WILEY | Hindawi

Research Article

A Big Data Mining and Blockchain-Enabled Security Approach for Agricultural Based on Internet of Things

Feng Zhang [ID] and Yongheng Zhang

School of Information Engineering, Yulin University, Yulin 719000, China

Correspondence should be addressed to Feng Zhang; zhangfeng@yulinu.edu.cn

In order to improve the utilization rate of agricultural big data and solve the security issues problem of multisource and heterogeneous agricultural big data, an improved agricultural big data ant colony optimization algorithm (BigDataACO) is proposed to complete the multisource agricultural big data information in the feature layer and decision-making, and the problem of multisource data fusion was solved. The swarm intelligence algorithm is a process of simulating the complex problem of populations in nature through the mutual cooperation between individuals. The algorithm has potential parallelism and strong robustness, and the algorithm does not depend on specific problems. The definition, principle, and implementation method of agricultural big data fusion problem are studied. Then, the insufficiency of big data fusion modeling algorithm is analyzed. Finally, the source and core steps of the ant colony big data fusion algorithm are studied. The experimental results show that the improved BigDataACO algorithm is verified by the measured data. Compared with K-means, D-S evidence theory, and Bayesian algorithm, the uncertainty of data fusion is greatly reduced by the improved algorithm proposed in this paper.

1. Introduction

Agricultural big data is a collection of data that has a wide range of sources, diverse types, complex structures, and potential value and is difficult to apply common methods of processing and analysis, after integrating its own characteristics such as regional, seasonal, diversity, and periodicity of agriculture [1, 2]. Agricultural big data retains the basic characteristics of big data, such as huge volume, variety, low value, fast processing speed, high veracity, and high complexity, and big data application research in agriculture is still relatively small [3, 4]. With the development of Internet, cloud computing, and other technologies, the Internet of things is used in more and more application fields. Many aspects of smart city also use the Internet of things technology. Smart city includes many aspects, such as intelligent transportation and telemedicine.

By building a decentralized system, application blockchain can provide infrastructure support for big data generated by the Internet of things and help solve the ubiquitous data security problems in the Internet of things, while the Internet of things provides a lot of landing scenarios for the blockchain. This paper applies the characteristics of blockchain, such as peer-to-peer, open and transparent, secure communication, hard to tamper with, and multiparty consensus, which will have an important impact on the Internet of things: the characteristics of multicenter and weak centralization will reduce the high operation and maintenance costs of data centric architecture, the characteristics of information encryption and secure communication will help to protect privacy, and identity rights management and multiparty consensus will help to identify noncompliance. Based on the chain structure, it is helpful to build a verifiable and traceable electronic evidence storage. The distributed architecture and the characteristics of subject equivalence help to break the shackles of multiple information islands in the Internet of things and promote the horizontal flow of information and multiparty cooperation. In order to continuously promote the optimization of the agricultural economy, to realize the sustainable industrial development and regional industrial structure optimization, and further promote the construction of smart agriculture, it is necessary to comprehensively and

timely grasp the development of agriculture, which needs to rely on agricultural big data and related big data fusion processing technology. However, it faces enormous challenges for prediction accuracy in traditional big data modeling algorithms. Since the data fusion is to build a classification model through the training set (i.e., through the classification algorithm), so the fusion rule set that best represents the training data is found. That is a process of gradual optimization [5, 6]; so many researchers applied the swarm intelligence algorithm to the data fusion learning model and achieved some results. The swarm intelligence algorithm is a process of simulating the complex problem of populations in nature through the mutual cooperation between individuals. The algorithm has potential parallelism and strong robustness, and it does not depend on specific problems. The construction of classification learning model based on the swarm intelligence algorithm has become a research hotspot in the field of data mining in recent years [7, 8].

In this paper, the representative of the ant colony algorithm and clustering algorithm in swarm intelligence algorithm are introduced into data fusion mining and decision-making. The problem of constructing based on traditional ant colony classification algorithm and clustering method is studied. And then, the two algorithms are improved from different angles, and a new ant colony data fusion modeling algorithm is proposed. Finally, a number of experiments verify the effectiveness of the improved algorithm in the construction of data fusion learning model.

2. Big Data Fusion and Security Algorithm Based on Ant Colony Algorithm

2.1. Definition of Big Data Fusion. Data fusion is a process in which multiple data are processed to produce more effective and more user-friendly data. Data fusion is the use of computer technology to perform multilevel, multifaceted, multilevel information detection, and correlation estimation and correlation analysis on various multisource and heterogeneous data under certain criteria [9, 10]; in order to obtain the target state and feature estimation, it is more accurate, complete, and reliable than a single data information.

The method of data fusion is generally applied in daily life. For example, when distinguishing a thing, it usually combines various sensory information to process and combines the process of more effective and more in line with the user's needs [11]. When identifying a thing, it is often not enough to synthesize the information obtained by various senses to make accurate judgments on things. Combining multiple sensory data, the description of things will be more accurate. In traditional agricultural big data applications, in some cases, it is not necessary to obtain a large amount of raw data, and only need to obtain the final result, and then, we can use data fusion technology to achieve this purpose.

2.2. Ant Colony Classification Algorithm. Ant colony algorithm is a bionic optimization algorithm, because of its good ability to find good solutions, potential parallelism, positive feedback, and easy to combine with other algorithms, people have applied it to solve many complex combinatorial optimization problems and have shown great potential [12].

The ant colony algorithm, which simulates the foraging behaviour of ant colony, is introduced as a new computational intelligence model. The algorithm is based on the following basic assumptions: ants communicate with each other through pheromones and environment, and each ant reacts only according to its local environment and only affects its local environment; the response of an ant to the environment is determined by its internal model. Because ants are genetic organisms, the behaviour of ants is actually the adaptive performance of their genes, that is, ants are reactive adaptive subjects [13]. At the individual level, each ant makes independent choices based on the environment. At the group level, the behaviour of a single ant is random, but ant colonies can form highly ordered group behaviours through self-organization processes. It can be seen from the above assumptions and analysis that the optimization mechanism of the basic ant colony algorithm includes two basic stages of adaptation and cooperation. In the adaptation phase, each candidate solution continuously adjusts its structure according to the accumulated information. The more ants passing through the path, the larger the amount of information, the easier the path is to be selected, and the smaller the amount of information. In the collaborative phase, the exchange of information between the candidate solutions is expected to produce a better performance solution, similar to the learning mechanism of the learning automaton [14, 15]. Ant colony algorithm is actually a class of multisubject system. Its self-organizing mechanism makes the ant colony algorithm not need to have a detailed understanding of every aspect of the problem. Self-organization is essentially a dynamic process in which the ant colony algorithm mechanism increases the system without external influences and reflects the dynamic evolution from disorder to order. The first ant algorithm was proposed by Dorigo and is called the ant system [16–18]. The ant system incorporates heuristic information and designs the transition probability p_{ij}^k, and the taboo table has been added to enhance the algorithm memory function. In the ant system, the probability that an ant will transfer from node i to node j is defined as [19]:

$$p_{ij}^k(t) = \begin{cases} \dfrac{s_{ij}^a(t)\eta_{ij}^{\beta}(t)}{\sum N_i^k(t)\tau_i^a(t)\eta_i^{\beta}(t)}, & \text{if } j \in N_i^k(t) \\ 0 & \text{if } j \notin N_i^k(t) \end{cases}. \quad (1)$$

In which, s_{ij} is the value of the information on the edge (i, j), t represents the postlatency effect of moving from node i to node j. η_{ij} is heuristic information, calculated by a heuristic function, which represents the a priori effect of moving from node i to node j. The pheromone concentration s_{ij} is a memory of past good quality movements, indicating the impact of past movements from node i to point j on the current selection. The choice of the search path in the ant system is to seek a balance between s_{ij} and η_{ij}. This method can well

handle the relationship between the exploration and development of the ant optimization process.

According to the biological principle of ants, the pheromone on each side introduces a volatilization mechanism, which can encourage ants to explore new paths and avoid premature convergence [20, 21]. In each iteration, the original pheromones need to be volatilized to release new information. For each pheromone on the side, volatilization is performed using equation (2).

$$s_{ij}(t) = (1 - \rho)s_{ij}(t). \tag{2}$$

In which, ρ is a constant whose value ranges from $\rho \in [0,1]$, indicating the degree to which ants have forgotten previous decisions. ρ is the influence of controlling the previous search history. The value of ρ is small, indicating that the volatilization rate is slow; when the value of ρ is large, the volatilization rate is fast.

According to different pheromone update strategies, Dorigo M proposes three different basic ant colony algorithm models, which are called Ant-Cycle model, Ant-Quantity model, and Ant-Density model, the difference is in the difference in $\Delta \tau_{ij}^t(t)$ seeking.

In the Ant-Cycle model,

$$\Delta \tau_{ij}^t(t) = \begin{cases} \dfrac{Q}{L_k}, & \text{If the } K \text{ ant passes through } (i, j) \text{ in this cycle} \\ 0, & \text{else} \end{cases}. \tag{3}$$

In which, Q represents the pheromone intensity, which affects the convergence speed of the algorithm to some extent, and L_k represents the total length of the path taken by the kth ant in this cycle.

In the Ant-Quantity model,

$$\Delta \tau_{ij}^t(t) = \begin{cases} \dfrac{Q}{d_{ij}}, & \text{If the } k \text{ ant passes through } (i, j) \text{ between } t \text{ and } t + 10, \text{else}. \end{cases} \tag{4}$$

In the Ant-Density model,

$$\Delta \tau_{ij}^t(t) = \begin{cases} Q, & \text{If the } k \text{ ant passes through } (i, j) \text{ between } t \text{ and } t + 10, \text{else}. \end{cases} \tag{5}$$

The main difference between them is that the formula (3) and formula (4) use local information, that is, the ant updates the pheromone on the path after completing one step, and the formula (5) uses the overall information, that is, the ant completes. The pheromone on all paths is updated after a loop, and the performance is better when solving.

3. Improved Ant Colony Big Data Fusion Modeling Algorithm

The traditional ant colony fusion modeling algorithm uses a sequential coverage strategy to mine rules one at a time. Since the training set samples covered by the mining rules are removed each time, the search space changes, and the algorithm does not consider the interaction between the discovered rules, such that the rules output earlier will affect the rules that are output later.

3.1. Big Data Fusion Algorithm. Suppose that a multivariate data node will overflow all the nodes of the whole network with its keywords. After the node receives the packet, it will calculate the relevance of the data association [22, 23]. When the source node wants to send data, suppose a certain time t, an ant k that is in node i, and its probability $p_k(i, j)$ to access the next hop node j will be selected according to the following probabilistic criteria, which can be written as formula (6).

$$p_k(i, j) = \begin{cases} \dfrac{\tau(i, j)^a \eta(i, j)^\beta}{\sum \tau(i, u)\eta(i, u)^\beta}, & j \in N_k(i) \\ 0 & j \notin N_k(i) \end{cases} \tag{6}$$

$$\eta(i, j) = \frac{1}{d(i, j)} \alpha\beta.$$

In which, $N_k(i)$ represents a collection of nodes that the node has not yet accessed, τ represents a pheromone, and $\tau(i, j)$ represents the amount of information on the path between nodes i and j; in $\eta(i, j) = 1/d(i, j)$, $\eta(i, j)$ takes the reciprocal of the distance between nodes, which is a heuristic factor, indicating the visibility of the path; α indicates the degree of importance of relative information; β indicates the relative importance of heuristic information.

After all, the ants complete the process of traversing the nodes; the global update rule of the information on each path can be written as

$$\tau_{(i,j)}(t + 1) = (1 - \rho)\tau(i, j) + \Delta\tau(i, j) \tag{7}$$

When the node receives a pheromone update packet, it will update its pheromone table according to Eq. (8) and Eq. (9).

$$\triangle\tau_{ij} = [\varepsilon + (h_i - h_j)] \times \triangle\omega_j, \tag{8}$$

$$\triangle\omega_j = \sum_{j \in R_j} (H_{ij})^{-1} + (h_j)^{-1}. \tag{9}$$

In which, ρ is the evaporation coefficient of the information amount, indicating the length of the pheromone volatilization, $\Delta\tau(i, j)$ represents the increment of the amount of information on the path between the nodes, and its value is determined as shown in equation (10).

$$\Delta\tau(i, j) = \begin{cases} 0; & (i, j) \notin \rho \\ \dfrac{Q}{L_k}; & (i, j) \in \rho \end{cases}. \tag{10}$$

In which, Q is a constant used to control the total amount of pheromone released by the ant after completing a path

search, L_k represents the total length of the path, and ρ represents the path accessed by the ant.

This completes the improvement of the level gradient field. Next, the center point fusion algorithm is used to find the center point, and the fusion tree and data report are established, thus completing the whole algorithm.

3.2. Big Data Fusion Ant Colony Optimization Algorithm. The big data fusion ant colony optimization algorithm (BigDataACO) improves the basic ant colony method. Ant-Miner, an ant colony classification algorithm, aims at mining classification rules with certain structural forms. Where the acquisition of classification rules is one of the main functions of Ant-Miner ant colony classification algorithm. The general structure of classification rules is

$$\text{IF} < \text{conditions} > \text{THEN} < \text{class} > . \tag{11}$$

In which, <conditions> rule antecedent referred to, which consists of a series of conjuncts, comprising a logical combination of the predicted property, its form is

$$< term_1 \text{ and } term_2 \text{ and } \cdots \text{ and } term_n > . \tag{12}$$

Each conjunction is a specific value of the attribute in the training set, and the same attribute can only appear once in the predecessor. The condition items of the precategory of the classification rule are a triple <*attribute, operator, attribute value*>, the attribute in the triplet belongs to the attribute space of the data to be classified; an operator can be a relational operator, often using "=", and attribute values are generally treated as discrete values. <Class> is called a post rule and is a class in the dataset.

The core operation of the ant colony classification search is to generate rules, that is, the current ants sequentially add a rule predecessor to the current partial rule.

Assume that the form of the rule item $term_{ij}$ is $G_i = D_{ij}$, where G_i is the i attribute and D_{ij} is the j value of G_i. The formula for calculating the probability that the term $term_{ij}$ is added to the current partial solution is as follows:

$$H_{ij} == \frac{\tau_{ij}(t)\eta_{ij}}{\sum_{i=1}^{n} x_i \sum_{j}^{b_j} \tau_{ij(t)}\eta_{ij}}. \tag{13}$$

In which, n is the number of attributes, if the current ant does not use the attribute G_i, then x_i is set to 1; otherwise, it is set to 0; b_j is the number of values in the ith attributes range; η_{ij} is the problem-dependent heuristic function of item $term_{ij}$, which is calculated as

$$\eta_{ij} = \frac{\log_2 k - B\left(T \middle| G_i = D_{ij}\right)}{\sum_{i=1}^{n} x_i \sum_{j=}^{b_j} \log_2 k - B\left(T \middle| G_i = D_{ij}\right)}. \tag{14}$$

In which, $B(T \mid G_i = D_{ij})$ is a $term_{ij}$ related entropy. The higher the value of η_{ij}, the greater the likelihood that $term_{ij}$ will be selected.

$\tau_{ij(t)}$ is the pheromone of the current ant path at the node position (i, j), that is, the pheromone at the time item $term_{ij}$. The pheromone of all paths is the same when the algorithm is initialized. The value of this pheromone is inversely proportional to the number of all attribute values.

Ant-Miner uses only a single ant to construct an ant colony in the ant colony construction. Only one ant is used in each iteration of the while loop, and the pheromone update is performed after the ant has completed the construction of the rule. The traditional Ant-Miner can easily select the attribute items in the discovered rules when performing ant colony search. Although the development ability is enhanced, it is easy to prematurely converge, and its calculation method of attribute selection probability is also complicated. This paper proposes a new method based on pheromone attraction and exclusion in the construction of rules. Based on this, the probability formula of state transition is modified, so that the pheromone of ant in the rule search process not only contains the attraction part but also contains the exclusion part. In the exclusion part, ants tend to explore in the initial stage of the search rule process and tend to develop in the latter part of the search.

To use the ant colony construction rules, we first need to initialize the classification modeling algorithm, set the parameter values required by the algorithm, and then place all the training sample data in the training set. Simulate the ant optimization model in the artificial ant colony algorithm to establish the attribute node path. Each node obtains the initial pheromone value according to formula (15):

$$\tau_{ij}(t = 0) = \frac{1}{\sum_{i=1}^{n} b_i}. \tag{15}$$

In which, n represents the total number of sample attributes of the training set, and b_i is the number of values in the value field of the attribute D_{ij}.

In the process of selecting attribute nodes, if the selection is random every time, the calculation time cost of the mining rules will be very large. This paper improves the probability transfer method. The probability formula for the item $term_{ij}$ to be added to the current part is

$$H_{ij}(t) == \frac{\tau_{ij}(t)\eta_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{b_j} \tau_{ij}(t)\eta_{ij}}, \forall i \in I. \tag{16}$$

In which, η_{ij} is problem-dependent term $term_{ij}$ heuristic function. The larger the value, the higher the correlation of $term_{ij}$ in the classification and the greater the possibility of being selected; $\tau_{ij}(t)$ is the pheromone at the $term_{ij}$ position at time t. The improved ant colony data fusion algorithm flow is shown in Figure 1.

4. Experimental Analysis

4.1. Experimental Methods and Data Sets. This experiment analyses and verifies the performance of the BigDataACO algorithm on multisource big data sets. In order to give a more intuitive analysis for the performance of the algorithm,
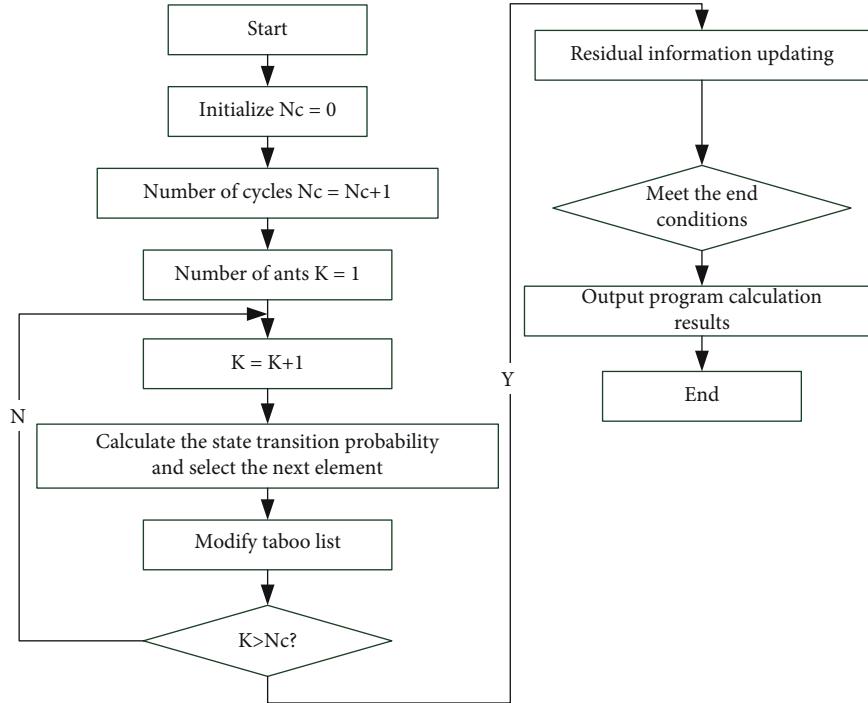
FIGURE 1: The improved ant colony data fusion algorithm flow.

we compare the BigDataACO algorithm with K-means Algorithm, D-S (Dempster-Shafer evidence theory) evidence theory, and Bayesian algorithm and verify the performance of the algorithm by clustering accuracy, purity, relevance, and time consumption.

In the experiment, we will compare and analyze the performance of the BigDataACO algorithm and other comparison algorithms proposed in this chapter on the three data sets of Agricultural Features Data, Agricultural Multilanguage Data, and Agricultural Multimedia Data. The Agricultural Features Data contains image characteristics of nine handwritten characters, each of which has 200 images and a total of 1800 images. Each picture can be represented as a 122-dimensional character shape Fourier coefficient, a 221-dimensional contour description, a 260-dimensional pixel average, and a 16-dimensional morphological feature. There are 12 kinds of visual features, that is, 12 modes of data. In the experiment, the first five modal feature sets were used for multimodal data clustering analysis. Table 1 gives a brief description of several data sets. The clustering distribution of multimodal data before data fusion is shown in Figure 2.

In the experiment, we compare the BigDataACO algorithm with K-means, D-S evidence theory, and Bayesian algorithm. Ant algorithm is a new simulated evolutionary algorithm based on population. K-means is a common clustering method based on segmentation. In this paper, the ant algorithm and K-means algorithm are combined to solve the local optimization problem to a large extent by using the randomness of the ant algorithm and overcome the sensitivity of initial parameters of the K-means algorithm. It improves the quality of clustering and overcomes the problem that the density-based algorithm cannot find arbitrary

shape clustering. And Bayesian classification model is a simple and effective classification method, which has a good theoretical foundation and high classification accuracy. Because of the independent hypothesis premise in naive Bayesian classification, it is particularly important whether the K-means feature selection step can accurately and effectively classify. D-S theory is a generalization of Bayesian reasoning method, which mainly uses Bayesian conditional probability in probability theory, and needs to know prior probability. However, D-S evidence theory does not need to know the prior probability, which can well express "uncertainty," and is widely used to deal with uncertain data.

In order to analyze and compare the performance of each multimodal clustering algorithm more comprehensively, we use the accuracy, correlation, and time consumption to measure and analyze all clustering results.

4.2. Experimental Environment. All the experiments are performed on the same PC (personal computer). The hardware configuration is as follows: Intel Core i7-7000U processor, 2.80 GHz main frequency, 32 GB memory; and the software use MATLAB2012. The data in each experimental data set is randomly divided into five parts. The first part contains 30% of the data in the whole data set, and all of them have labels. In the experiment, each clustering model is initialized with the labels. The remaining data are divided into three blocks, and all of them are labeled-free. In the experiment, it is added three times for incremental clustering fusion.

4.3. Experimental Relation Algorithm Description. The first set of experiments will verify the clustering performance of the four algorithms in the three data sets of Agricultural

TABLE 1: Experimental data set.

Name	Number of instances	Modal number	Feature dimension	Multiple source classes
Agricultural Features Data	2000	6	121/232/24/324/6	12
Agricultural Multilanguage Data	8000	4	2100/2100/4200/2100/2100	6
Agricultural Multimedia Data	260009	3	600/1000/2000	121
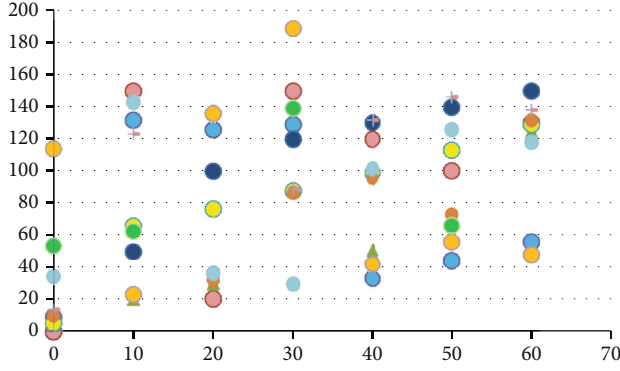
FIGURE 2: The clustering distribution of multimodal data before data fusion.

FIGURE 3: Comparisons of four different data fusion clustering algorithms.

Features Data, Agricultural Multilanguage Data, and Agricultural Multimedia Data. In the specific experiment, 1200 data instances were randomly selected for label clustering mode initialization, and then, other unlabeled data were equally divided into three in random order to join the existing clustering results to complete the incremental clustering. For the K-means and BigDataACO algorithms, the initial cluster number set [5, 6, 12, 13, 16] five different values to complete the experiment, the DS evidence theory and the Bayesian algorithm's iteration number and shared feature dimension are set to [121, 260, 400, 6, 8, 12]. Each experiment was performed 20 times randomly, and the cluster average results were recorded. The specific experimental results of each algorithm are compared as shown in Figure 3.

4.4. Experimental Results and Analysis. As can be seen from Figure 3, with the addition of the amount of each data block, the accuracy, correlation, and execution time of most algorithms decrease, but the execution time of the algorithms increases significantly.

The K-means algorithm may find the local optimal clustering, rather than the global optimal clustering. However, the ACO and Bayesian algorithm can achieve global optimal clustering. The original K-means algorithm calculates the distance between each observation point and all cluster centers in each iteration. When the number of observation points is large, the performance of the algorithm is not good. The accuracy of Bayesian classification is the highest, and D-S theory is a generalization of Bayesian reasoning method, which mainly uses Bayesian conditional probability in probability theory, and needs to know the prior probability. D-S evidence theory does not need to know the prior probability, which can express "uncertainty" well, and is widely used to deal with uncertain data. It is mainly suitable for information fusion, expert system, intelligence analysis, legal case analy-
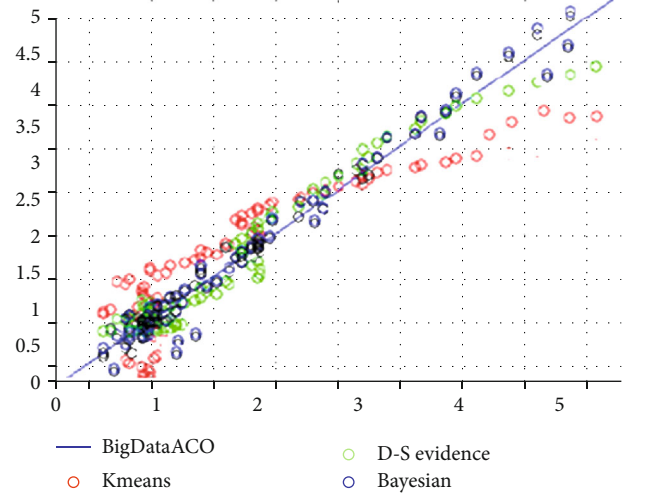
sis, and multiattribute decision analysis as an uncertain reasoning method.

In the performance comparison index, we use SSE (sum of square due to error) to represent the sum of the positions from the center point of the current classification situation to the points of its own cluster. The calculation formula of SSE can be written as

$$SSE = \sum_{i=1}^{k} \sum_{p \in c_i} |p - m_i|^2, \tag{17}$$

where p is the position of the point (x, y) and m_i is the position of the center point.

In the process of clustering algorithm iteration, we evaluate the current classification effect by calculating the SSE value under the current center point. If the SSE value is greatly reduced after an iteration, the clustering process is basically completed, and there is no need for many iterations. Compared with the K-means, D-S evidence theory, and Bayesian algorithm, the SSE value of the BigDataACO algorithm is the smallest, while that of the K-means, D-S evidence theory, and Bayesian algorithm is larger. It can also be seen from Figure 4 that the BigDataACO algorithm has the best clustering performance, and the clustering performance is relatively stable with the dynamic change of data. The implementation of the D-S evidence theory and Bayesian algorithmic algorithm under the best parameter setting also has no incremental data processing capability, so it has a similar time performance with the K-means algorithm.
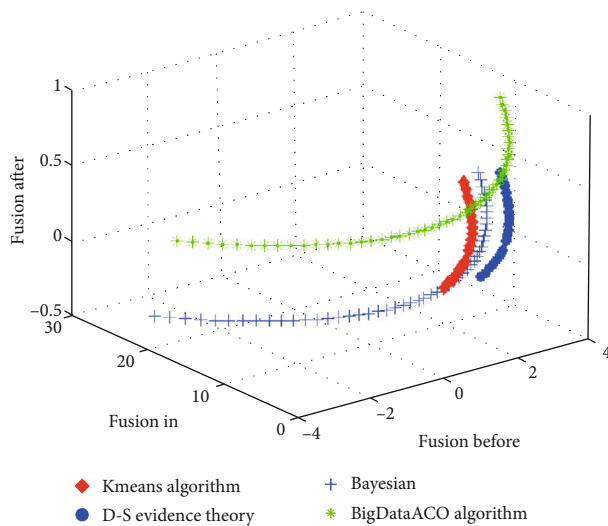
FIGURE 4: Data Fusion Results Using Different Algorithms.

5. Conclusions

In this paper, an improved BigDataACO algorithm based on ACO is proposed, that is, taking agricultural large data fusion as the research object; this paper studies the construction and prediction methods of classification models for different data sets by using an improved ant colony algorithm. In the wireless sensor monitoring of large agricultural data, data fusion technology can be combined with multiprotocol levels of sensor networks. The real-time monitoring data of sensors with certain uncertainty and ambiguity and soil moisture retrieved by hyperspectral data are used as agricultural large data sets.

In the process of fusion, the improved ant colony optimization algorithm and Bayesian maximum entropy method are used to complete the integration of the three data sets of Agricultural Features Data, Agricultural Multilanguage Data, and Agricultural Multimedia Data at the regional scale. On this basis, the improved BigDataACO algorithm is used to complete the fusion of multisource information in the data set, solve the problem of information fusion in the process of agricultural management and decision-making, and eliminate the possible redundancy between multisource agricultural information. Contradictions have improved the reliability of agricultural decision-making and the utilization of agricultural big data information. Further studies are expected to understand the connotation of the problem of big data fusion; in the era of big data, the analysis and mining for the Agricultural Multilanguage Data is a research field and which attracts much attention. To effectively learn the characteristics of massive, low-quality, heterogeneous, high-dimensional, and fast-changing big data, there are still a series of problems and challenges. Our study provides a corresponding Agricultural Multilanguage Data fusion algorithm for the incompleteness of multimodal data, real-time processing, and multisource data fusion. The purpose of this paper is to propose the organization and blockchain-enabled security method of big data to realize the mining and analysis of monitoring data of the Internet of things.

References

[1] B. Zhou, D. Polap, and M. Wozniak, "A regional adaptive variational PDE model for computed tomography image reconstruction," *Pattern Recognition*, vol. 92, no. 1, pp. 64–81, 2019.

[2] X. Xia, M. Marcin, and X. Fan, "Multi-sink distributed power control algorithm for cyber-physical-systems in coal mine tunnels," *Computer Networks*, vol. 161, no. 1, pp. 210–219, 2019.

[3] H. Song, W. Li, and P. Shen, "Gradient-driven parkingnavigation using a continuous information potential field based on wireless sensor network," *Information Sciences*, vol. 408, no. 2, pp. 100–114, 2017.

[4] Q. Xu, L. Wang, and X. H. Hei, "GI/Geom/1 queue based on communication model for mesh networks," *International Journal of Communication Systems*, vol. 27, no. 11, pp. 3013–3029, 2014.

[5] W. Wei, X. Fan, H. Song, X. Fan, and J. Yang, "Imperfect information dynamic Stackelberg Game based resource allocation using hidden Markov for cloud computing," *IEEE Transactions on Services Computing*, vol. 11, no. 1, pp. 78–89, 2018.

[6] F. Zhang, H. F. Xue, D. S. Xu, Y. H. Zhang, and F. You, "Big data cleaning algorithms in cloud computing," *International Journal of Online Engineering (iJOE)*, vol. 9, no. 3, pp. 77–81, 2013.

[7] X. Li, L. Zhang, X. Cao et al., "Retrieval of precipitable water vapor using MFRSR and comparison with other multisensors over the semi-arid area of northwest China," *Atmospheric Research*, vol. 172-173, pp. 83–94, 2016.

[8] O. P. Rahi, A. K. Chandel, and M. G. Sharma, "Optimization of hydro power plant design by particle swarm optimization (PSO)," *Procedia Engineering*, vol. 30, pp. 418–425, 2012.

[9] F. Zhang, H. F. Xue, and J. C. Zhang, "Multi-source big data dynamic compressive sensing and optimization method for water resources based on IoT," *Sustainable Computing: Informatics and Systems*, vol. 20, no. 1, pp. 210–219, 2018.

[10] S. Singh and Y. Liu, "A cloud service architecture for analyzing big monitoring data," *Tsinghua Science and Technology*, vol. 21, no. 1, pp. 55–70, 2016.

[11] D. S. Yang, J. J. Chen, and M. Zhang, "Research and application on key technology of high-speed storage and retrieval for big data," *Electronic Testing*, vol. 3, pp. 62-63, 2014.

[12] H. Wu, J. Wang, and X. Zhang, "Combining hidden Markov model and fuzzy neural network for continuous recognition of complex dynamic gestures," *The Visual Computer*, vol. 33, no. 10, pp. 1265–1278, 2015.

[13] G. Reeves, J. Liu, S. Nath, and F. Zhao, "Managing massive time series streams with multi-scale compressed trickles," *Proceedings of the VLDB Endowment*, vol. 2, no. 1, pp. 97–108, 2009.

[14] J. Huang, J. Liu, and X. Yao, "A multi-agent evolutionary algorithm for software module clustering problems," *Soft Computing*, vol. 21, no. 12, pp. 3415–3428, 2017.

[15] Y. Wang and J. C. Zeng, "A survey of a multi-objective particle swarm optimization algorithm," *CAAI Transactions on Intelligent Systems*, vol. 5, no. 5, pp. 377–384, 2010.

[16] N. Li, T. Zou, D. B. Sun, and Y. Q. Qin, "Multi-objective optimization utilizing particle swarm," *Computer Engineering and Applications*, vol. 3, no. 4, pp. 193–262, 2005.

[17] Y. Zhou, D. Zhang, and N. Xiong, "Post-cloud computing paradigms: a survey and comparison," *Tsinghua Science and Technology*, vol. 22, no. 6, pp. 714–732, 2017.

[18] H. Liang, J. Zou, K. Zuo, and M. J. Khan, "An improved genetic algorithm optimization fuzzy controller applied to the wellhead back pressure control system," *Mechanical Systems and Signal Processing*, vol. 142, article 106708, 2020.

[19] Z. Huang, X. Xu, J. Ni, H. Zhu, and C. Wang, "Multimodal representation learning for recommendation in internet of things," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10675–10685, 2019.

[20] H. Liang, J. Zou, Z. Li, M. J. Khan, and Y. Lu, "Dynamic evaluation of drilling leakage risk based on fuzzy theory and PSO-SVR algorithm," *Future Generation Computer Systems*, vol. 95, pp. 454–466, 2019.

[21] A. Shahzad, M. Lee, Y. K. Lee et al., "Real time MODBUS transmissions and cryptography security designs and enhancements of protocol sensitive information," *Symmetry*, vol. 7, no. 3, pp. 1176–1210, 2015.

[22] Y. Zhang, R. Zhu, Z. Chen, J. Gao, and D. Xia, "Evaluating and selecting features via information theoretic lower bounds of feature inner correlations for high-dimensional data," *European Journal of Operational Research*, 2020.

[23] H. Liang, A. Xian, M. Mao, P. Ni, and H. Wu, "A research on remote fracturing monitoring and decision-making method supporting smart city," *Sustainable Cities and Society*, vol. 62, article 102414, 2020.