

Brain Imaging Big Data Mining and Fusion Method Based on Cognitive Intelligence

Xiao-Yan Li

Yulin University

Feng Zhang (✉ 896316414@qq.com)

Yulin University

Yong-Heng Zhang

Yulin University

Research

Keywords: brain imaging data, big data fusion, information security, data mining, cognitive intelligence

Posted Date: June 18th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-36155/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Brain Imaging Big Data Mining and Fusion Method Based on Cognitive Intelligence

Xiao-Yan Li^a, Feng Zhang^{b*}, Yong-Heng Zhang^b

^a Mathematics and Statistics, Yulin University, Yulin 719000, Shaanxi, China

^bSchool of Information Engineering, Yulin University, 719000, Yulin, China

Abstract

In order to improve the utilization rate of brain imaging big data and solve the fusion problem of multi-source and heterogeneous brain imaging big data, an improved brain imaging big data ant colony optimization algorithm (BigDataACO) is proposed to complete the multi-source brain imaging big data information in the feature layer and decision-making and the problem of multi-source 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 brain imaging big data fusion problem are studied. Then the insufficiency of big data fusion modeling algorithm is analyzed. Finally, the source and core steps of 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.

Keywords: brain imaging data, big data fusion, information security, data mining, cognitive intelligence

1. Introduction

Brain imaging 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]. Brain imaging 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 brain imaging is still relatively small [2].

In order to continuously promote the optimization of the brain imaging economy, to realize the sustainable industrial development and regional industrial structure optimization, and further promote the construction of cognitive intelligence for brain imaging, it is necessary to comprehensively and timely grasp the development of brain imaging, which needs to rely on brain imaging big data and related big data fusion processing technology. However, it's face 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 (ie, through the classification algorithm), so the fusion rule set that best represents the training data is found. That is a process of gradual optimization [3, 4]; 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 swarm intelligence algorithm has become a research hotspot in the field of data mining in recent years [5].

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

* Corresponding author.

E-mail address: 896316414@qq.com.

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 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 multi-level, multi-faceted, multi-level information detection, and correlation estimation and correlation analysis on various multi-source and heterogeneous data under certain criteria [5, 6], 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 [7]. 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 brain imaging 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, easy to combine with other algorithms, people have applied it to solve many complex combinatorial optimization problems, and have shown great potential [8].

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 [9]. At the individual level, each ant makes independent choices based on the environment. At the group level, the behaviours 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 [10, 11]. Ant colony algorithm is actually a class of multi-subject 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 [12, 13]. 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 [14]:

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

In formula (1), s_{ij} is the value of the information on the edge (i, j) , τ represents the post-latency 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 [15, 16]. 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).

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

In formula (2), ρ is a constant whose value ranges from $\rho \in [0,1]$, indicates 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 $\Delta\tau'_{ij}(t)$ seeking.

In the Ant-Cycle model,

$$\Delta\tau'_{ij}(t) = \begin{cases} \frac{Q}{L_k} & , \text{If the } k \text{ ant passes through } i, j \text{ in this cycle} \\ 0 & , \text{else} \end{cases} \quad (3)$$

In equation (3), Q represents the pheromone intensity, which affects the convergence speed of the algorithm to some extent, L_k represents the total length of the path taken by the k th ant in this cycle.

In the Ant-Quantity model,

$$\Delta\tau'_{ij}(t) = \begin{cases} \frac{Q}{d_{ij}} & , \text{If the } k \text{ ant passes through } i, j \text{ between } t \text{ and } t+1 \\ 0 & , \text{else} \end{cases} \quad (4)$$

In the Ant-Density model,

$$\Delta\tau'_{ij}(t) = \begin{cases} Q & , \text{If the } k \text{ ant passes through } i, j \text{ between } t \text{ and } t+1 \\ 0 & , \text{else} \end{cases} \quad (5)$$

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 [17, 18]. 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).

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

$$\eta(i, j) = \frac{1}{d(i, j)} \alpha \beta \quad (6)$$

Where $N_k(i)$ represents a collection of nodes that the node has not yet accessed; τ represents a pheromone, $\tau(i, j)$ represents the amount of information on the path between nodes i and j ; in $\eta(i, j) = \frac{1}{d(i, j)}$, $\eta(i, j)$ takes the reciprocal of the distance between nodes, 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:

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

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

$$\Delta\tau_{ij} = [\epsilon + (h_i - h_j)] \times \Delta\omega_j \quad (8)$$

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

Where, 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).

$$\Delta\tau(i, j) = \begin{cases} 0; (i, j) \notin \rho \\ \frac{Q}{L_k}; (i, j) \in \rho \end{cases} \quad (10)$$

Where 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:

IF <conditions> THEN <class>

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

<term₁ and term₂ and ...and term_n >

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 pre-category 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, 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:

$$H_{ij} = \frac{\tau_{ij}(t)\eta_{ij}}{\sum_{i=1}^n x_i \sum_j^{b_j} \tau_{ij(t)}\eta_{ij}} \quad (11)$$

Where 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 i -th attributes range; η_{ij} is the problem-dependent heuristic function of item term_{ij}, which is calculated as:

$$\eta_{ij} = \frac{\log_2 k - B(T | G_i = D_{ij})}{\sum_{i=1}^n x_i \sum_{j=1}^{b_i} \log_2 k - B(T | G_i = D_{ij})} \quad (12)$$

Where $B(T | G_i = D_{ij})$ 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 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 (13):

$$\tau_{ij}(t=0) = \frac{1}{\sum_{i=1}^n b_i} \quad (13)$$

In formula (13), 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:

$$H_{ij}(t) = \frac{\tau_{ij}(t)\eta_{ij}}{\sum_{i=1}^n \sum_{j=1}^{b_i} \tau_{ij}(t)\eta_{ij}}, \forall i \in I \quad (14)$$

Wherein, 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.

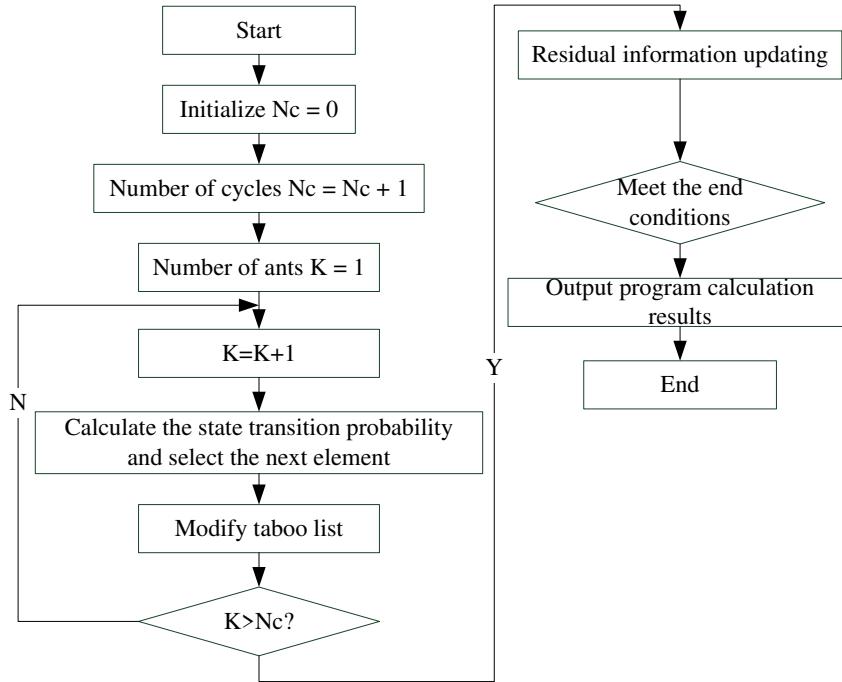


Figure 1. The improved ant colony data fusion algorithm flow

4. Experimental analysis

This experiment analyses and verifies the performance of the BigDataACO algorithm on multi-source big data sets. In order to give a more intuitive analysis for the performance of the algorithm, we compare the BigDataACO algorithm with K-means Algorithm, D-S 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 Brain imaging Features Data, Brain imaging Multilanguage Data and Brain imaging Multimedia Data. The Brain imaging 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 multi-modal data before data fusion is shown in Figure 2.

Table 1. Experimental data set

Name	Number of instances	Modal number	Feature dimension	Multiple source classes
Brain imaging Features Data	2000	6	121/232/24/324/6	12
Brain imaging Multilanguage Data	8000	4	2100/2100/4200/2100/2100	6
Brain imaging Multimedia Data	260009	3	600/1000/2000	121

In the experiment, we compare BigDataACO algorithm with K-means, D-S evidence theory and Bayesian algorithm.

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

All the experiments are performed on the same PC. The hardware configuration is as follows: Intel Core i7-7000U processor, 2.80GHz main frequency, 32GB 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 mode 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.

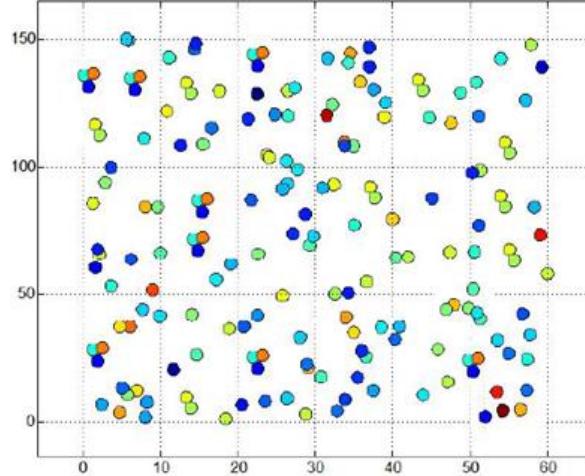


Figure 2. The clustering distribution of multi-modal data before data fusion

The first set of experiments will verify the clustering performance of the four algorithms in the three data sets of Brain imaging Features Data, Brain imaging Multilanguage Data and Brain imaging 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 [6, 5, 12, 16, 13] 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] and [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.

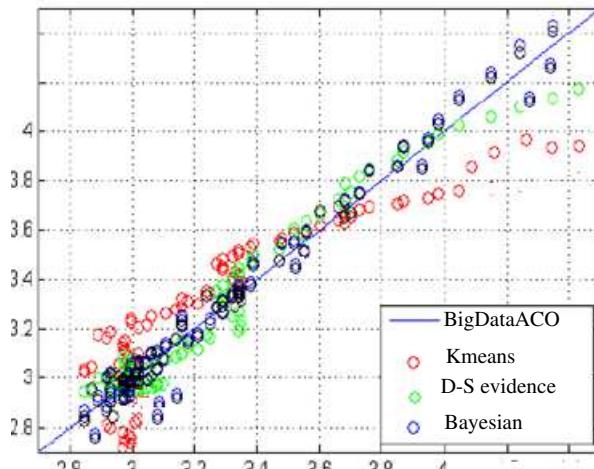


Figure 3. Comparisons of Four Different Data Fusion Clustering Algorithms

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.

It can also be seen from Figure 4 that BigDataACO algorithm has the best clustering performance, and the clustering performance is relatively stable with the dynamic change of data. The implementation of D-S evidence theory and Bayesian algorithmic algorithm under the best parameter setting also has no incremental data processing capability, so it has similar time performance with K-means algorithm.

5. Conclusions

In this paper, an improved BigDataACO algorithm based on ACO is proposed, that is, taking brain imaging large data fusion as the research object, this paper studies the construction and prediction methods of classification models for different data sets by using improved ant colony algorithm. In the wireless sensor monitoring of large brain imaging data, data fusion technology can be combined with multi-protocol 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 brain imaging large data sets.

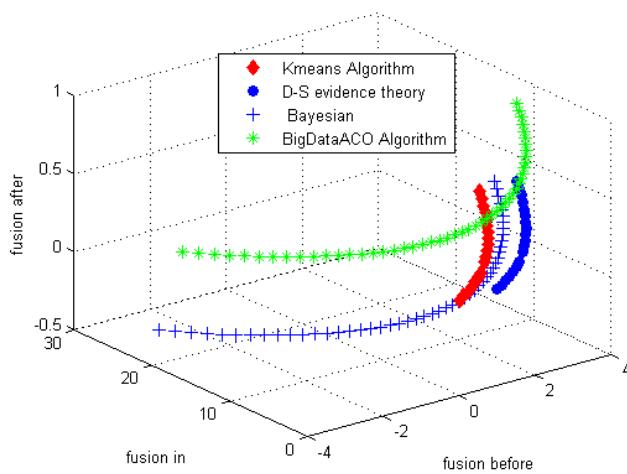


Figure 4. Data Fusion Results Using Different Algorithms

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 brain imaging features data, brain imaging Multilanguage data and brain imaging Multimedia data at the regional scale; On this basis, the improved BigDataACO algorithm is used to complete the fusion of multi-source information in the data set, solve the problem of information fusion in the process of brain imaging management and decision-making, and eliminate the possible redundancy between multi-source brain imaging information. Contradictions have improved the reliability of brain imaging decision-making and the utilization of brain imaging 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 brain imaging 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 brain imaging Multilanguage data fusion algorithm for the incompleteness of multimodal data, real-time processing and multi-source data fusion.

Declarations

Availability of data and materials

All data, models, and code generated or used during the study appear in the submitted manuscript.

Competing interests

All the authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

Funding

This work is partially supported by the Funding Project for Department of Education of Shaanxi Province in 2018 (No. 18JK0906); Natural Science and Technology Project Plan in Yulin of China (2016CXY-12,2019-78-2, 2019-78-1, 2019-76-2, 2016CXY-09); Funding Project for Department of Yulin University (16GK24,TZRC1801).

Authors' contributions

All the authors have contributed and have taken part in verifying and revising this manuscript.

Acknowledgements

We would like to thank everyone for their help.

References

1. Çanakçı H, Baykasoglu A, Güllü H, "Prediction of compressive and tensile strength of Gaziantep basalts via neural networks and gene expression programming," *Neural Computing and Applications*, vol. 18, no. 8, pp. 1031-1041, 2009
2. F Zhang, H-F Xue, D-S Xu, "Big data cleaning algorithms in cloud computing," *International Journal of Online Engineering*, vol.9, no. 3, pp. 77-81, 2013

3. X Li , L Zhang and 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.17, no. 1, pp. 83-94, 2013
4. Li Zhao-Xing, "A multi objective optimization algorithm for recommender system based on PSO," *Computer Modelling and New Technologies*, vol.18, no. 7, pp. 231-235, 2014
5. Alícia Pou, Maria P. Diago, Hipólito Medrano, et al, "Validation of thermal indices for water status identification in grapevine," *Brain imaging Water Management*, vol.134, no. 1, pp. 60-72, 2014
6. Rahi, O.P, "Optimization of hydro power plant design by Particle Swarm Optimization (PSO)," *Procedia Engineering*, vol.30, no. 1, pp. 418-425, 2011
7. Dorigo M, Maniezzo V, Colorni A, "Ant system: optimization by a colony of cooperating agents," *IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man & Cybernetics Society*, vol.26, no. 1, pp. 29-41, 1996
8. Donkal, Gita,Verma, Gyanendra K, "A multimodal fusion based framework to reinforce IDS for securing Big Data environment using Spark," *Journal of Information Security and Applications*, vol.43, no. 1, pp. 1-11, 2018
9. Zhang Feng, Xue Hui-feng,Zhang Jing-Cheng, "Multi-source Big data Dynamic Compressive Sensing and Optimization Method for Water Resources Based on IoT," *Sustainable Computing: Informatics and Systems*, Vol. 20, pp. 210-219,2018
10. S. Singh and Y. Liu, "A Cloud Service Architecture for Analyzing Big Monitoring Data," *Tsinghua Science & Technology*, vol. 21, no. 1, pp. 50-70, 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. D. Lu, Y. Yu, and H. Liu, "Gesture Recognition Using Data Glove: An Extreme Learning Machine Method," in *Proceedings of IEEE International Conference on Robotics and Biomimetics*, pp. 1349-1354, Qingdao, China, March 2017
13. H. Wu, J. Wang, and X. Zhang, "Combining Hidden Markov Model and Fuzzy Neural Network for Continuous Recognition of Complex Dynamic Gestures," *Visual Computer*, pp. 1-14, August 2015
14. G. Reeves, J. Liu, and S. Nath, "Managing Massive Time Series Streams with Multi-scale Compressed Trickles," *Proceedings of the Vldb Endowment*, vol. 2, no. 1, pp. 97-108, 2009
15. J. H. Huang and X. Yao, "A Multi-Agent Evolutionary Algorithm for Software Module Clustering Problems," *Soft Computing*, Vol. 21, No. 13, pp. 3415-3428, 2017
16. K. Praditwong and X. Yao, "A New Multi-Objective Evolutionary Optimization Algorithm: The Two-Archive Algorithm," in *Proceedings of the 2006 International Conference on Computational Intelligence and Security*, Vol. 1, No. 4, pp. 286-291, 2015
17. 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
18. 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

Figures

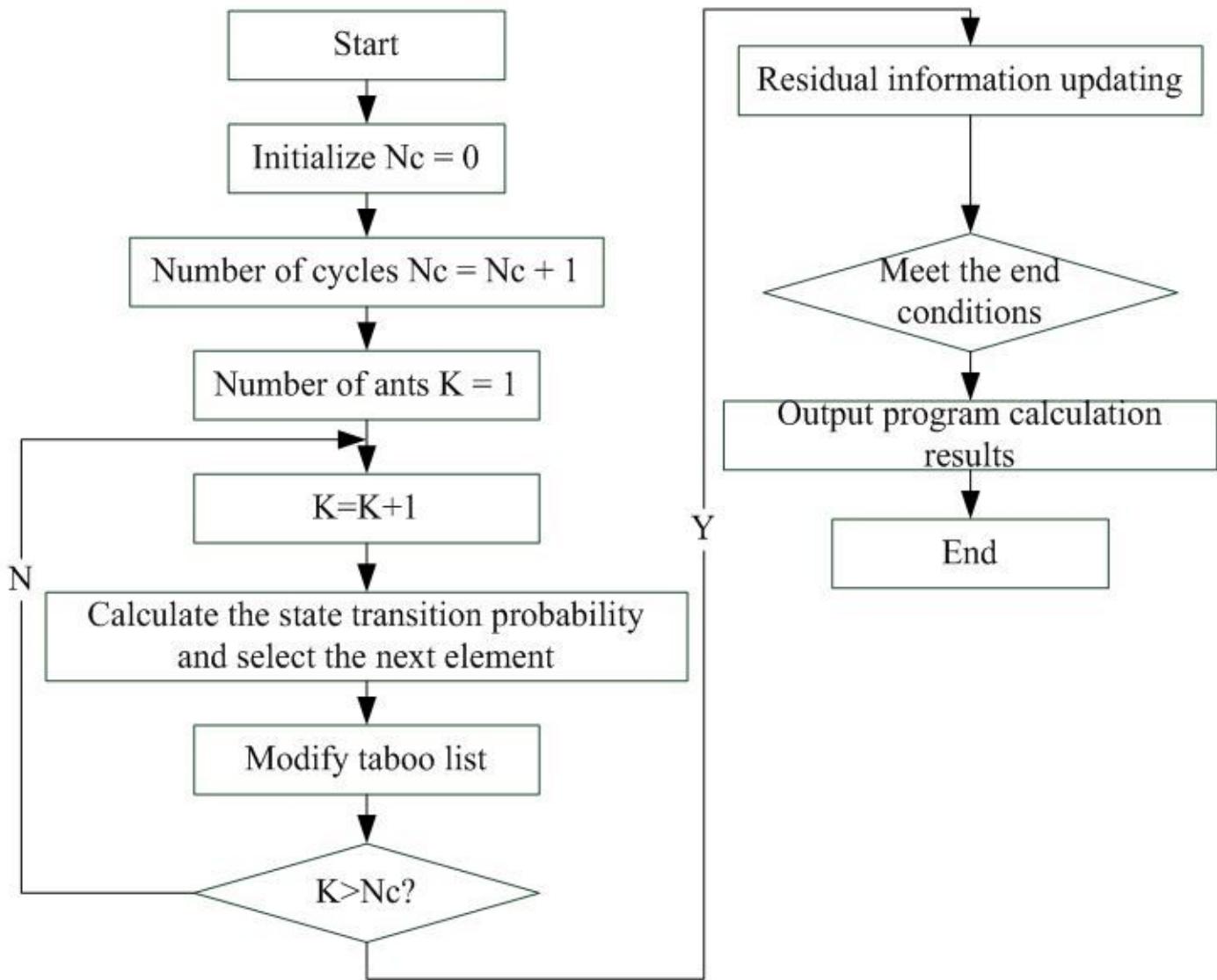


Figure 1

The improved ant colony data fusion algorithm flow.

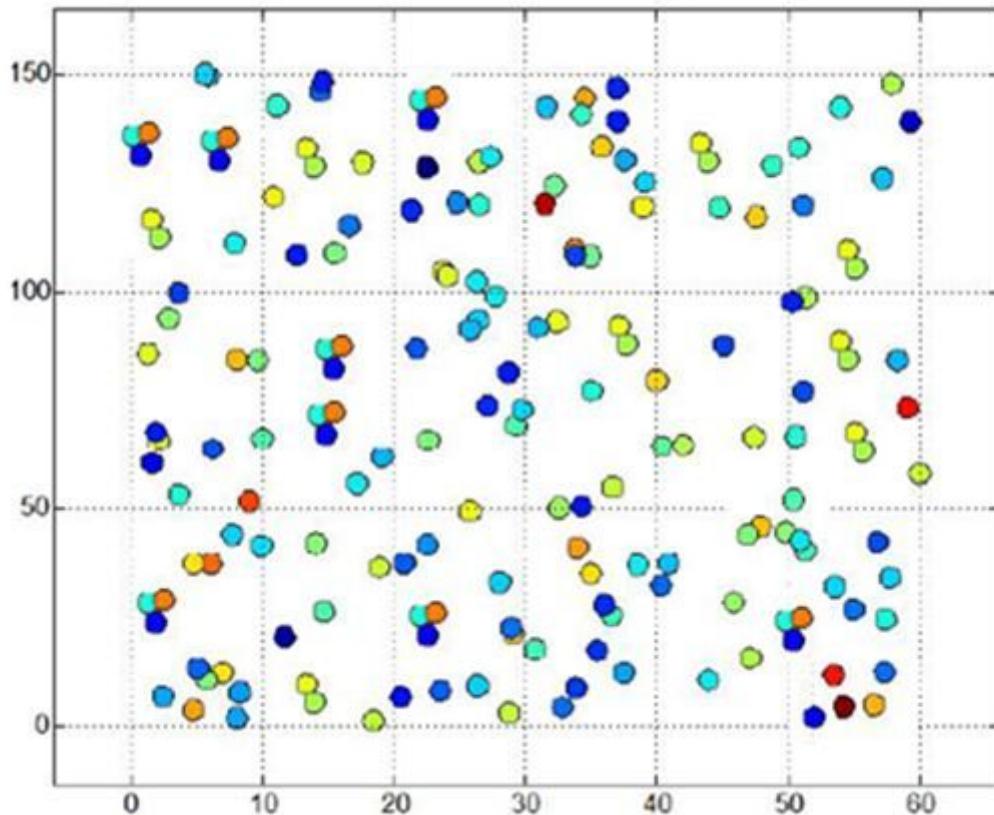


Figure 2

The clustering distribution of multi-modal data before data fusion.

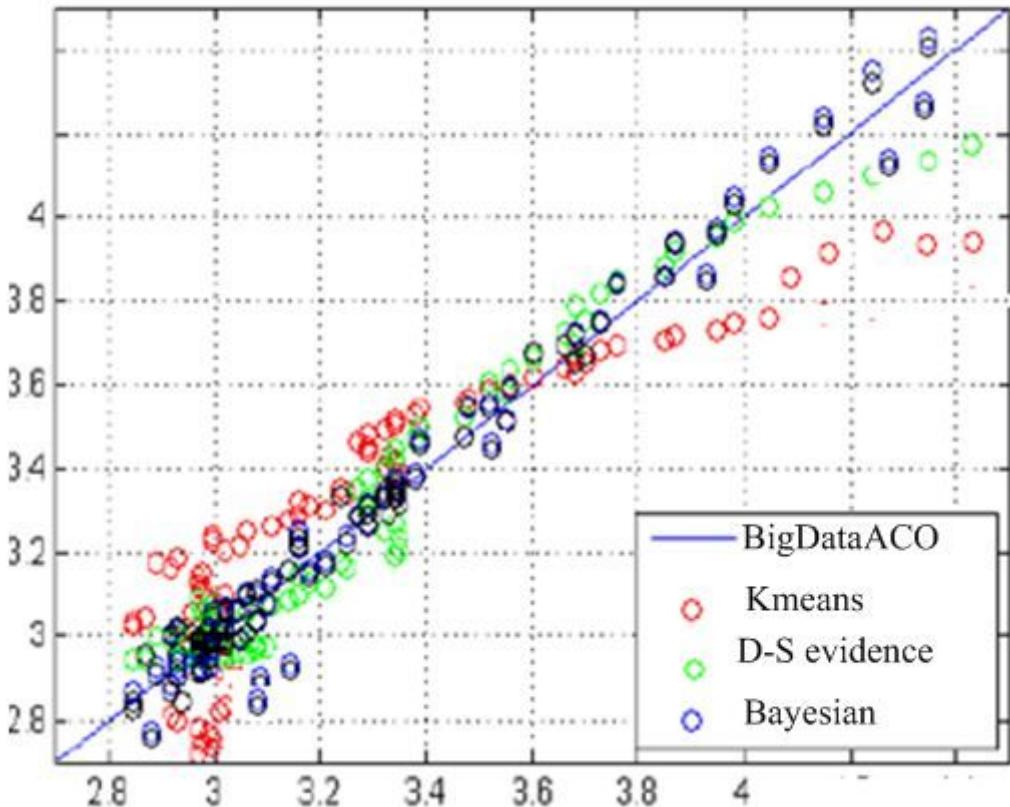


Figure 3

Comparisons of Four Different Data Fusion Clustering Algorithms.

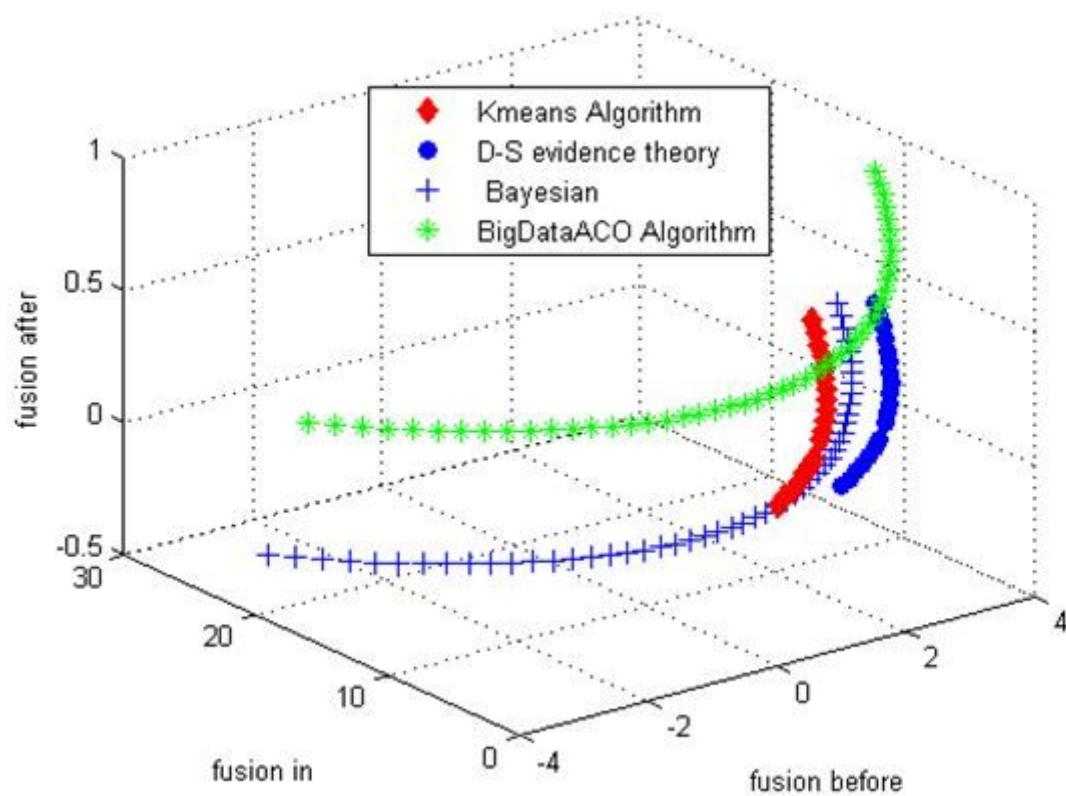


Figure 4

Data Fusion Results Using Different Algorithms.