RESEARCH ARTICLE

Optimal location of logistics distribution centres with swarm intelligent clustering algorithms

Tsung-Xian Lin\textsuperscript{1,2}, Zhong-huan Wu\textsuperscript{1,2}\textsuperscript{*}, Wen-Tsao Pan\textsuperscript{3}

\textsuperscript{1} Department of Management, Guangzhou Huashang College, Guangzhou, China, \textsuperscript{2} Institute for Economic and Social Research, Guangzhou Huashang College, Guangzhou, China, \textsuperscript{3} School of business, Guangdong University of Foreign Studies, Guangzhou, China

* M18819453269_1@163.com

Abstract

A clustering algorithm is a solution for grouping a set of objects and for distribution centre location problems. But the common K-means clustering algorithm may give local optimal solutions. Swarm intelligent algorithms simulate the social behaviours of animals and avoid local optimal solutions. We employ three swarm intelligent algorithms to avoid these solutions. We propose a new algorithm for the clustering problem, the fruit-fly optimization K-means algorithm (FOA K-means). We designed a distribution centre location problem and three clustering indicators to evaluate the performance of algorithms. We compare the algorithms of K-means with the ant colony optimization algorithm (ACO K-means), particle swarm optimization algorithm (PSO K-means), and fruit-fly optimization algorithm. We find K-Means modified by the fruit-fly optimization algorithm (FOA K-means) has the best performance on convergence speed and three clustering indicators, compactness, separation, and integration. Thus, we can apply FOA K-means to improve the distribution centre location solution and the efficiency for distribution in the future.

1 Introduction

The development of Machine Learning provides supervised learning and unsupervised learning algorithms. Supervised learning uses labelled data sets to train algorithms to classify data and predict outcomes. We can use a linear model to make an economic data forecast, like the influence of educational input on gross national income (GNI) [1]. A deep neural network is a widely used non-linear model for animal and human face recognition [2] and stock returns and options market forecast [3–5]. Unlike supervised learning, unsupervised learning can classify unlabelled data. The clustering algorithm is an unsupervised learning method that groups a set of unlabelled objects.

The clustering algorithm can help solve many problems, like large databases and customer segmentation. Many researchers developed clustering algorithms for various customers segmentation, like in the automobile retailer industry [6], airline customers [7], e-commerce companies [8,9]. Clustering can help forecast customer demand and improve marketing
strategy for companies [10]. There have been many clustering methods. The K-means clustering algorithm is widely used in previous cases, and k-means clustering has been widely studied with various extensions in the literature and applied in many substantive areas [11–14]. However, these k-means clustering algorithms are usually initialized and need to be given some cluster priors. In general, the cluster number is unknown. In this case, the validity index can be used to find the cluster numbers where they should be independent of the clustering algorithm [15]. Many clustering effectiveness metrics for k-means clustering algorithms are already in the literature, such as Bayesian Information Criterion (BIC) [16], Akaike Information Criterion (AIC) [17], Dunn Index [18], Silhouette Width (SW) [19], Calinski and Harabasz Index (CH) [20], Gap Statistics [21], Generalized Dunn Index (DNg) [22], and modified Dunn indices (DNs) [23]. so we focus on the K-means clustering algorithm in this paper [24].

The K-means clustering algorithm is efficient, but the simple design may suffer from local optimal solutions. Clustering is combinations of data points in different clusters and, thus, is a combinatorial optimization problem. In practice it has been found that metaheuristic algorithms like Genetic Algorithms (GAs) are better choice for combinatorial optimization [25]. Some researchers try the hybrid genetic algorithm to improve K-means [26,27]. We try to develop K-Means with a swarm intelligence algorithm because the swarm intelligence algorithms have advantages to find a better solution in non-convex programming problems and the swarm intelligence algorithms have advantages to find a better solution in non-convex programming problems compared with other algorithms. Particle swarm optimization (PSO) [28] and Ant Colony Optimization (ACO) [29] have good features to simulate animals behaviours to avoid local solutions. PSO is a new biological evolution algorithm. It starts from a random solution, finds the optimal solution through iteration, and evaluates the quality of the optimal solution through fitness. It is simpler than the genetic algorithm, and there is no "crossover" and "mutation" compared with genetic algorithm, which finds the global optimum by searching optimum value. Each in the ACO algorithm can change the surrounding environment by releasing pheromones, and each can perceive the real-time changes in the surrounding environment, and the individuals communicate indirectly. At the same time, the distributed computing method is adopted in the search process, and multiple individuals perform parallel computing, which greatly improves the computing ability and operating efficiency. We applied the clustering algorithm modified by PSO [30] and ACO [31]. Additionally, we develop a new swarm intelligence algorithm, Fruit fly Optimization (FOA), to the clustering algorithm. Fruit fly Optimization has been proven to find a satisfactory solution and good convergence speed [32–34]. By simulating the predation process of fruit flies using their keen sense of smell and vision, FOA realizes a group iterative search. The principle of FOA is easy to understand, simple to operate, easy to implement, and has strong local search ability. The three algorithms probability search method is not easy to fall into the local optimum, and it is easy to find the global optimum solution. Given the advantage of K-Means and FOA, we can obtain a good algorithm for the clustering problem.

Logistics distribution centres are intensive and specialized, serving a set or region of customers. The logistics distribution centres store, sort, distribute, and deliver goods based on customer requirements. Selecting algorithms to achieve lower costs will help companies work efficiently and create more value for society [35]. There are many positive cases about location, such as e-commerce in a city [36], the clothing industry [37], and the locations for relief distribution and victim evacuation after a disaster [38]. Clustering and location problems have similar objectives, so we use a clustering algorithm to help cluster customers and locate the distribution centres.

This paper is organized as follows: Section 2 Methodology shows the details of common K-Means clustering, ACO, PSO, and FOA K-means. Section 3 Results and discussion conducts
empirical analysis and describes the superior performance of FOA K-means. Section 4 presents the conclusions.

2 Methodology

The K-means algorithm is an efficient and widely used algorithm for clustering. The simple design suffers from local optimal solutions in many cases. But the swarm intelligence algorithms can find better solutions in non-convex programming problems. Three swarm intelligence algorithms are employed to solve the location problem for distribution problems. The ant colony optimization algorithm, particle swarm optimization algorithm, and fruit-fly optimization algorithm are introduced as follows.

2.1 K-Means algorithm

K-Means clustering is a fast, robust, and simple algorithm with an effective convergence property that gives results when data sets are distinctly separated from each other. It performs well when the number of cluster centres is specified by a well-defined list of types shown in the data. The process of common K-Means is:

Step 1: Randomly select K objects from N objects as the cluster centre, where K is the required target classification number.

Step 2: Calculate the distance of all remaining objects to each cluster centre. Use the Euclidean distance to assign objects to the same categories and nearest cluster centre.

Step 3: Determine the cluster centres based on the mean properties of each object in the same categories.

Step 4: Repeat steps (2) and (3) until the new cluster centre is equal to or less than the target threshold and the clustering algorithm ends.

2.2 Ant colony optimization algorithm

The ant colony optimization algorithm (ACO) is a probabilistic algorithm for route optimization and clustering analysis. This application performs well for transportation and telecommunication networks [31]. The ant colony algorithm was proposed by Marco Dorigo in 1992 and shows how ants find their way to food. The main characteristics of this algorithm include distributed computing, heuristic search, and feedback. It is an effective algorithm to find a satisfactory solution and avoid falling into local solutions.

The clustering algorithm modified by the ant colony algorithm appears in Fig 1. The process is as follows:

Step 1: Initialize the ant colony parameters. Set the maximum number of iterations to 1000. Set the number of cluster samples, attributes, and cluster centres. Set the pheromone evaporation rate rho = 0.1, which will update the pheromone matrix. Set the number R of ants (the solutions) to 100.

Step 2: Initialize the pheromone matrix. The pheromone matrix is a matrix of N × K (sample number × cluster number). Set all pheromone initial values to 0.01.

Step 3: All ants construct a solution set according to the pheromone matrix $\tau_{i,N\times K}$. The higher $\tau_{i,k}$, object i is assigned to categories k and centre k.
Fig 1. Flowchart of ant colony optimization algorithm.

https://doi.org/10.1371/journal.pone.0271928.g001
Step 4: Generate the vector $r_{1 \times N}$ for each ant. If the corresponding value $r_i$ is less than the set pheromone threshold $q = 0.9$, the path takes the maximum pheromone. If there are multiple maximum values, randomly take one of them from the same maximum value as the path. If the pheromone is greater than the set threshold $q = 0.9$, the ratio of each path pheromone to the total pheromone of the sample is obtained. The initial path is determined by probabilities based on the pheromone proportion of each path. In this step, each ant may have a different solution.

Step 5: Determine the new centres based on the mean attributes of objects in the same categories.

Step 6: The deviation error $F_{ij}$ is the Euclidean distances of each object to the cluster centre $j$. The smaller the $F_{ij}$ is, the better the clustering effect is. Calculate the sum $D_i$ of deviation smallest error $F_{ij}$ for each sample $i$ to the nearest cluster centre $j$ in each ant and find the best ant solutions $D_{best}$.

Step 7: Update the pheromone value according to the best $L$ solutions ($L = 2$) in the last step. The pheromone matrix update is as follows:

$$\tau_{ij}^{new} = (1 - \rho) \times \tau_{ij}^{best solution} + \frac{1}{F}$$

$$F = \sum_{i=1}^{t} D_i$$

$\tau_{ij}^{best solution}$ is updated according to the best solution.

Step 8: Determine whether the maximum number of iterations is reached. If not reached, return step (3); if reached, the optimal cluster solution is output.

### 2.3 Particle swarm optimization algorithm

Kennedy and Eberhart first introduced the particle swarm optimization algorithm (PSO). PSO is effective and highly efficient as well [39]. This algorithm simulates the social behaviours of birds and is a good optimization algorithm with relatively low iteration numbers. A study finds that adding environmental factors can improve the performance of PSO in searching strategy, but the question is how to select the environmental factor more effectively [40].

In the PSO algorithm, each particle can remember the best solution that it has searched for. $P_{id}$ is the best position that the entire particle swarm has experienced. $P_{gd}$ is the best solution in the current search. Each particle has a speed, $V$ represents the velocity of the $i$-th particle in the $d$-th dimension, $w$ is the inertia weight, $\eta_1$, $\eta_2$ are the importance parameters for $P_{id}$ and $P_{gd}$, and rand() is the random function that generates a floating variable in the range [0,1]. Calculate:

$$V_{id} = wV_{id} + \eta_1 \text{rand}() (P_{id} - X_{id}) + \eta_2 \text{rand}() (P_{gd} - X_{id}), \quad \eta_1 = \eta_2 = 1.2$$

$$w = w_{max} - t * (w_{max} - w_{min}) / M, \quad w_{max} = 0.9, \quad w_{min} = 0.4, \quad M \sim \text{iterations}$$

Then the updated position of particle is:

$$X_{id} = X_{id} + V_{id}$$

The process of the PSO K-Means Algorithm is as follows:
Step 1: Initialize the particles and randomly assign each object to a category. Determine cluster centres based on the mean of the objects in the same category. Initialize the velocity of the particles, iterations, and group size.

Step 2: For each particle, calculate the fitness value, and compare it to the recorded best fitness value of $P_{id}$. If there is one better than $P_{id}$, update $P_{id}$.

Step 3: Compare each particle fitness value with the best position in the current group $P_{gd}$. If there is a better one, update $P_{gd}$.

Step 4: Determine the velocity and position of the particles based on Formula (1) and Formula (2).

Step 5: Apply the K-means optimization. Update the cluster centre, apply the nearest neighbour rule, and update the particle.

Step 6: If the end condition is reached (maximum iterations $M = 200$ in this case), the process ends. Otherwise, return to step (2).

2.4 Fruit fly optimization algorithm

The fruit fly optimization algorithm (FOA) was proposed by Professor Pan (2011) and has been widely applied in solving programming problems. These problems include logistics storage selection [22] and a mutual fund forecasting model [41], which optimize the parameters of a general regression neural network in the forecasting model and obtain a better solution than PSO. The fruit-fly optimization algorithm simulates the fruit-fly behaviours of seeking food and contributes a swarm intelligence optimization algorithm. The process of the standard fruit fly algorithm is as follows:

Step 1: Initialize the population size $Size_{pop}$, the number of iterations $Max_{gen}$, and the random location of the fruit fly.

Step 2: Each fruit fly randomly searches for food in all directions.

Step 3: Calculate the distance $Dist_i$ from each fruit fly to the origin and set the smell value $S_i$.

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

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

Step 4: Substitute the obtained smell value to the fitness function and determine the best fruit fly.

Step 5: Identify the best-performing fruit flies in the fruit fly population.

Step 6: Record the optimal taste density value and the specific $X$, $Y$ coordinates.

$$Smell_i = \text{Function}(S_i)$$

$$[\text{bestSmell, bestIndex}] = \max (\text{Smell}_i)$$

$$\text{Smell}_{best} = \text{best Smell}$$

$$X_{axis} = X(\text{bestIndex})$$

$$Y_{axis} = Y(\text{bestIndex})$$
Step 7: Start the iteration, repeat (2) through (6), and if the current iteration time is less than maximum iterations Maxgen, then perform step (2). Otherwise, end the calculation.

The Basic FOA will be reconstructed to make it more suitable for solving clustering problems based on the K-Means algorithm.

2.5 K-Means modified by FOA

In the previous result, we can see some flaws while we are using the other algorithms. ACO avoids falling into a local optimal solution, but the convergence speed is low. On the other hand, PSO updates the decisive variables well, but compared with K-Means, it does not show its advantage of being a swarm intelligent algorithm. So, we want to develop an algorithm to combine the characteristics of good convergence speed of K-Means and FOA and Modified FOA. The K-Means algorithm is the new algorithm to satisfy our demands.

We develop the modified FOA K-Means algorithm as follows:

Step 1: Design the fruit fly structure for clustering. Each fruit fly is a cluster centre. There are many objects, assigned to N categories. There are N centres, and each centre has k attributes. \(x_1, y_1\) represent the coordinate vectors of the i-th fruit fly. The size of the fruit-fly population is sizepop. When initializing, \(x_i, y_i\) are randomly generated, and \(\text{rands}(1, N \times k)\) represents a random number of \((0, 1)\), which produces one row and \(N \times k\) columns. \(rcc\) is the max attribute value. In the following case, we set it to 6000. \(D_i\) is the distance from each fruit fly to the origin.

\[
\begin{align*}
x_i, y_i & = \text{rand} \times rcc \\
rcc & = \text{max attribute} \\
X_i & = \{x_1, x_2, x_3, \ldots, x_{N \times k}\} \\
Y_i & = \{y_1, y_2, y_3, \ldots, y_{N \times k}\} \\
D_i & = \sqrt{X_i^2 + Y_i^2 + Z_{i,\text{modi}}} \quad \text{(when initializing vector } Z_{i,\text{modi}} = 0) 
\end{align*}
\]

Step 2: The objects closest to the same cluster centre are categorized into the same class. Sum up the Euclidean distances from the cluster centre to each object. Apply the K-means method to generate a new cluster centre CentreBest with the shortest total distance of all solutions.

Step 3: Determine the adjust vector \(Z_{i,\text{modi}}\)

\[
Z_{i,\text{modi}} = \text{CentreBest}_i - \left( X_{i,\text{axis}}^2 + Y_{i,\text{axis}}^2 \right)^{1/2}
\]
Step 4: Determine the variation from best fruit fly. Compare the original best fruit fly with other flies with mutations. The method of mutation is as follows.

\[ X_i = X_{\text{best axis}} + (2 \times \text{rands}(1, N \times k) - 1) \times \text{ssc} \]

\[ Y_i = Y_{\text{best axis}} + (2 \times \text{rands}(1, N \times k) - 1) \times \text{ssc} \]

\[ \text{ssc} = \frac{\max \text{ attribute}}{2} \]

Step 5: Determine whether max iterations is reached. If not, return to step 2. Otherwise, output the result. We set 200 iterations.

3 Results and discussion

To evaluate the performance of swarm intelligence algorithms, we conduct a calculation test case.

3.1 Location problem

In a city, retailers need to send goods to many customers. They must build distribution centres to serve all customers. The locations of customers appear in Fig 2.

We need to determine the number of distribution centres. Within a group, the sum of squared error (SSE) is a common method to determine the optimal number of clusters. In Fig 3, the optimal number of clusters is three. With more categories or clusters, there is a shorter distance within the group’s sum of squared error. We focus on the change of slope. The decline from one to three categories is rapid and then decreases slowly, so to reduce the construction...
cost, the optimal number of clusters is three. To serve the needs of the retailers, three logistics
distribution centres should be built.

3.2 Results of the algorithms

3.2.1 Common K-Means. Means results appear in Fig 5. The sample number N is 300,
and the cluster number K is 3. Fig 4 shows the result.
The distribution centres result appears in Table 1.

### 3.2.2 K-Means modified by ACO

The ant colony algorithm is used to solve the problem. The initial parameters are as follows: the iteration is 200, the ant number is 200, the sample number is 300, and the cluster number is 3. After running ACO, we find that the ant colony clustering algorithm is unsuitable for this problem. The convergence is poor with more iterations. The result appears in Fig 5.

The distribution centres appear in Table 2:

### 3.2.3 K-Means modified by PSO

We set initial parameters: the iteration is 200, the particle size is 200, the sample number is 300, and the cluster number is 3. The result appears in Fig 6.

The distribution centres are shown in Table 3:

### 3.2.4 K-Means modified by FOA

In the previous result, we saw some flaws in the other algorithms. ACO avoids falling into local optimal solution, and PSO updates the decision variables but does not show its advantage of a swarm intelligent algorithm compared with

---

**Table 1. K-Means distribution centres.**

|                         | Horizontal axis(X/m) | Vertical axis(Y/m) |
|-------------------------|----------------------|--------------------|
| Distribution centre 1   | 3198.532555          | 3458.119546        |
| Distribution centre 2   | 1531.344733          | 2906.774243        |
| Distribution centre 3   | 2570.17              | 1218.828           |

The total distance from each retailer to the distribution centres is 184400.41 m.

**Table 2. ACO distribution centres.**

|                         | Horizontal axis(X/m) | Vertical axis(Y/m) |
|-------------------------|----------------------|--------------------|
| Distribution Centre 1   | 2872.944             | 3401.305           |
| Distribution Centre 2   | 1930.966             | 2349.514           |
| Distribution Centre 3   | 2417.269             | 1607.756           |

The total distance from each retail customer to the distribution centres is 280379.72 m.
K-Means. We develop a new FOA K-Means algorithm with good convergence speed and performance to avoid local optimal solutions.

The result of the fruit fly algorithm appears in Fig 7:

The distribution centres appear in Table 4:

### 3.2.5 Comparison of results.

In Table 5, Total Distance is a compactness indicator, so smaller is better. Total Centre Distance is the separation indicator, so higher is better. Davies-Bouldin index is an integrated indicator, so smaller is better.

ACO is not suitable for cluster analysis in large clustering samples in the study due to the poor performance of three indicators. Total Distance is the highest, 280379.72. The total centre distance is the lowest, 4149.43. Davies-Bouldin Index is the highest, 4.53. Compared with Common K-Means, MFOA K-Means, and PSO K-Means, the convergence speed of ACO is obviously slower in the computational experiment.

The PSO K-Means and FOA K-Means algorithms can prevent the Common K-Means from falling into a local optimum as well. In this case, PSO K-Means does not perform better than common K-Means due to the higher compactness indicator and higher DBI. PSO K-Means is more likely to fall into a local optimal solution than FOA. The performance of FOA K-Means is better than that of common K-Means and has the best performance in three indicators. FOA K-Means has the smallest DBI, 1.846.

Through the results in Fig 9, we can also draw further conclusions:

In the results, we can see that when the FOA K-Means algorithm reached a local optimal state, it still could optimize the solution. If we run more tests on common K-Means, it might generate a worse solution than the previous result. FOA K-Means has a more stable

---

### Table 3. PSO K-Means distribution centres.

| Distribution Centre | Horizontal axis(X/m) | Vertical axis(Y/m) |
|---------------------|-----------------------|--------------------|
| 1                   | 3282.131              | 3430.185           |
| 2                   | 2562.433              | 1209.77            |
| 3                   | 1601.736              | 2966.247           |

The total distance from each retailer to the distribution centres is 185227.7256 m.

---
Table 4. MFOA distribution centres.

| Distribution Centre | Horizontal axis(X/m) | Vertical axis(Y/m) |
|----------------------|----------------------|-------------------|
| Distribution Centre 1| 3231.8               | 3431.8            |
| Distribution Centre 2| 1562.7               | 2930.4            |
| Distribution Centre 3| 2559.9               | 1198.7            |

The total distance from each retailer to the distribution centres is 184365.38 m.

https://doi.org/10.1371/journal.pone.0271928.t004
performance. Another cluster intelligent algorithm, the particle swarm K-Means algorithm, compared to the modified fruit fly optimization K-Means algorithm, has little advantage in this problem.

![MFOA K-Means Iteration Result](https://doi.org/10.1371/journal.pone.0271928.g008)

![PSO K-Means Iteration Result](https://doi.org/10.1371/journal.pone.0271928.g008)

![Common K-Means Iteration Result](https://doi.org/10.1371/journal.pone.0271928.g008)

Fig 9. Three optimal clustering algorithms iteration result.

https://doi.org/10.1371/journal.pone.0271928.g009

Table 5. Comparable indicators of the four algorithms.

| Algorithms                  | Common K-Means | ACO    | MFOA K-Means | PSO K-Means |
|-----------------------------|----------------|--------|--------------|-------------|
| Total Distance (m)          | 184400.41      | 280379.72 | 184365.38    | 185227.73   |
| Compactness indicator       |                |        |              |             |
| Total Centre Distance (m)   | 6063.77        | 4149.43 | 6073.07      | 6079.44     |
| Separation indicator        |                |        |              |             |
| Davies-Bouldin Index        | 1.846841       | 4.532969| 1.84223476   | 1.85038066  |
| Integrated indicator        |                |        |              |             |

The comparison figure of the algorithms appears in Fig 8 and Table 5:

https://doi.org/10.1371/journal.pone.0271928.t005
4 Conclusions

We solved the location problem with swarm intelligent clustering algorithms (including the ant colony optimization algorithm, the fruit fly optimization K-Means algorithm, and the particle swarm optimization algorithm) and the common K-Means algorithm. We proposed a new swarm intelligent algorithm, the fruit fly optimization K-Means algorithm (FOA K-Means). We compared the indicators of algorithm results. The FOA K-Means algorithm has the best performance in the DBI indicator, which has great significance for clustering problems in further studies. FOA K-means has the best performance on convergence speed and three clustering indicators, compactness, separation, and integration. ACO is not suitable for cluster analysis of the large clustering samples in the study due to the poor performance in three indicators. PSO K-Means is easier to fall into a local optimal solution than FOA. Therefore, FOA K-means is the best solution for distribution centre location problems and clustering.

The study has some deficiencies. To analyse the clustering problem further, some actual constraints are neglected in location problems. First, the location of the distribution centres is obtained using the mean of the coordinates. Second, the relative importance of different customers is not considered. Different customers have various needs, which can affect the location of the distribution centres. The need for excellent customer service becomes important. And any business’ ability to understand each segment of its customers’ needs will give it a greater advantage in providing targeted customer service and developing custumed centres. Future research can explore addressing these problems further.

Supporting information

S1 Dataset.
(XLSX)

Author Contributions

Data curation: Tsung-Xian Lin.

Formal analysis: Zhong-huan Wu.

Funding acquisition: Tsung-Xian Lin.

Software: Wen-Tsao Pan.

Supervision: Tsung-Xian Lin.

Writing – review & editing: Tsung-Xian Lin.

References

1. Pan W. T. (2017). "A Newer Equal Part Linear Regression Model: A Case Study of the Influence of Educational Input on Gross National Income." Eurasia Journal of Mathematics Science & Technology Education 13(8): 5765–5773.
2. Jiang H. and Learned-Miller E. (2017). Face Detection with the Faster R-CNN. IEEE International Conference on Automatic Face & Gesture Recognition.
3. Chen L., Pelger M. and Zhu J. (2019). "Deep Learning in Asset Pricing." SSRN Electronic Journal.
4. Gu S., Kelly B. T. and Xiu D. (2018). "Empirical Asset Pricing via Machine Learning." Social Science Electronic Publishing.
5. Kaplan S. and Mansilla K. (2020). "Machine Learning and Asset Pricing: an Application for Option Pricing across Markets."
6. Baharisaravi S. (2012). "New approach clustering algorithm for customer segmentation in automobile retailer industry."

7. Wei and L. Feng(2012). "Design and Implementation of Airline Customer Segmentation System Based on Ant Colony Clustering Algorithm." Advanced Materials Research 433– 440: 3357–3361.

8. Chen X., Fang Y., Yang M., Nie F., Zhao Z. and Huang J. Z. (2018). "PurTreeClust: A Clustering Algorithm for Customer Segmentation from Massive Customer Transaction Data." IEEE Transactions on Knowledge & Data Engineering PP(3): 1–1.

9. Nurmalasari A. Mukhayaroh S. Martlina S.Hartinianad Adiwihardja C. (2020). "Implementation of Clustering Algorithm Method for Customer Segmentation." Journal of Computational and Theoretical Nanoscience 17(2): 1388–1395.

10. Caniato F., Matteo K., Ronchi S., Verganti R. and Zottieri G. (2005). "Clustering customers to forecast demand." Production Planning & Control 16(1): 32–43.

11. Alhawarat M. and Hegazi M., “Revisiting k-means and topic modeling, a comparison study to cluster arabic documents,” IEEE Access, vol. 6, pp. 42740–42749, 2018.

12. Meng Y., Liang J., Cao F., He Y., “A new distance with derivative information for functional k-means clustering algorithm," Information Sciences, vol. 463–464, pp. 166–185, 2018.

13. Lv Z., Liu T., Shi C., Benediktsson J.A., Du H., “Novel land cover change detection method based on k-means clustering and adaptive majority voting using bitemporal remote sensing images,” IEEE Access, vol. 7, pp. 34425–34437, 2019.

14. Zhu J., Jiang Z., Evangelidis G.D., Zhang C., Panga S., Li Z., “Efficient registration of multi-view point sets by k-means clustering,” Information Sciences, vol. 488, pp. 205–218, 2019.

15. Halkidi M., Batistakis Y., Vazirgiannis M., “On clustering validation techniques,” J. Intell. Inf. Syst., vol. 17, pp. 107–145, 2001.

16. Kass R.E., Raftery A.E., “Bayes Factors,” Journal of the American Statistical Association, vol. 90, pp. 773–795, 1995.

17. Bozdogan H., “Model selection and Akaike’s information criterion (AIC): The general theory and its analytical extensions,” Psychometrika, vol. 52, pp. 345–370, 1987.

18. Dunn J.C., “A fuzzy relative of the ISODATA process and its use in detecting compact, well-separated clusters,” J. Cybernetics, vol. 3, pp. 32–57, 1974.

19. Rousseew P.J., “Silhouettes: A graphical aid to the interpretation and validation of cluster analysis,” Journal of Computational and Applied Mathematics, vol. 20, pp. 53–65, 1987.

20. Calinski T., Harabasz J., “A dendrite method for clusteranalysis,” Commun. Stat.-Theory Methods, vol. 3, pp. 1–27, 1974.

21. Tibshirani R., Walther G., and hastie T., “Estimating the number of clusters in a data set via the gap statistic,” Journal of the Royal Statistical Society: Series B, vol. 63(2), pp. 411–423, 2001.

22. Pal N. R. and Biswas J., “Cluster validation using graph theoretic concepts,” Pattern Recognition, vol. 30, pp. 847–857, 1997.

23. Ila N., “Modified Dunn’s cluster validity index based on graph theory,” Przeglad Elektrotechniczny (Electrical Review), vol. 2, pp. 126–131, 2012.

24. Hartigan J. A. and Wong M. A. (1979). “A K-Means Clustering Algorithm.” Applied Statistics 28(1).

25. Michalewicz Z., Genetic Algorithms + Data Structures = Evolution Programs, 3rd edition. Springer, 1996.

26. Bandyopadhyay S. and Maulik U. (2002). “An evolutionary technique based on K-Means algorithm for optimal clustering in RN.” Information Sciences 146(1–4): 221–237.

27. Krishna K. and Murty M. N. (1999). “Genetic K-Means Algorithm." IEEE TRANSACTIONS ON CYBERNETICS 29(3): 433–439. https://doi.org/10.1109/3477.764879 PMID: 18252317

28. Poli R., Kennedy J. and Blackwell T. (2007). "Particle swarm optimization." Swarm Intelligence 1(1).

29. Dorigo M., Birattari M. and Stützle T. (2006). "Ant Colony Optimization." IEEE Computational Intelligence Magazine 1(4): 28–39.

30. Chen C. Y. and Ye F. (2012). Particle swarm optimization algorithm and its application to clustering analysis. IEEE International Conference on Networking.

31. Mirabi M. and Seddighi P. (2017). "Hybrid ant colony optimization for capacitated multiple-allocation cluster hub location problem." Artificial Intelligence for Engineering Design Analysis & Manufacturing: 1–15.

32. Du Ting-Song, Ke Xian-Ting, Liao Jia-Gen, Shen and Jun(2018). "DSLC-FOA: Improved fruit fly optimization algorithm for application to structural engineering design optimization problems." Applied Mathematical Modelling.
33. Salehi M., Farhadi S., Moieni A., Safaie N. and Hesami M. (2021). "A hybrid model based on general regression neural network and fruit fly optimization algorithm for forecasting and optimizing paclitaxel biosynthesis in Corylus avellana cell culture." Plant Methods 17(1): 1–13. https://doi.org/10.1186/s13007-021-00714-9

34. Wang X., Chen H., Heidari A. A., Zhang X. and Huang H. (2020). "Multi-population following behavior-driven fruit fly optimization A Markov chain convergence proof and comprehensive analysis (Knowledge-Based Systems, Impact Factor: 5.921)." Knowledge-Based Systems http://aliasgharheidari.com:106437.

35. Zhao S. A. and Chi-Wen Q. U. (2017). "An Improved Cuckoo Algorithm for Solving the Problem of Logistics Distribution Center Location." Mathematics in Practice and Theory.

36. Brasileiro L. A. and Chiuffa V. (2018). "Simulation of the Optimal Location for a Distribution Center of E-Commerce Operations in São Paulo State." IEEE Latin America Transactions.

37. Liu M., Xu X. and Zhang D. (2019). "Integrated optimization model for distribution network design: a case study of the clothing industry." International Transactions in Operational Research.

38. Setiawan E., Liu J. and French A. (2019). "Resource location for relief distribution and victim evacuation after a sudden-onset disaster." IIE Transactions 51(8): 830–846.

39. Kennedy J. and Eberhart R. (1995) Particle Swarm Optimization. Proceedings of the IEEE International Conference on Neural Networks, 4, 1942–1948.

40. Song W., Ma W. and Qiao Y. (2017). Particle swarm optimization algorithm with environmental factors for clustering analysis. SOFT COMPUTING 21: 283–293.

41. Han S.-Z., Huang L.-H., Zhou Y.-Y. and Liu Z.-L. (2018). Mixed chaotic FOA with GRNN to construction of a mutual fund forecasting model. Cognitive Systems Research 52: 380–386.