Journal of Information Hiding and Privacy Protection
DOI:10.32604/jihpp.2020.010362
Article

A Clustering Method Based on Brain Storm Optimization Algorithm

Tianyu Wang, Yu Xue, Yan Zhao, Yuxiang Wang*, Yan Zhang and Yuxiang He

Nanjing University of Information Science & Technology, Nanjing, 210044, China
*Corresponding Author: Yuxiang Wang. Email: wangyx@nuist.edu.cn
Received: 02 August 2020; Accepted: 29 August 2020

Abstract: In the field of data mining and machine learning, clustering is a typical issue which has been widely studied by many researchers, and lots of effective algorithms have been proposed, including K-means, fuzzy c-means (FCM) and DBSCAN. However, the traditional clustering methods are easily trapped into local optimum. Thus, many evolutionary-based clustering methods have been investigated. Considering the effectiveness of brain storm optimization (BSO) in increasing the diversity while the diversity optimization is performed, in this paper, we propose a new clustering model based on BSO to use the global ability of BSO. In our experiment, we apply the novel binary model to solve the problem. During the period of processing data, BSO was mainly utilized for iteration. Also, in the process of K-means, we set the more appropriate parameters selected to match it greatly. Four datasets were used in our experiment. In our model, BSO was first introduced in solving the clustering problem. With the algorithm running on each dataset repeatedly, our experimental results have obtained good convergence and diversity. In addition, by comparing the results with other clustering models, the BSO clustering model also guarantees high accuracy. Therefore, from many aspects, the simulation results show that the model of this paper has good performance.

Keywords: Clustering method; brain storm optimization algorithm (BSO); evolutionary clustering algorithm; data mining

1 Introduction

Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups. In simple words, the aim is to segregate groups with similar traits and assign them into clusters. Whether for understanding or utility, clustering analysis has long played an important role in a wide variety of fields: Psychology and other social sciences, biology, statistics, pattern recognition, information retrieval, machine learning, and data mining.

A typical clustering process mainly includes steps such as data (or sample or pattern) preparation, feature selection and feature extraction, proximity calculation, clustering (or grouping), and effectiveness evaluation of clustering results [1].

Although clustering analysis has a history of decades of research, many clustering algorithms have been proposed and related applications have been developed, but the clustering problem still has huge challenges.

There is no clustering technique (clustering algorithm) can be universally applied to reveal the various structures presented by various multidimensional data sets [2]. According to the data accumulation rules in clustering and the methods to apply these rules There are multiple clustering algorithms. There is multiple classification for clustering algorithms, like partitioned clustering, grid-based clustering, hierarchical clustering and other [3].

However, different types of clustering are similar and interoperable in some ways. Most clustering algorithms require parameters to be given in advance. In fact, this is not feasible in most cases without relevant knowledge and experience [4].

In recent years, the research on clustering has risen to a higher level. A large number of new clustering methods have emerged, such as partition-based fuzzy clustering analysis [5], neighbor propagation clustering analysis, sparse subspace clustering methods, and robust unsupervised clustering.

As one of the important problems of optimization, Clustering is similar to classification in that data are grouped to a certain degree. In the field of data mining and machine learning, for solving large datasets, many researchers have created a lot of clustering algorithm to find the optimal solution or satisfying solution. The purpose of it is to classify things that are similar in some aspects in order to find patterns from them, and better predict the development trend of objective things. Sometimes we call clustering the unsupervised classification, and some algorithms including K-means, Evolutionary Algorithm (EA), FCM and Particle Swarm Optimization (PSO) were created for it, which solved many practical problems effectively [6].

For clustering, unlike classification, the groups we get usually are not predefined and it is a process of unsupervised learning. Clustering algorithms attempt to divide the samples in the given datasets into several disjoint subsets, each of which is called a "cluster". Two definitions for clusters have been proposed: Set of like elements or elements from different clusters are not alike. And the distance between points in a cluster is less than the distance between a point in the cluster and any point outside it [7].

Due to the effectiveness and the convenience of the clustering method, many researchers have used it in many application domains, including economics, customer classification, anthropology, and biology. In addition, for different practical problems, the researchers have created clustering algorithms that adapt to different models. Traditional clustering algorithms are mainly single-edge based, for example, gradient decent algorithms which move from the present position along the opposite direction of the function's gradient at this very point. Traditional optimization algorithms are mainly good at solving unimodal problems [8].

However, when many clustering method is applied to a real-world database, there exist some long-lasting problems: Lots of them may well fall into local optimum (The result of clustering often falls into local optimal solution as their uncertainty and randomness [9]).

Among traditional population based optimization algorithms, PSO-oriented feature selection algorithms have shown satisfactory results because of their simple structure and fast convergence [10]. Brain storm optimization (BSO) is a type of swarm intelligence (SI)-based optimization algorithm that imitates the human brainstorming process. Due to its flexibility and conciseness, BSO has produced satisfactory results in solving various optimization problems.

There are many applications in which the optimization variables are discrete or could be better represented by a binary string. Hence, a novel binary version of BSO (BBSO) was also proposed and applied to serve for an array thinning design and a dual-band microstrip patch antenna design [11]. In addition, some researchers proposed multiple modifications to improve the performance of BSO. These modifications include adopting a fitness based grouping mechanism, using the global-best idea information for updating the population, and applying the update scheme on every problem variable respectively. It was called Global-best BSO (GBSO) [12]. Actually, the BBSO and GBSO also hit us to proposed the new model in our paper.

In recent years, as an optimization algorithm, BSO was widely used in solving some of the supervised classification problems. Considering it is significant advantages like clear and effectiveness, we proposed a new model based on BSO. By finding local optimal clustering in local optimum anchored in a new generation via population variation [13].

Clustering is a characteristic of BSO, and it has also been a research hotspot for many BSO researchers. In many studies, it is mainly based on the study of clustering methods and clustering algorithms. For example, the clustering algorithm in the initial BSO is k-means clustering. However, the analysis of individual diversity and the number of clusters has few results. In this paper, the solution space is divided into a central solution and a common solution through adaptive parameter adjustment, which saves a lot of workload of clustering. However, the effectiveness of the clustering operation in the clustering algorithm and the selection of the number of clusters need further study.

In this paper, we focus on introducing BSO to process datasets to solve the unsupervised clustering problems. In addition, four datasets were applied in our experiment, getting the satisfactory result. Also, we will conduct the further experiment to optimize the clustering model.

2 Brain Storm Optimization Algorithm

The BSO algorithm is inspired by the above-mentioned brainstorming model. As an simple and effective swarm intelligence algorithm, BSO features the process of clustering and population variation. For different practical problems, its different implementation process parameters may play different roles. Each parameter may well influence the optimal results. On the one hand, in the process of clustering, BSO applies K-means to find local optimum. On the other hand, in the process of population variation, BSO produce new individuals via some methods of mutation on the basis of local optimum [14]. That means clustering and variation play the key role in BSO algorithm.

Considering BSO is mainly composed of the clustering part and the variation part, we can recognize that the algorithm process ensure the model's diversity and accuracy. At the beginning of the algorithmic process, BSO initialed the population and used clustering method to gather individuals for the K clusters. Each cluster center represented the local optimal value of the cluster. And in the process of update and variation, the algorithm optimized the information through learning to promote local search [15]. The algorithm enhances the diversity of the group and reduces the coupling degree between the classes through the fusion mutation operation of individuals and inter-class individuals to promote global search.

BSO algorithm [16] is shown in Algorithm 1.

Algorithm 1: Procedure of BSO

1. Randomly generate n individuals to initialize the population;
2. Evaluate the adaptive value of each individual;
3. Cluster n individuals into m clusters by clustering algorithm;
4. Comparing the adaptive values of individuals in each cluster, and placing the best individuals in the cluster as the cluster center;
5. Randomly generate a number between (0, 1);
   a) If the value is less than the predefined probability parameter p_{5a}:
      i. Randomly select a cluster center;
      ii. Randomly generate an individual to replace the cluster center;(This is a one-step divergence operation, mainly to expand the search ability and avoid falling into the local optimal solution.)
   b) No less than p_{5a} ,do nothing;
6. Start generating new individuals and generate a random number between (0, 1);
   a) If the value is less than the predefined probability parameter p_{6a}:
      i. Randomly select a cluster;
      ii. Randomly generate a value between (0, 1);
      iii. If the value is less than the predefined probability parameter p_{6aiii} , select a cluster center ,adding random values to generate new individuals;
      iv. Otherwise, randomly select a common individual from the class, and add a random value to generate a new individual;
   b) Otherwise, randomly select two clusters to generate new individual;
      i. Randomly generate a number between (0, 1);
      ii. If the value is less than the predefined probability parameter p_{6bii}, combine the cluster centers of the two clusters and add random values to generate new individuals;
      iii. Otherwise, combine the common individuals of the two clusters and add random values to generate new individuals;

c) The newly generated individual's adaptive value is compared with the same subscript old individual, and the individual with good fitness value is the new individual of the next iteration;

7. If n new individuals have been generated, then go to Step 8, otherwise return to Step 6;

8. If the predefined maximum number of iterations is reached, stop, otherwise re turn to Step 2.

In Step 6, which is a process to generate new individuals, the selected variation information is worth adding to a Gaussian random new information, the mathematical Eq. (1) described below [16]:

$$x_{new}^{d} = x_{old}^{d} + \xi * n(\mu,\sigma) \tag{1}$$

Among the equation above, x_{old}^{d} indicates the d-dimension of the selected individual, X_{new}^{d} is the d-dimension of the new individual, n(μ, σ) is a Gaussian random function with μ mean and σ variance, and ξ is a measure of Gaussian random value weight. The coefficient (2) is expressed as follows [17]:

$$\xi = \text{logsig}\left(\frac{\frac{T}{2} - t}{K}\right) * rand() \tag{2}$$

Among the coefficient above, logsig() is an S-type logarithmic transfer function, T is the maximum number of iterations, t is the current number of iterations, K is used to change the slope of logsig(), and the speed of the algorithm can be adjusted from global search to local search, rand() is generated random value between (0,1).

3 Clustering Method Based on BSO

In this paper, we focus on proposing a new clustering model based on BSO to search the global optimal solution. Applying this clustering method, we are looking forward to map the individuals being clustered into a linear space of (0,1) and divide the m intervals in advance.

The variation of individuals within and between clusters is shown in the Fig. 1. The choice of individuals also involves ordinary individuals and the cluster centers.

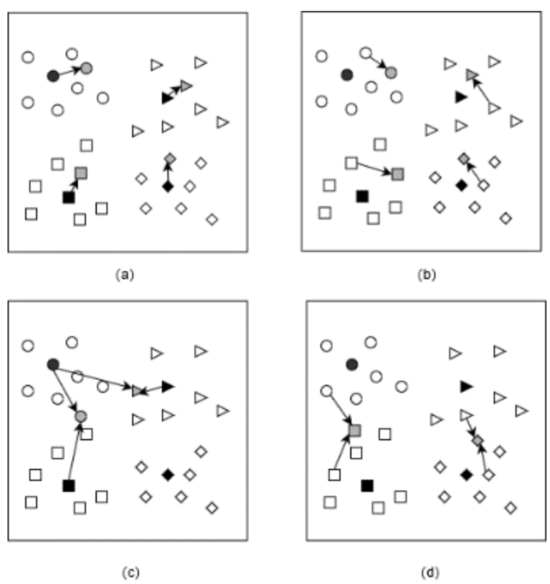

Figure 1: The variation of individuals and clusters

Given a data set $D = \{x_1, x_2, \dots, x_n\}$, x_i in it means the i^{th} example. The purpose of clustering is to learn a model $f(x)$: Optimize the n individuals and clustering the n individuals into m categories.

First of all, we open up a linear space of (0,1) and divide it into m intervals (m is the parameter we defined for the clusters we want).

For example, we want to cluster the population into 6 categories. So we need to divide the linear space into 5 blocks in advance. Starting from 0, we will give the right end of each interval a pre-set parameter. Like now, we set (0, 0.2, 0.4, 0.6, 0.8, 1) for each cluster.

[0, 0.2]	(0.2, 0.4]	(0.4, 0.6]	(0.6, 0.8]	(0.8, 1]

Then, we focus on using BSO to initialize and cluster the population. During this period, for every individual, we give it a random label. The random label is generated from rand (0,1).

The presentations of training data can be shown as Tab. 1:

Table 1: The display of training data

Individuals	**Labels**
x_1	t_1 = rand (0,1)
x_2	t_2 = rand (0,1)
x_i	t_i = rand (0,1)
x_n	t_n = rand (0,1)

When the individual gets its label, we adopt our mapping mechanism to place the individual in the specific interval. This clustering period is completed in the BSO process. Using the example above, if we get one group of individuals whose label values fall in the interval [0.4, 0.6], then we match them with the third interval. During this kind of learning period, the clusters we want are shaped gradually.

The expected mapping picture is shown as Fig. 2:

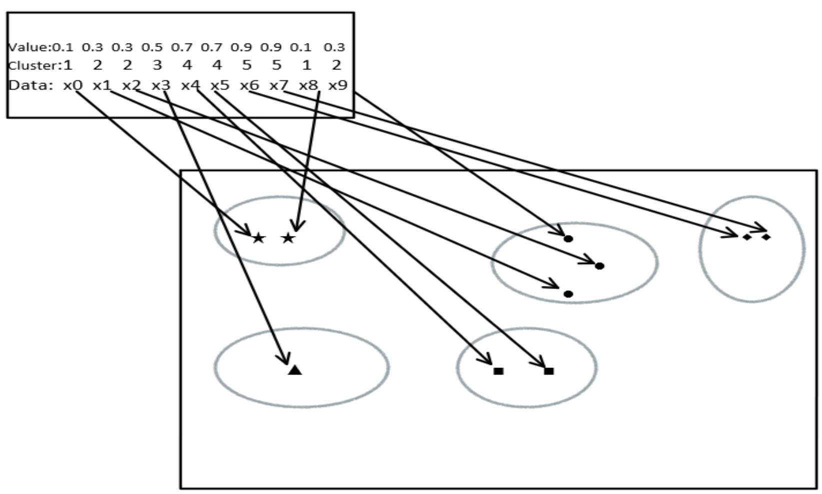

Figure 2: Mapping model

Obviously, in our clustering model, optimization work and clustering work are done simultaneously.

This is an effective and feasible process, and the mapping mechanism is carried out rapidly and probably. The features above ensure the diversity of solutions and the accuracy and validity of clustering.

4 Experimental Design

Figures and tables should be inserted in the text of the manuscript.

A. Datasets

In the experiment, four different datasets in UCI Machine Learning Repository were selected. The Tab. 2 lists the details of all datasets, including the number of dataset features, samples, and classes. For each dataset, 70% of the samples were used as training sets and the rest were used as test sets.

Table 2: Description of data

ID	Datasets	NoE	NoF	NoC
DS1	wbcd	569	30	2
DS2	iris	150	4	3
DS3	iono2	351	33	2
DS4	hill	100	606	3

B. Parameter Settings

In the experiment, we apply the novel binary model to solve the clustering problems. First and foremost, BSO mainly plays a key role in the iterative process. For the part of parameter setting, we set the more appropriate parameters selected to complete the process of K-means. In addition, during the iterative process of BSO, the four updating methods and the clustering part were carried out gradually to optimize the objective function [18]. Each algorithm runs 30 times on each dataset and the details of the BSO algorithm are given in Tab. 3.

Table 3: Parameter values setting

N	m	max_iteration	P	P_cluster	P_one/P_two
100	2	12000	0.8	0.8	0.5

5 Experimental Results and Analysis

The experimental results are shown in Tab. 4. In these table, “min” and “max” denote the minimum and maximum value of classification accuracy obtained in 30 runs, “mean” indicates the average of the final classification accuracy or the solution size, and “std” indicates the minimum and maximum deviation.

Table 4: Clustering accuracy on the wbcd

Clustering accuracy		wbcd			
		min	max	mean	std
Test	data	25.83%	75.22%	56.07%	21.78%

Table 5: Clustering accuracy on the iris

Clustering accuracy		iris			
		min	max	mean	std
Test	data	16.00%	79.33%	53.13%	25.39%

Table 6: Clustering accuracy on the iono2

Clustering accuracy		Iono2			
		min	max	mean	std
Test	data	33.90%	66.10%	55.05%	12.88%

Table 7: Clustering accuracy on the hill

Clustering accuracy		hill			
		min	max	mean	std
Test	data	47.36%	54.46%	49.97%	1.54%

From Tab. 4, Tab. 5, Tab. 6 and Tab. 7 above, it is obvious to find that the four datasets got the relatively better results. We keep the value of deviation-"std" below 0.3, which is a satisfactory results we want. The values of 'std' on our clustering model are much greater than these on other clustering algorithm especially in some complex practical problems. For hill and iono2 datasets, our model shows great stability, making the experiment carried out effectively. For wbcd and iris datasets, the table shows that the value of max is relatively high, which means our model is feasible but still have a lot of room for improvement. Considering that the shortcomings of original BSO are sensitive to the selection of initial cluster centers and easy to fall into local optimal solution rather than global best, it may well reduce the efficiency of clustering to a certain extent. However, fortunately, the value of mean was kept near 50%. Under such circumstances, we can make specific adjustment for some parameter to enable them apply to the new clustering model, improving the flexibility and accuracy of the model. Generally, the data we got proves the feasibility of the model we proposed. With the further optimized model, we are looking forward to getting the more satisfactory information.

6 Conclusion

There is no denying that clustering has been a widely concerned problem in the practical applications, which attracts many researchers in the field of data mining and machine learning. With more and more swarm intelligence algorithm proposed, we make the optimization get better and better. Nevertheless, falling into local optimum easily is still a headache for many researchers. Considering the overall superiority and convenience of clustering in BSO, we introduce it to optimize populations and cluster individuals. A series of experiments conducted, we verify the feasibility of our proposed model and get the satisfactory result. We are looking forward to optimize the model better, extending the field in which BSO can work and get the excellent clustering result.

Acknowledgements: Thanks to JIHPP for providing us with a paper template and allowing us to modify. This work is supported by National Natural Science Foundation of China (Grant No. 61876089, 61403206), by Natural Science Foundation of Jiangsu Province (Grant No. BK20141005), by Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 14KJB520025), by Science and technology program of Ministry of Housing and Urban-Rural Development (2019-K-141), by Entrepreneurial team of sponge City (2017R02002), and by Priority Academic Program Development of Jiangsu Higher Education Institutions, Innovation and entrepreneurship training program for College Students.

Funding Statement: This work was supported by Natural Science Foundation of Jiangsu Province (Grant No. BK20141005) and by Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 14KJB520025).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] A. Aldhafeeri, *Brain Storm Optimization for Electromagnetic Applications*. Diss, UCLA, 2018.

[2] M. Callon, R. Arie and L. John, *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*. Springer, 1986.

[3] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[4] H. Duan and C. Li, “Quantum-behaved brain storm optimization approach to solving Loney’s solenoid problem,” *IEEE Transactions on Magnetics*, vol. 51, no. 1, pp. 1–7, 2014.

[5] J. Stefanowski, “Data Mining-Clustering,” University of Technology, Poland, 2009.

[6] A. K. Jain, M. N. Murty and P. J. Flynn, “Data clustering: A review,” *ACM Computing Surveys*, vol. 31, no. 3, pp. 164–323, 1999.

[7] K. Li, K. Deb, Q. Zhang and S. Kwong, “An evolutionary many-objective optimization algorithm based on dominance and decomposition,” *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 5, pp. 694–716, 2014.

[8] J. P. M. De Sa, “Pattern recognition: Concepts, methods and applications,” *Springer Science & Business Media*, 2012.

[9] M. El-Abd, “Global-best brain storm optimization algorithm,” *Swarm and Evolutionary Computation*, vol. 37, pp. 27–44, 2017.

[10] P. K. Roy, C. Paul and S. Sultana, “Oppositional teaching learning based optimization approach for combined heat and power dispatch,” *International Journal of Electrical Power & Energy Systems*, vol. 57, pp. 392–403, 2014.

[11] S. Sambasivam and N. Theodosopoulos, “Advanced data clustering methods of mining web documents,” *Issues in Informing Science & Information Technology*, vol. 3, 2006.

[12] G. G. Wang, G. S. Hao, S. Cheng, Y. Shi and Z. Cui, “An improved brain storm optimization algorithm based on graph theory,” *IEEE Congress on Evolutionary Computation (CEC)*, *IEEE*, pp. 509–515, 2017.

[13] J. Xue, Y. Wu, Y. Shi and S. Cheng, “Brain storm optimization algorithm for multi-objective optimization problems,” in *Int. Conf. in Swarm Intelligence*, Springer, Berlin, Heidelberg, pp. 513–519, 2012.

[14] Y. Yang, D. Duan and H. Zhang, “Kinematic recognition of hidden markov model based oncimproved brain storm optimization algorithm,” *Space Medicine & Medical Engineering*, vol. 28, no. 6, pp. 403-407, 2015.

[15] Y. Shi, “Brain storm optimization algorithm,” in *Int. Conf. in Swarm Intelligence*, Springer, Berlin, Heidelberg, pp. 303–309, 2011.

[16] Y. Yang, Y. Shi and S. Xia, “Advanced discussion mechanism-based brain storm optimization algorithm,” *Soft Computing*, vol. 19, no. 10, pp. 2997–3007, 2015.

[17] Y. Xue, B. Zhao and T. Ma, “Performance analysis for clustering algorithms,” *International Journal of Computing Science and Mathematics*, vol. 7, no. 5, pp. 485–493, 2016.

[18] Y. Xue, T. Tang and T. Ma, “Classification based on brain storm optimization algorithm,” in *Int. Conf. on Bio-Inspired Computing: Theories and Applications*, Springer, Singapore, pp. 371-376, 2016.