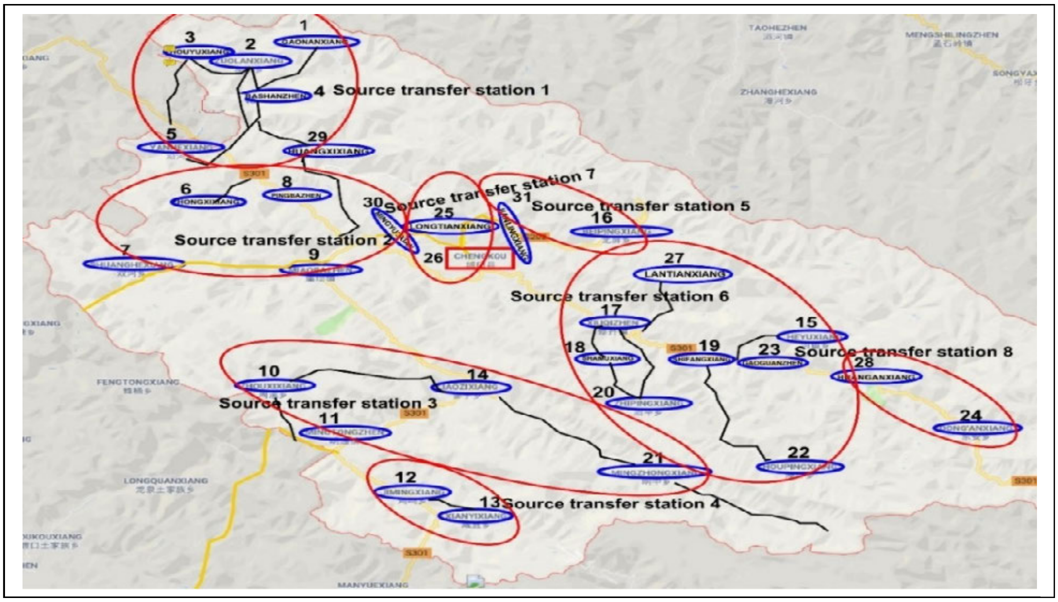

FIGURE 7. Clustering results of township logistics nodes based on the AP clustering algorithm.

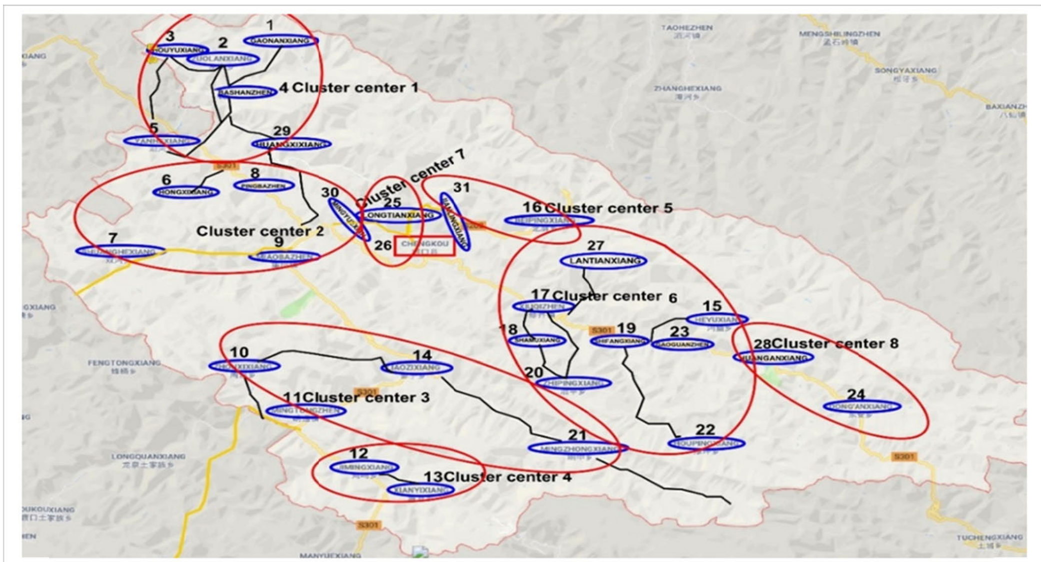

FIGURE 8. Location results of the source transfer station based on the AP clustering algorithm.

TABLE 4. Data and parameter settings.

Total number of township logistics nodes	n =31
Number of source transfer stations	A =8
Number of joint distribution centres	$B = AS/D$ =3
Number of demand points	C=7
Required area for a single distribution centre	L =300 (m2)
Individual distribution centre facility and operating costs per year	C_w=550 (yuan/ m2)
Maximum capacity of the agricultural product distribution centre	D=22000 (t)
Fixed life of each agricultural product distribution centre	t=25
Average annual total supply in the region	AS= 66000 (t)
Agricultural product corruption rate	θ=0.01
Average selling price of agricultural products	p=20 (yuan/kg)
Unit cost per delivery vehicle	h=1.5 (yuan/t*km)

TABLE 5. Distance between alternative distribution centres Unit: km.

Alternative distribution centres	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_1	0	38.5	82	100.6	66	63.6	49.2	99.5
X_2	38.5	0	43.6	62.3	38.4	34.8	21.4	71.9
X_3	82	43.6	0	18.7	81.8	79.4	65	115.3
X_4	100.6	62.3	18.7	0	100.5	96.9	83.6	134
X_5	66	38.4	81.8	100.5	0	36.6	17.6	73.7
X_6	63.6	34.8	79.4	96.9	36.6	0	20.8	38.2
X_7	49.2	21.4	65	83.6	17.6	20.8	0	56.8
X_8	99.5	71.9	115.3	134	73.7	38.2	56.8	0

TABLE 6. Distance between alternative distribution centres and demand points Unit: km.

Alternative distribution centres	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
X_1	48	274	198.2	79.3	226.3	199.7	279.2
X_2	20.4	349.4	206.7	218.9	131.3	92	176.2
X_3	63.9	249.2	114.3	115	176.2	161.3	171.4
X_4	82.5	241.4	106.5	133.7	170.7	176.1	163.5
X_5	21.4	248.2	196.2	107.8	226.2	119.2	253.3
X_6	17.4	211.7	193.8	105.4	223.8	197.2	250.9
X_7	4.7	231.3	179.4	90.9	209.3	182.7	236.2
X_8	53.1	179	229.6	141.8	259.5	233	286.7

TABLE 7. Annual supply of agricultural products at various supply points Unit: tons.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Supply	8956	9505	6985	7580	6350	9710	9980	6980

TABLE 8. Unit land price, yearly construction cost and capacity limit of alternatives.

Alternative distribution centre	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Unit land price (m2)	2750	6382	3441	3500	4421	5892	7852	2950
Average construction cost (year)	33000	76584	41292	42000	53052	70704	94224	35400

TABLE 9. Simplified name of the demand points.

Simplified name	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
Original name	Chengkou County	Wuxi County	Kaizhou	Wanyuan City	Kaijiang County	Dazhou	Wanzhou

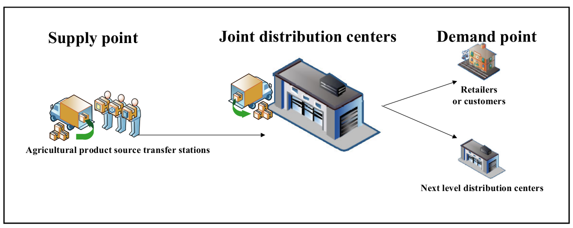

FIGURE 9. Agricultural product uplink distribution route diagram.

First, 8 supply source transfer stations will be used as the supply points and candidate points of the joint distribution centres, and the centres of neighbouring counties and the next-level distribution centre will be used as the demand points (a total of seven demand points) to construct the joint distribution centre location model. The distribution path diagram is shown in Fig. 9.

The following data information is obtained. The set values of the relevant parameters in the model are shown in Table 4. The distance between the alternative distribution centres is shown in Table 5. The distance between the alternative distribution centres and demand points is shown in Table 6. The annual supply of agricultural products at various supply points is shown in Table 7. The unit land price, average annual construction cost and capacity limit of each alternative distribution centre are shown in Table 8. The source transfer stations of Bashan town, Miaoba town, Mingtong town, Xianyi town, Beiping township, Xiuqi town, Longtian township and Huang'an township, are marked as X_1,X_2, X_3,X_4, X_5,X_6, X_7, and X_8, respectively. The seven demand points of Chengkou County, Wuxi County, Kaizhou, Wanyuan city, Kaijiang County, Dazhou, and Wanzhou, are marked as Y_1, Y_2,Y_3, Y_4,Y_5, Y_6, and Y_7. They are shown in Table 9.

VII. RESULTS AND DISCUSSION

Three of eight source transfer stations need to be chosen as the logistics distribution centres. Setting the fruit fly population

TABLE 10. Final location and distribution path from source transfer stations to joint distribution centres.

Final location	Distribution path	From source transfer stations to joint distribution centres
X_4(Xianyi town)	←---Delivery----	X_3(Mingtong town)
X_6(Xiuqi town)	←---Delivery----	X_2(Miaoba town), X_7 (Longtian township)
X_8(Huang'an township)	←---Delivery----	X_1(Bashan town), X_5(Beiping township)

TABLE 11. Distribution path from joint distribution centres to demand points.

Final location	Distribution path	From joint distribution centres to demand points
X_4(Xianyi town)	----Delivery---→	Y_2(Wuxi County), Y_3(Kaizhou), Y_6(Dazhou)
X_6(Xiuqi town)	----Delivery---→	Y_5(Kaijiang County)
X_8(Huang'an township)	----Delivery---→	Y_1(Chengkou County), Y_4(Wanyuan city), Y_7(Wanzhou)

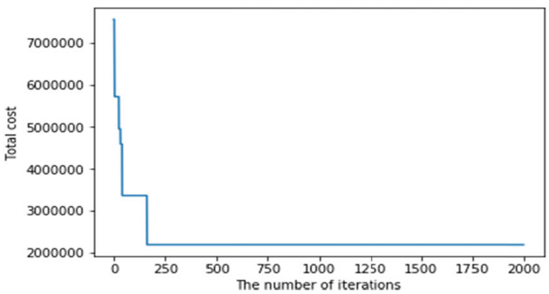

FIGURE 10. The optimization result of 2000 iterations based on FOA.

size Sizepop = 20, Fig. 10 shows the optimization result of 2000 iterations of the algorithm, and Fig. 11 shows the total cost and distribution routing. Table 8 shows the final location and dist Problems and Policy Suggestions of ribution path from the source transfer stations to the joint distribution centres. Table 10 shows the distribution path from the joint distribution centres to the demand points.

As shown in Fig. 9, the number of iterations is 2000, and the total cost of the remaining joint distribution centre locations has reached the optimal value of 2,190,402 yuan, or approximately 2.2 million yuan. Based on the geographical factors, the final iteration result of this paper is the final plan of site selection, namely, Xianyi town, Xiuqi town and Huang'an township, which are the most preferred site centres. The agricultural products of Mingtong town are distributed to Xianyi town, the agricultural products of Longtian township and Miaoba town are distributed to Xiuqi town, and the agricultural products of Bashan town and Beiping township are distributed to Huang'an township. The agricultural products of Xianyi town are distributed to Wuxi County, Kaizhou and Dazhou. The agricultural products of Xiuqi town are distributed to Kaijiang County. The agricultural products of Huang'an township are distributed to Chengkou County, Wanyuan city and Wanzhou. The details are shown in Table 10 and Table 11.

Through the study, it is found that the final joint distribution centers can meet the conditions of close distance to the county center and convenient transportation. It shows that the logistics network optimization for online sales of remote rural agricultural products designed by this way is feasible, and proves the theoretical and practical significance of the paper. □

```
Distribution routing:
 [[0.  0.  0.  0.  0.  0.  0.  1.]
 [0.  0.  0.  0.  0.  1.  0.  0.]
 [0.  0.  0.  1.  0.  0.  0.  0.]
 [0.  0.  0.  1.  0.  0.  0.  0.]
 [0.  0.  0.  0.  0.  0.  0.  1.]
 [0.  0.  0.  0.  0.  1.  0.  0.]
 [0.  0.  0.  0.  0.  1.  0.  0.]
 [0.  0.  0.  0.  0.  0.  0.  1.]]

 [[0.  0.  0.  0.  0.  0.  0.  1.]
 [0.  0.  0.  1.  0.  0.  0.  0.]
 [0.  0.  0.  1.  0.  0.  0.  0.]
 [0.  0.  0.  0.  0.  0.  0.  1.]
 [0.  0.  0.  0.  0.  1.  0.  0.]
 [0.  0.  0.  1.  0.  0.  0.  0.]
 [0.  0.  0.  0.  0.  0.  0.  1.]
 [0.  0.  0.  0.  0.  0.  0.  0.]]

Total cost:
 2190402.5204715296
```

FIGURE 11. The optimization result of 500 iterations based on FOA.

VIII. CONCLUSION

This study found that final joint distribution centres can meet the conditions of close distance to a county centre and con-

TABLE 12. Spatial distance between township logistics nodes 1.

Nodes	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8	Node 9	Node 10
Node 1	0	22.9	32.2	9.3	27.3	37.5	66.7	37.1	47.6	103.1
Node 2	22.9	0	9.9	13.7	19	29.2	58.5	28.9	39.4	95.6
Node 3	32.2	9.9	0	23	28.4	36.5	65.7	36.1	46.6	104.2
Node 4	9.3	13.7	23	0	18.1	28.3	56.7	27.9	38.5	94.6
Node 5	27.3	19	28.4	18.1	0	18.5	47.7	18.1	28.6	84.2
Node 6	37.5	29.2	36.5	28.3	18.5	0	44.8	15.1	25.7	81.8
Node 7	66.7	58.5	65.7	56.7	47.7	44.8	0	33.2	20.6	76.1
Node 8	37.1	28.9	36.1	27.9	18.1	15.1	33.2	0	12	67.6
Node 9	47.6	39.4	46.6	38.5	28.6	25.7	20.6	12	0	55.6
Node 10	103.1	95.6	104.2	94.6	84.2	81.8	76.1	67.6	55.6	0
Node 11	91	82.9	90.1	82	72.1	69.2	64	55.5	43.6	12.7
Node 12	102.9	94.6	101.8	93.7	83.8	80.9	75.8	69.4	55.3	20.5
Node 13	109.7	101.6	108.8	100.6	90.8	87.8	82.7	76.3	62.3	27.4
Node 14	78.7	70.6	77.8	69.7	59.8	56.9	51.7	63.2	31.3	24.9
Node 15	94.2	85.8	95.3	84.9	75.2	72	76.5	60.6	57.4	113.5
Node 16	75.2	66.9	74.2	66	56.3	53.2	57.5	61.6	38.4	94.5
Node 17	71.6	64.5	72.7	63.6	52.7	50.8	53.9	38	34.8	90.9
Node 18	75.2	67	76.4	66.1	56.3	53.3	57.5	41.7	38.4	94.6
Node 19	81	72.8	82.1	71.9	62.1	59.1	63.3	47.4	44.2	100.3
Node 20	84.1	75.9	85.2	74.9	65.1	62.1	66.4	50.5	47.3	103.4
Node 21	92.8	84.7	91.9	83.6	73.9	70.8	65.9	59.5	45.4	47.3
Node 22	102.9	94.8	104.1	93.9	84	81.1	85.2	69.4	66.1	122.3
Node 23	89.3	81.1	90.4	80.1	70.4	67.3	71.6	55.7	52.5	108
Node 24	113.2	104.9	114.3	104	94.2	91.2	95.4	79.6	76.3	131.9
Node 25	58.3	50.2	57.4	49.2	39.4	36.3	40.6	24.7	21.4	77.6
Node 26	57.2	49	58.3	48	38.3	35	39.2	21.7	20.4	76.5
Node 27	88.5	80.2	89.7	79.3	69.5	66.5	70.7	54.8	51.6	107.8
Node 28	108.7	100.5	110	99.5	89.7	86.6	90.8	74.9	71.9	128
Node 29	24.4	26.3	35.6	15.3	30.6	40.9	48.9	30.2	29.9	86
Node 30	47.1	38.9	46.2	38	28.2	25.2	29.4	11.6	10.3	66.4
Node 31	67.1	58.7	66	57.8	48	45	49.2	33.4	30.1	86.3

venient transportation. The model shows that a logistics network optimization for online sales of remote rural agricultural products designed in this manner is successful. Rural logistics construction has significant importance and necessity for rural revitalization. The logistics of online sales of agricultural products is a key part of rural logistics construction. For remote villages with relatively inadequate infrastructure, it is extremely important to address the problem of online sales of agricultural products. Based on the background of rural revitalization, this paper analyses the problem of agricultural products in remote rural areas, combining the characteristics of agricultural products and the logistics characteristics in the Internet era and remote rural areas. This paper proposes a logistics distribution model for agricultural products in remote villages. On the basis of fully considering the distribution characteristics of agricultural products in remote

TABLE 13. Spatial distance between township logistics nodes 2.

Nodes	Node 11	Node 12	Node 13	Node 14	Node 15	Node 16	Node 17	Node 18	Node 19	Node 20
Node 1	91	102.9	109.7	78.7	94.2	75.2	71.6	75.2	81	84.1
Node 2	82.9	94.6	101.6	70.6	85.8	66.9	64.5	67	72.8	75.9
Node 3	90.1	101.8	108.8	77.8	95.3	74.2	72.7	76.4	82.1	85.2
Node 4	82	93.7	100.6	69.7	84.9	66	63.6	66.1	71.9	74.9
Node 5	72.1	83.8	90.8	59.8	75.2	56.3	52.7	56.3	62.1	65.1
Node 6	69.2	80.9	87.8	56.9	72	53.2	50.8	53.3	59.1	62.1
Node 7	64	75.8	82.7	51.7	76.5	57.5	53.9	57.5	63.3	66.4
Node 8	55.5	69.4	76.3	63.2	60.6	61.6	38	41.7	47.4	50.5
Node 9	43.6	55.3	62.3	31.3	57.4	38.4	34.8	38.4	44.2	47.3
Node 10	12.7	20.5	27.4	24.9	113.5	94.5	90.9	94.6	100.3	103.4
Node 11	0	11.7	18.7	12.3	100.7	81.8	79.4	81.9	87.7	90.7
Node 12	11.7	0	7.2	24	112.6	93.6	90	93.6	99.4	102.5
Node 13	18.7	7.2	0	31	119.5	100.5	96.9	100.6	106.3	109.4
Node 14	12.3	24	31	0	88.5	69.6	66	69.6	75.4	78.4
Node 15	100.7	112.6	119.5	88.5	0	59	22.7	25	13.1	33.9
Node 16	81.8	93.6	100.5	69.6	59	0	36.6	40.2	46	49.1
Node 17	79.4	90	96.9	66	22.7	36.6	0	3.8	9.5	12.6
Node 18	81.9	93.6	100.6	69.6	25	40.2	3.8	0	12.1	8.8
Node 19	87.7	99.4	106.3	75.4	13.1	46	9.5	12.1	0	20.9
Node 20	90.7	102.5	109.4	78.4	33.9	49.1	12.6	8.8	20.9	0
Node 21	34.9	46.4	53.3	22.4	102.5	83.6	81.2	83.7	89.5	92.6
Node 22	109.7	121.4	128.3	97.3	31.2	67.9	31.5	34	22	42.8
Node 23	95.9	107.7	114.6	83.7	4.8	54.3	17.8	20.4	8.3	29.2
Node 24	119.8	131.5	138.5	107.5	24.9	78.1	41.8	44.4	32.1	53
Node 25	65	76.7	83.6	52.7	42.1	17.6	20.8	23.4	29.1	32.2
Node 26	63.9	75.6	82.5	51.6	38.6	21.4	17.4	19.9	25.8	28.8
Node 27	95.1	106.8	113.8	82.8	33.9	53.4	18.3	20.8	21.1	28.9
Node 28	115.3	127.1	134	103.1	20.3	73.7	38.2	39.7	27.5	48.4
Node 29	73.4	85.1	92	61.1	76.2	57.4	55	57.5	63.2	63.3
Node 30	53.8	65.5	72.4	41.5	46.9	28	25.6	28.2	33.9	37
Node 31	73.6	85.3	92.3	61.3	50.7	8.2	29.4	32	37.7	40.8

rural areas and the principle and influencing factors of site selection, logistics network optimization is carried out for the first kilometre of remote villages. The first stage clusters the township logistics nodes to obtain alternative joint distribution centres. The second stage builds a joint distribution centre location model for remote rural agricultural products based on the existing agricultural product distribution centre location model. Combined with specific cases, the AP clustering algorithm is used to cluster the township logistics nodes to obtain the source transfer stations. Finally, the source transfer stations are regarded as the alternative distribution centres, and the fruit fly optimization algorithm and MATLAB software are used to solve the problem. The joint distribution centre location model for remote rural agricultural products has a certain reference value for the same problem in other remote areas.

TABLE 14. Spatial distance between township logistics nodes 3.

Nodes	Node 21	Node 22	Node 23	Node 24	Node 25	Node 26	Node 27	Node 28	Node 29	Node 30	Node 31
Node 1	92.8	102.9	89.3	113.2	58.3	57.2	88.5	108.7	24.4	47.1	67.1
Node 2	84.7	94.8	81.1	104.9	50.2	49	80.2	100.5	26.3	38.9	58.7
Node 3	91.9	104.1	90.4	114.3	57.4	58.3	89.7	110	35.6	46.2	66
Node 4	83.6	93.9	80.1	104	49.2	48	79.3	99.5	15.3	38	57.8
Node 5	73.9	84	70.4	94.2	39.4	38.3	69.5	89.7	30.6	28.2	48
Node 6	70.8	81.1	67.3	91.2	36.3	35	66.5	86.6	40.9	25.2	45
Node 7	65.9	85.2	71.6	95.4	40.6	39.2	70.7	90.8	48.9	29.4	49.2
Node 8	59.5	69.4	55.7	79.6	24.7	21.7	54.8	74.9	30.2	11.6	33.4
Node 9	45.4	66.1	52.5	76.3	21.4	20.4	51.6	71.9	29.9	10.3	30.1
Node 10	47.3	122.3	108	131.9	77.6	76.5	107.8	128	86	66.4	86.3
Node 11	34.9	109.7	95.9	119.8	65	63.9	95.1	115.3	73.4	53.8	73.6
Node 12	46.4	121.4	107.7	131.5	76.7	75.6	106.8	127.1	85.1	65.5	85.3
Node 13	53.3	128.3	114.6	138.5	83.6	82.5	113.8	134	92	72.4	92.3
Node 14	22.4	97.3	83.7	107.5	52.7	51.6	82.8	103.1	61.1	41.5	61.3
Node 15	102.5	31.2	4.8	24.9	42.1	38.6	33.9	20.3	76.2	46.9	50.7
Node 16	83.6	67.9	54.3	78.1	17.6	21.4	53.4	73.7	57.4	28	8.2
Node 17	81.2	31.5	17.8	41.8	20.8	17.4	18.3	38.2	55	25.6	29.4
Node 18	83.7	34	20.4	44.4	23.4	19.9	20.8	39.7	57.5	28.2	32
Node 19	89.5	22	8.3	32.1	29.1	25.8	21.1	27.5	63.2	33.9	37.7
Node 20	92.6	42.8	29.2	53	32.2	28.8	28.9	48.4	63.3	37	40.8
Node 21	0	111.2	97.6	121.6	66.2	65.4	96.9	117	75.1	55.4	75.2
Node 22	111.2	0	26.3	50.3	51.1	47.7	43.1	45.7	85.2	55.9	59.8
Node 23	97.6	26.3	0	23.9	37.4	34.1	29.4	19.4	71.6	42.1	45.9
Node 24	121.6	50.3	23.9	0	61.4	58.1	53.5	9.2	95.6	66	69.9
Node 25	66.2	51.1	37.4	61.4	0	4.7	36.6	56.8	40.5	11.2	9.4
Node 26	65.4	47.7	34.1	58.1	4.7	0	33.4	53.1	39.4	9.8	13.2
Node 27	96.9	43.1	29.4	53.5	36.6	33.4	0	48.4	70.7	41.3	45.1
Node 28	117	45.7	19.4	9.2	56.8	53.1	48.4	0	90.7	61.4	65.2
Node 29	75.1	85.2	71.6	95.6	40.5	39.4	70.7	90.7	0	29.4	49.2
Node 30	55.4	55.9	42.1	66	11.2	9.8	41.3	61.4	29.4	0	19.8
Node 31	75.2	59.8	45.9	69.9	9.4	13.2	45.1	65.2	49.2	19.8	0

There are still some limitations in this paper, which need to be improved. First, this paper lacks spot investigation and expert consultation. The main considerations are economic benefits in modeling, but without considering the natural conditions, the calculated location results may not be the final site. In the future research process, it is necessary to use a variety of methods and consider a variety of factors to conduct a more comprehensive analysis of the location problem. Second, Fruit Fly Optimization Algorithm has the disadvantages of easy to fall into premature convergence and low precision. In the future research, if Fruit Fly Optimization Algorithm is used to study related problems, it is necessary to improve Fruit Fly Optimization Algorithm by combining various heuristic algorithms.

TABLE 15. Preprocessed distance data between township logistics nodes 1.

Node	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8	Node 9	Node 10
Node 1	0.37005	0.061663	0.044123	0.147342	0.051901	0.037968	0.021469	0.038372	0.0299947	0.0139257
Node 2	0.061663	0.37005	0.138828	0.10164	0.074001	0.04858	0.024453	0.049076	0.0361597	0.0150126
Node 3	0.044123	0.138828	0.37005	0.061401	0.049925	0.038994	0.021793	0.03942	0.0306316	0.0137794
Node 4	0.147342	0.10164	0.061401	0.37005	0.077583	0.050099	0.025223	0.050804	0.0369942	0.0151704
Node 5	0.051901	0.074001	0.049925	0.077583	0.37005	0.075949	0.029933	0.077583	0.049582	0.0170332
Node 6	0.037968	0.04858	0.038994	0.050099	0.075949	0.37005	0.031849	0.092512	0.0550714	0.0175299
Node 7	0.021469	0.024453	0.021793	0.025223	0.029933	0.031849	0.37005	0.042813	0.068387	0.0188344
Node 8	0.038372	0.049076	0.03942	0.050804	0.077583	0.092512	0.042813	0.37005	0.1154772	0.0211853
Node 9	0.029995	0.03616	0.030632	0.036994	0.049582	0.055071	0.068387	0.115477	0.37005	0.0257172
Node 10	0.013926	0.015013	0.013779	0.01517	0.017033	0.01753	0.018834	0.021185	0.0257172	0.37005
Node 11	0.015767	0.017299	0.015924	0.017487	0.019872	0.020699	0.022368	0.025763	0.0327156	0.1093474
Node 12	0.013953	0.01517	0.014103	0.015315	0.017114	0.017724	0.018908	0.02064	0.0258554	0.0687128
Node 13	0.013092	0.01413	0.013199	0.01427	0.015802	0.016339	0.01734	0.018785	0.0229733	0.051715
Node 14	0.018216	0.020291	0.018425	0.020552	0.023926	0.025135	0.027639	0.022649	0.0453715	0.0568064
Node 15	0.015235	0.016717	0.01506	0.016894	0.019058	0.0199	0.018737	0.023613	0.0249176	0.0126553
Node 16	0.019058	0.021405	0.019313	0.021695	0.0254	0.026867	0.024875	0.023232	0.0370893	0.0151864
Node 17	0.02001	0.022196	0.019709	0.022507	0.027119	0.028124	0.026521	0.037475	0.0408723	0.0157846
Node 18	0.019058	0.021374	0.018761	0.021662	0.0254	0.026817	0.024875	0.034189	0.0370893	0.0151704
Node 19	0.017702	0.019682	0.017466	0.019927	0.023047	0.024207	0.022613	0.03012	0.0322764	0.0143126
Node 20	0.017053	0.018884	0.016834	0.019134	0.021993	0.023047	0.021565	0.028289	0.030183	0.0138855
Node 21	0.015463	0.016933	0.015614	0.017155	0.019391	0.020234	0.021728	0.024045	0.0314325	0.030183
Node 22	0.013953	0.015139	0.013793	0.015283	0.017074	0.01768	0.016834	0.02064	0.0216625	0.0117484
Node 23	0.016066	0.01768	0.015871	0.0179	0.020349	0.021279	0.02001	0.025671	0.0272215	0.0132968
Node 24	0.012689	0.013688	0.012567	0.013806	0.015235	0.015733	0.015044	0.018011	0.0187854	0.0108965
Node 25	0.024536	0.028456	0.024918	0.029029	0.03616	0.039206	0.035104	0.057257	0.0658879	0.0184727
Node 26	0.025004	0.029146	0.024536	0.029747	0.037185	0.040642	0.036342	0.064997	0.0690416	0.0187366
Node 27	0.01621	0.017877	0.015995	0.018079	0.02061	0.021533	0.020263	0.026089	0.0276917	0.0133214
Node 28	0.013212	0.014284	0.013056	0.014427	0.015995	0.016564	0.015802	0.019134	0.019927	0.0112273
Node 29	0.057947	0.053838	0.039966	0.09134	0.046393	0.034849	0.029205	0.046997	0.0474614	0.0166787
Node 30	0.03031	0.036619	0.030894	0.037475	0.050273	0.056143	0.048255	0.119299	0.1336784	0.0215653
Node 31	0.021342	0.02437	0.021695	0.024747	0.029747	0.031709	0.029029	0.04256	0.0471511	0.0166211

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

APPENDIX A

See Table 12–17.

APPENDIX B

Affinity Propagation Clustering Main program

```
from sklearn.cluster import AffinityPropagation
from sklearn import metrics
from sklearn.datasets.samples_generator import
make_blobs
import numpy as np
import xlrd
# Define input matrix function
def excel(path):
```

TABLE 16. Preprocessed distance data between township logistics nodes 2.

Node	Node 11	Node 12	Node 13	Node 14	Node 15	Node 16	Node 17	Node 18	Node 19	Node 20	Node 21
Node 1	0.0158	0.0140	0.0131	0.0182	0.0152	0.0191	0.0200	0.0191	0.0177	0.0171	0.0155
Node 2	0.0173	0.0152	0.0141	0.0203	0.0167	0.0214	0.0222	0.0214	0.0197	0.0189	0.0169
Node 3	0.0159	0.0141	0.0132	0.0184	0.0151	0.0193	0.0197	0.0188	0.0175	0.0168	0.0156
Node 4	0.0175	0.0153	0.0143	0.0206	0.0169	0.0217	0.0225	0.0217	0.0199	0.0191	0.0172
Node 5	0.0199	0.0171	0.0158	0.0239	0.0191	0.0254	0.0271	0.0254	0.0230	0.0220	0.0194
Node 6	0.0207	0.0177	0.0163	0.0251	0.0199	0.0269	0.0281	0.0268	0.0242	0.0230	0.0202
Node 7	0.0224	0.0189	0.0173	0.0276	0.0187	0.0249	0.0265	0.0249	0.0226	0.0216	0.0217
Node 8	0.0258	0.0206	0.0188	0.0226	0.0236	0.0232	0.0375	0.0342	0.0301	0.0283	0.0240
Node 9	0.0327	0.0259	0.0230	0.0454	0.0249	0.0371	0.0409	0.0371	0.0323	0.0302	0.0314
Node 10	0.1093	0.0687	0.0517	0.0568	0.0127	0.0152	0.0158	0.0152	0.0143	0.0139	0.0302
Node 11	0.3701	0.1183	0.0752	0.1128	0.0143	0.0175	0.0181	0.0175	0.0164	0.0158	0.0408
Node 12	0.1183	0.3701	0.1876	0.0589	0.0128	0.0153	0.0159	0.0153	0.0144	0.0140	0.0308
Node 13	0.0752	0.1876	0.3701	0.0458	0.0120	0.0143	0.0148	0.0143	0.0135	0.0131	0.0268
Node 14	0.1128	0.0589	0.0458	0.3701	0.0162	0.0206	0.0217	0.0206	0.0190	0.0183	0.0630
Node 15	0.0143	0.0128	0.0120	0.0162	0.3701	0.0242	0.0622	0.0566	0.1061	0.0419	0.0140
Node 16	0.0175	0.0153	0.0143	0.0206	0.0242	0.3701	0.0389	0.0354	0.0310	0.0291	0.0172
Node 17	0.0181	0.0159	0.0148	0.0217	0.0622	0.0389	0.3701	0.3370	0.1444	0.1102	0.0177
Node 18	0.0175	0.0153	0.0143	0.0206	0.0566	0.0354	0.3370	0.3701	0.1146	0.1553	0.0171
Node 19	0.0164	0.0144	0.0135	0.0190	0.1061	0.0310	0.1444	0.1146	0.3701	0.0674	0.0160
Node 20	0.0158	0.0140	0.0131	0.0183	0.0419	0.0291	0.1102	0.1553	0.0674	0.3701	0.0155
Node 21	0.0408	0.0308	0.0268	0.0630	0.0140	0.0172	0.0177	0.0171	0.0160	0.0155	0.3701
Node 22	0.0131	0.0118	0.0112	0.0148	0.0455	0.0211	0.0451	0.0418	0.0641	0.0333	0.0129
Node 23	0.0150	0.0133	0.0125	0.0171	0.2730	0.0263	0.0789	0.0690	0.1641	0.0486	0.0147
Node 24	0.0120	0.0109	0.0104	0.0134	0.0568	0.0184	0.0341	0.0321	0.0443	0.0270	0.0118
Node 25	0.0220	0.0187	0.0172	0.0271	0.0339	0.0797	0.0677	0.0604	0.0487	0.0441	0.0216
Node 26	0.0224	0.0190	0.0174	0.0277	0.0369	0.0659	0.0806	0.0707	0.0549	0.0492	0.0219
Node 27	0.0151	0.0134	0.0126	0.0173	0.0419	0.0268	0.0768	0.0677	0.0668	0.0491	0.0148
Node 28	0.0125	0.0113	0.0107	0.0139	0.0694	0.0194	0.0373	0.0359	0.0515	0.0295	0.0123
Node 29	0.0195	0.0169	0.0156	0.0234	0.0188	0.0249	0.0260	0.0249	0.0226	0.0226	0.0191
Node 30	0.0266	0.0219	0.0198	0.0344	0.0304	0.0506	0.0553	0.0503	0.0419	0.0385	0.0258
Node 31	0.0195	0.0168	0.0155	0.0233	0.0282	0.1660	0.0483	0.0444	0.0378	0.0349	0.0191

```
data = xlrd.open_workbook(path)
table = data.sheets()[0]
nrows = table.nrows # Number of rows
ncols = table.ncols # Number of columns
c1 = np.arange(0, nrows, 1)
datamatrix = np.zeros((nrows, ncols))
for x in range(ncols):
    cols = table.col_values(x)
    minVals = min(cols)
    maxVals = max(cols)
    cols1 = np.matrix(cols) # Convert list to matrix for matrix operation
ranges = maxVals - minVals
    b = cols1 - minVals
```

TABLE 17. Preprocessed distance data between township logistics nodes 3.

Node	Node 22	Node 23	Node 24	Node 25	Node 26	Node 27	Node 28	Node 29	Node 30	Node 31
Node 1	0.0140	0.0161	0.0127	0.0245	0.0250	0.0162	0.0132	0.0579	0.0303	0.0213
Node 2	0.0151	0.0177	0.0137	0.0285	0.0291	0.0179	0.0143	0.0538	0.0366	0.0244
Node 3	0.0138	0.0159	0.0126	0.0249	0.0245	0.0160	0.0131	0.0400	0.0309	0.0217
Node 4	0.0153	0.0179	0.0138	0.0290	0.0297	0.0181	0.0144	0.0913	0.0375	0.0247
Node 5	0.0171	0.0203	0.0152	0.0362	0.0372	0.0206	0.0160	0.0464	0.0503	0.0297
Node 6	0.0177	0.0213	0.0157	0.0392	0.0406	0.0215	0.0166	0.0348	0.0561	0.0317
Node 7	0.0168	0.0200	0.0150	0.0351	0.0363	0.0203	0.0158	0.0292	0.0483	0.0290
Node 8	0.0206	0.0257	0.0180	0.0573	0.0650	0.0261	0.0191	0.0470	0.1193	0.0426
Node 9	0.0217	0.0272	0.0188	0.0659	0.0690	0.0277	0.0199	0.0475	0.1337	0.0472
Node 10	0.0117	0.0133	0.0109	0.0185	0.0187	0.0133	0.0112	0.0167	0.0216	0.0166
Node 11	0.0131	0.0150	0.0120	0.0220	0.0224	0.0151	0.0125	0.0195	0.0266	0.0195
Node 12	0.0118	0.0133	0.0109	0.0187	0.0190	0.0134	0.0113	0.0169	0.0219	0.0168
Node 13	0.0112	0.0125	0.0104	0.0172	0.0174	0.0126	0.0107	0.0156	0.0198	0.0155
Node 14	0.0148	0.0171	0.0134	0.0271	0.0277	0.0173	0.0139	0.0234	0.0344	0.0233
Node 15	0.0455	0.2730	0.0568	0.0339	0.0369	0.0419	0.0694	0.0188	0.0304	0.0282
Node 16	0.0211	0.0263	0.0184	0.0797	0.0659	0.0268	0.0194	0.0249	0.0506	0.1660
Node 17	0.0451	0.0789	0.0341	0.0677	0.0806	0.0768	0.0373	0.0260	0.0553	0.0483
Node 18	0.0418	0.0690	0.0321	0.0604	0.0707	0.0677	0.0359	0.0249	0.0503	0.0444
Node 19	0.0641	0.1641	0.0443	0.0487	0.0549	0.0668	0.0515	0.0226	0.0419	0.0378
Node 20	0.0333	0.0486	0.0270	0.0441	0.0492	0.0491	0.0295	0.0226	0.0385	0.0349
Node 21	0.0129	0.0147	0.0118	0.0216	0.0219	0.0148	0.0123	0.0191	0.0258	0.0191
Node 22	0.3701	0.0538	0.0284	0.0280	0.0299	0.0331	0.0312	0.0168	0.0256	0.0239
Node 23	0.0538	0.3701	0.0591	1.3701	0.0417	2.3701	0.0725	3.3701	0.0339	4.3701
Node 24	0.0284	0.0591	0.3701	0.0233	1.3701	0.0267	2.3701	0.0150	3.3701	0.0205
Node 25	0.0280	0.0381	0.0233	0.3701	0.2783	0.0389	0.0252	0.0352	0.1234	0.1459
Node 26	0.0299	0.0417	0.0246	0.2783	0.3701	0.0426	0.0269	0.0362	0.1402	0.1054
Node 27	0.0331	0.0483	0.0267	0.0389	0.0426	0.3701	0.0295	0.0203	0.0345	0.0316
Node 28	0.0312	0.0725	0.1489	0.0252	0.0269	0.0295	0.3701	0.0158	0.0233	0.0220
Node 29	0.0168	0.0200	0.0150	0.0352	0.0362	0.0203	0.0158	0.3701	0.0483	0.0290
Node 30	0.0256	0.0339	0.0217	0.1234	0.1402	0.0345	0.0233	0.0483	0.3701	0.0711
Node 31	0.0239	0.0311	0.0205	0.1459	0.1054	0.0316	0.0220	0.0290	0.0711	0.3701

```
            normcols = b / ranges # Data
            normalization
            datamatrix[:, x] = normcols # Store
            data
        return datamatrix
# Compute Affinity Propagation
Similarity = excel(''d: \ desktop \ meizi.xlsx'')
af = AffinityPropagation(preference = None,
affinity = 'precomputed').fit(Similarity)
cluster_centers_indices = af.cluster_centers_indices_
labels = af.labels_
n_clusters_ = len(cluster_centers_indices)
```

```
print("Cluster center \n",cluster_centers_indices)
print("Classification label \n",labels)
print("Number of clusters \n",n_clusters_)
# Plot result
import matplotlib.pyplot as plt
from itertools import cycle
plt.close('all')
plt.figure(1)
plt.clf()
colors = cycle('bgrcmykbgrcmykbgrcmykbgrcmyk')
for k, col in zip(range(n_clusters_), colors):
        class_members = labels == k
        cluster_center = X[cluster_centers_indices[k]]
        plt.plot(X[class_members, 0],
        X[class_members, 1], col + '.')
        plt.plot(cluster_center[0], cluster_center[1], 'o',
        markerfacecolor = col,
markeredgecolor = 'k', markersize = 14)
        for x in X[class_members]:
        plt.plot([cluster_center[0],
        x[0]], [cluster_center[1], x[1]], col)
plt.title('Estimated number of clusters: %d' % n_clusters_)
plt.savefig("AP_feed_similarity.png")
plt.show()
```

Fruit Fly Optimization Algorithm Main program

```
import xlrd
import numpy as np
import matplotlib.pyplot as plt
def generate(total, choose):
  p = float(choose) / total
  data = np.array(np.random.binomial(1, p, total))
  if np.sum(data) == choose:
    return data
  else:
    return generate(total, choose)
def excel(path,a):
  data = xlrd.open_workbook(path)
  table = data.sheets()[a]
nrows = table.nrows # the number of rows
ncols = table.ncols # the number of columns
c1 = np.arange(0, nrows, 1)
datamatrix = np.zeros((nrows, ncols))
for x in range(ncols):
  cols = table.col_values(x)
  #minVals = min(cols)
  #maxVals = max(cols)
  cols1 = np.matrix(cols) # Converting list to matrix
  for matrix operation
  #ranges = maxVals - minVals
  #b = cols1 - minVals
  #normcols = b / ranges # Data normalization
  #datamatrix[:, x] = normcols # Store data
  datamatrix[:, x] = cols1
 return datamatrix
def cf(a,b,c): # Generate a duplicate matrix of a row
  m = np.zeros([b,c])
  for i in range(b):
    m[i,:] = a
  return m
def ps(a,b,c): # Generate a random distribution
matrix based on the index
  n = [a,b,c]
  m = np.zeros([8,8])
  m[a][a] = 1
  m[b][b] = 1
  m[c][c] = 1
  for i in range(8):
    if sum(m[i,:]) == 1:
      continue
    else:
      c = np.random.choice(n)
      m[i][c] = 1
    return m
def cost_AB():
  m = 0 #Create cyclic variables of index matrix
  p = 20 # Transportation cost per unit product
  cita = 0.01 # Deterioration rate
  x_B = generate(8,3) # Generate the judgment
  matrix of whether to build a distribution center in a
  certain
  place or not
  #print(x_B)
  Con_B = [33000,76584,41292,42000,53052,70704,
  94224,35400] # Construction cost of each distribution
  center
  Con = sum(x_B*Con_B)+495000 #
  Construction cost of distribution center
    #print(x_B*Con_B)
  Dis_AB = excel("G: \\ desktop \\q_AB.xlsx",0) #
  Create a distance matrix between 8 alternative
  distribution centers
  #print(Dis_AB)
  q_AB = excel("G: \\ desktop \\q_AB.xlsx",2)
  # Create traffic between 8 alternative distribution centers
  # print(q_AB)
  yes = np.zeros([1,8]) # Create an index matrix
  m = 0 # As a loop variable for the index matrix
  for i in range(8):
    if x_B[i] == 1:
      yes[0][m] = i
    m = m +1 #At the same time, m = 2
  #print(yes)
  fa = ps(int(yes[0][0]),int(yes[0][1]),int(yes[0][2]))
  # Create whether to distribute matrix
  c = (Dis_AB*q_AB*fa.T).T # c is
  the transport scheme matrix
  b = sum(sum(c)) # Total distribution costs
  aa = q_AB*p
  bb = (1-cita)**Dis_AB
  cc = 1/bb-1
  dd = (aa*cc*fa.T).T # Damage cost matrix for
```

```
        each location
        ff = sum(sum(dd)) # Total damage cost
        gg = ff+b
        return gg,fa
    cost_AB()
    # Define the function to be solved (named fun1)
    def fun1():
        y = cost_AB()
        return y
    ####### Fruit Fly Optimization Algorithm ######
    ## Initialize the parameters
    popsize = 30 # Initialize the population size
    maxgen = 500 # Initialize the maximum number of itera-
tions
    R = 1    # the optimization step
    D = 1    # Optimize the number of variables
    Dist = np.zeros([popsize,D])
    S = np.zeros([popsize,D])
    Smell = np.zeros([popsize,1])
    X = np.zeros([popsize,D])
    Y = np.zeros([popsize,D])
    fitness = np.zeros([maxgen,1])
    # The initial position is given
    X_axis = np.random.rand(1,D)
    Y_axis = np.random.rand(1,D)
    # the search distance is given
    for i in range(popsize):
        X[i,:] = X_axis + R*(2*np.random.rand(1,D)-1)
        Y[i,:] = Y_axis + R*(2* np.random.rand(1,D)-1)
        #Calculate Dist
        Dist[i,:] = np.sqrt(X[i,:]**2+Y[i,:]**2)
        # the smell concentration judgment value is calculated,
        which is the reciprocal of distance
        S[i,:] = 1/Dist[i,:]
            # Substitute smell concentration judgment value into
            smell concentration judgment function
            (Fitness function)
            #Smell[i] = fun1(S[i,:])
        m,n = fun1()
        Smell[i] = m
    # Find out the fruit fly with optimal smell concentration
    Smellbest,index = np.min(Smell),np.argmin(Smell)
    bestSmell = Smellbest
    # Keep the best smell concentration value
    and X, Y coordinate
    X_axis = X[int(index),:]
    Y_axis = Y[int(index),:]
    plt.figure(1)
    # Start the iteration
    for j in range(maxgen):
        for i in range(popsize):
            X[i,:] = X_axis + R*(2*np.random.rand(1,D)-1)
            Y[i,:] = Y_axis + R*(2*np.random.rand(1,D)-1)
            # Calculate Dist
            Dist[i,:] = np.sqrt(X[i,:]**2+Y[i,:]**2)
            # the smell concentration judgment value is
            calculated, which is the reciprocal of distance
            S[i,:] = 1/Dist[i,:]
                # Substitute smell concentration judgment value
                into smell concentration judgment function
                (Fitness function)
            #Smell[i] = fun1(S[i,:])
            m,n = fun1()
            Smell[i] = m
        Smellbest,index = np.min(Smell),np.argmin(Smell)
        if Smellbest < bestSmell:
            bestSmell = Smellbest
            X_axis = X[int(index),:]
            Y_axis = Y[int(index),:]
        fitness[j] = bestSmell
        if j == maxgen-1:
            print("Distribution routing: \n",n)
    print("Total cost: \n" ,bestSmell)
    plt.scatter(X,Y)
    plt.figure(20)
    plt.plot(range(maxgen),fitness)
    plt.xlabel('The number of iterations')
    plt.ylabel('Total cost')
```

REFERENCES

[1] W. Weiqun and L. Zhixin, "Research on the problems and model innovation of e-commerce development of fresh agricultural products in China," *World Agricult.*, vol. 6, pp. 213–217, 2017.

[2] Y. Xue-Guo, Z. Ping, Z. Jun, and W. Di, "Development trend, problems and countermeasures for cold chain logistics industry in China," *J. Agricult. Sci. Technol.*, vol. 1, pp. 7–14, 2015.

[3] H. Song, "Realistic analysis and coping strategies of constructing fresh agricultural product supply chain in China," *Market Modernization*, vol. 8, pp. 132–133, 2007.

[4] A. Kazemi, M. H. Fazel Zarandi, and S. M. Moattar Husseini, "A multi-agent system to solve the production–distribution planning problem for a supply chain: A genetic algorithm approach," *Int. J. Adv. Manuf. Technol.*, vol. 44, nos. 1–2, pp. 180–193, Sep. 2009.

[5] G. Sheng, "Research on the development of rural e-commerce logistics in China based on long tail theory," *China Storage Transp.*, vol. 7, pp. 124–126, 2015.

[6] A. A. Kuehn and M. J. Hamburger, "A heuristic program for locating warehouses," *Manage. Sci.*, vol. 9, no. 4, pp. 643–666, Jul. 1963.

[7] C. Y. Lee, "The multiproduct warehouse location problem: Applying a decomposition algorithm," *Int. J. Phys. Distrib. Logistics Manage.*, vol. 23, no. 6, pp. 3–13, Jun. 1993.

[8] M. A. Efroymson and T. L. Ray, "A branch-bound algorithm for plant location," *Oper. Res.*, vol. 14, no. 3, pp. 361–368, Jun. 1966.

[9] J. G. Klincewicz and H. Luss, "A dual-based algorithm for multiproduct uncapacitated facility location," *Transp. Sci.*, vol. 21, no. 3, pp. 198–206, 1987.

[10] C. S. ReVelle, H. A. Eiselt, and M. S. Daskin, "A bibliography for some fundamental problem categories in discrete location science," *Eur. J. Oper. Res.*, vol. 184, no. 3, pp. 817–848, Feb. 2008.

[11] L. Xiaochun and Z. Hesheng, "Study on address selection of distributing center using barycentric method," *J. Beijing Jiaotong Univ.*, vol. 24, no. 6, pp. 108–110, 2000.

[12] M. Yang and X. Li, "Application of improved gravity method in location of logistics distribution center," *Logistics Technol.*, vol. 26, no. 6, pp. 60–62, 2007.

[13] C. H. Aikens, "Facility location models for distribution planning," *Eur. J. Oper. Res.*, vol. 22, no. 3, pp. 263–279, Dec. 1985.

[14] A. Langevin and D. Riopel, *Logistics Systems: Design and Optimization*. 2005.

[15] O. Berman, D. Krass, and M. B. C. Menezes, "Facility reliability issues in networkp-median problems: Strategic centralization and co-location effects," *Oper. Res.*, vol. 55, no. 2, pp. 332–350, Apr. 2007.

[16] T. Paksoy and N. Y. Pehlivan, "A fuzzy linear programming model for the optimization of multi-stage supply chain networks with triangular and trapezoidal membership functions," *J. Franklin Inst.*, vol. 349, no. 1, pp. 93–109, Feb. 2012.

[17] Z. Gao, K. Yoshimoto, and S. Ohmori, "Application of integrative analytic hierarchy process and 0-1 integer programming to retail stores location in logistic system," in *Proc. 4th Int. Joint Conf. Comput. Sci. Optim.*, Apr. 2011, pp. 584–586.

[18] I. Ertugrul and N. Karakasoglu, "Comparison of fuzzy AHP and fuzzy TOPSIS methods for facility location selection," *Int. J. Adv. Manuf. Technol.*, vol. 39, nos. 7–8, pp. 783–795, Nov. 2008.

[19] B. Elevli, "Logistics freight center locations decision by using fuzzy-promethee," *Transport*, vol. 29, no. 4, pp. 412–418, Dec. 2014.

[20] C.-T. Chen, "A fuzzy approach to select the location of the distribution center," *Fuzzy Sets Syst.*, vol. 118, no. 1, pp. 65–73, Feb. 2001.

[21] V. Klapita and Z. Švecová, "Logistics centers location," *Transport*, vol. 21, no. 1, pp. 48–52, Mar. 2006.

[22] Z. Turskis and E. K. Zavadskas, "A new fuzzy additive ratio assessment method (ARAS-F). Case study: The analysis of fuzzy multiple criteria in order to select the logistic centers location," *Transport*, vol. 25, no. 4, pp. 423–432, Dec. 2010.

[23] W. Xu, L. N. Liu, Q. S. Zhang, and P. Liu, "Location decision-making of equipment manufacturing enterprise under dual-channel purchase and sale mode," *Complexity*, vol. 2018, Dec. 2018, Art. no. 3797131.

[24] D. Karaboga and B. Akay, "A comparative study of artificial bee colony algorithm," *Appl. Math. Comput.*, vol. 214, no. 1, pp. 108–132, Aug. 2009.

[25] Y. Marinakis and M. Marinaki, "A particle swarm optimization algorithm with path relinking for the location routing problem," *J. Math. Model. Algorithms*, vol. 7, no. 1, pp. 59–78, Mar. 2008.

[26] M. Salehi Maleh, S. Soleymani, R. Rasouli Nezhad, and N. Ghadimi, "Using particle swarm optimization algorithm based on multi-objective function in reconfigured system for optimal placement of distributed generation," *J. Bioinf. Intell. Control*, vol. 2, no. 2, pp. 119–124, Jun. 2013.

[27] F. V. D. Bergh and A. P. Engelbrecht, "A new locally convergent particle swarm optimiser," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, Oct. 2002, p. 6.

[28] Y. Liu, J. J. Sun, and X. Wang, "Research on multi-distribution center location based on fruit fly–immune algorithm," *Appl. Mech. Mater.*, vol. 743, pp. 338–342, Mar. 2015.

[29] A. Hiassat, A. Diabat, and I. Rahwan, "A genetic algorithm approach for location-inventory-routing problem with perishable products," *J. Manuf. Syst.*, vol. 42, pp. 93–103, Jan. 2017.

[30] B. D. Song and Y. D. Ko, "A vehicle routing problem of both refrigerated- and general-type vehicles for perishable food products delivery," *J. Food Eng.*, vol. 169, pp. 61–71, Jan. 2016.

[31] Y. Bo, "Application of the improved fruit flying optimization algorithm in urban logistics distribution center location," *J. Shandong Agricult. Univ. (Natural Sci. Ed.)*, vol. 46, no. 4, pp. 632–635, 2015.

[32] A. T. Mulloorakam and N. M. Nidhiry, "Combined objective optimization for vehicle routing using genetic algorithm," *Mater. Today, Proc.*, vol. 11, no. 3, pp. 891–902, 2019.

[33] F. Simsir and D. Ekmekci, "A Metaheuristic solution approach to capacitied vehicle routing and network optimization," *Eng. Sci. Technol., Int. J.*, vol. 22, no. 3, pp. 727–735, Jun. 2019.

[34] F. Habibi, E. Asadi, and S. J. Sadjadi, "A location-inventory-routing optimization model for cost effective design of microalgae biofuel distribution system: A case study in Iran," *Energy Strategy Rev.*, vol. 22, pp. 82–93, Nov. 2018.

[35] W. Xu and Y. Yin, "Functional objectives decision-making of discrete manufacturing system based on integrated ant colony optimization and particle swarm optimization approach," *Adv. Prod. Eng. Manage.*, vol. 13, no. 4, pp. 389–404, Dec. 2018.

[36] D. Tuzun and L. I. Burke, "A two-phase tabu search approach to the location routing problem," *Eur. J. Oper. Res.*, vol. 116, no. 1, pp. 87–99, Jul. 1999.

[37] T.-H. Wu, C. Low, and J.-W. Bai, "Heuristic solutions to multi-depot location-routing problems," *Comput. Oper. Res.*, vol. 29, no. 10, pp. 1393–1415, Sep. 2002.

[38] M. Albareda-Sambola, J. A. Diaz, and E. Fernández, "A compact model and tight bounds for a combined location-routing problem," *Comput. Oper. Res.*, vol. 32, no. 3, pp. 407–428, Mar. 2005.

[39] A. G. Qureshi, E. Taniguchi, and T. Yamada, "A microsimulation based analysis of exact solution of dynamic vehicle routing with soft time windows," *Procedia-Social Behav. Sci.*, vol. 39, pp. 205–216, 2012.

[40] H. Sadjady and H. Davoudpour, "Two-echelon, multi-commodity supply chain network design with mode selection, lead-times and inventory costs," *Comput. Oper. Res.*, vol. 39, no. 7, pp. 1345–1354, Jul. 2012.

[41] K. Holmberg, "Exact solution methods for uncapacitated location problems with convex transportation costs," *Eur. J. Oper. Res.*, vol. 114, no. 1, pp. 127–140, Apr. 1999.

[42] H.-S. Hwang, "A stochastic set-covering location model for both ameliorating and deteriorating items," *Comput. Ind. Eng.*, vol. 46, no. 2, pp. 313–319, Apr. 2004.

[43] W. Gong, D. Li, X. Liu, J. Yue, and Z. Fu, "Improved two-grade delayed particle swarm optimisation(TGDPSO) for inventory facility location for perishable food distribution centres in Beijing," in *Proc. Int. Conf. Comput. Comput. Technol. Agricult.*, 2007.

[44] W.-T. Pan, "A new fruit fly optimization algorithm: Taking the financial distress model as an example," *Knowl.-Based Syst.*, vol. 26, pp. 69–74, Feb. 2012.

[45] D. Kaifeng, "Research on location method of rural logistics center based on AP clustering," Lanzhou Univ., Lanzhou, China, Tech. Rep., 2015.

[46] *China Statistical Publishing House*. Accessed: 2017. [Online]. Available: http://data.tjj.cq.gov.cn/easyquery.htm?cn=D0102/

HAO ZHANG was born in China, in January 1978. He received the bachelor's degree in management and the Ph.D. degree in systems engineering from Harbin Engineering University, in 2004 and 2009, respectively.

From 2009 to 2011, he was a Lecturer with Beijing Technology Business University (BTBU), where he was an Associate Professor with the Logistics Management Department, from 2011 to 2018. Since 2018, he has been a Professor and the Deputy Director with the Science and Technology Department, BTBU. He is the author of six books, more than 50 articles, and more than six software copyrights. His research interests include supply chain and logistics management, digital management, and modeling and simulation of complex systems.

Dr. Zhang was a recipient of the Outstanding Paper Award from the China Society of Logistic, the BaoGong Logistics Award, the Humanities and Social Sciences Award, Heilongjiang, and the Beijing Teaching Achievement Award from the Beijing Technology and Business University.

HUIXIA FENG was born in Xinzhou, China, in 1996. She received the bachelor's degree in management from the Shanxi University of Finance and Economics, in 2018.

Since 2019, she has been a Postgraduate Research Student with the School of E-Business and Logistic, Beijing Technology and Business University. Her research interests include supply chain and logistics management, digital management, and modeling and simulation of complex systems.

Dr. Feng was a recipient of the Third Prize in the Mathematical Contest in Modeling, Shanxi, in 2016, and the Freshman Admission Scholarship from BTBU, in 2019.

HONGMEI WANG received the bachelor's degree in management from the Business School, Beijing Technology and Business University, in 2019.

She is currently a Postgraduate Research Student with Beijing Jiaotong University. Her specialization is logistics management. Her research interests include supply chain management and data mining and analysis. She was a recipient of the Second Prize of the University Data Analysis Contest twice in a row, the Third Prize in Thesis Writing Contest during the undergraduate period, and the Freshman Admission Scholarship from BJTU, in 2019.