

An Improved Possibilistic Fuzzy Entropy Clustering Based on Artificial Bee Colony Algorithm

Baofeng Guo and Mingyan Jiang*

School of Information Science and Engineering, Shandong university, Jinan, China

*Corresponding author

Abstract—In this paper, a possibilistic fuzzy entropy clustering algorithm (PFECM) based on unsupervised possibilistic clustering (UPC) algorithm and partition entropy (PE) has been proposed. Meanwhile, an efficient global optimization method—artificial bee colony (ABC) algorithm is introduced to optimize the proposed model. ABC-PFECM has two significant advantages compared with other algorithm. Firstly, it inherited the merits of PFCM including strong robust to noise and exclusion of the consistent clustering. Secondly, ABC-PFECM could eliminate the defects that PFCM is sensitive to the initial value and easily fall into the local optimal solution. Experimental results show splendid performance of our algorithm in decreasing computational complexity, improving clustering accuracy and enhancing global optimization capabilities.

Keywords—*possibilistic fuzzy clustering; partition entropy; unsupervised possibilistic clustering; artificial bee colony algorithm*

I. INTRODUCTION

Fuzzy clustering algorithms have well performances in clustering and data expression by enlarging the value range of membership. Recently, it has been active development in this field because of its strong ability in incomplete, inaccurate, no clear boundaries clustering problem. Fuzzy C-means clustering (FCM) [1] algorithm invented by Dunn and Bezdek is a perfect and widely used theory in all clustering algorithms based on objective functions [2]. However, because gradient descent algorithm is used to adjust parameters of FCM, the algorithm is sensitive to the initial value and easily falls into the local optimal solution. In addition, FCM algorithm requires the membership sums of the data points in all classes are 1, so the membership matrix can not truly reflect the data relation, and the robustness of the algorithm is poor as a result of sensitivity to noise. To overcome the above defects, Krishnapuram proposed possibilistic C-means clustering (PCM) [3] algorithm by relaxing the constraint conditions and constructing a new objective function. PCM enhances the noise robustness [4] compared to FCM. However, it still cannot avoid the consistency clustering problem due to the algorithm is sensitive to the initial value and easily fall into the local optimal solution. For disadvantages in FCM and PCM mentioned above, possibilistic Fuzzy C-means clustering (PFCM) [5] was presented. PFCM has strong robust to noise, and it will not bring about the defects of coincide clustering. Unfortunately, the algorithm involves too many parameters that need be artificially specified, and it will lead to the algorithm instable.

In this paper, we put forward a new clustering model called improved possibilistic fuzzy entropy clustering (PFECM), and an efficient global optimization method—artificial bee colony algorithm is introduced to optimize the proposed model. PFECM modifies PFCM based on UPC [6] and validity indexes of PE. PFECM takes advantage of the partition entropy [7] functions to solve the instability problem, which is caused by the reason of unknown structure of data, high-dimensional and multi-cluster situation. ABC-PFECM has two outstanding advantages compared with other algorithm. on one hand, the algorithm inherited the benefits of strong robust to noise and exclusion of the consistent clustering, on the other hand, it could also get rid of the defects that PFCM is sensitive to the initial value and easily fall into the local optimal solution. Experimental results show excellent performance of ABC-PFECM in decreasing computational complexity, improving clustering accuracy and enhancing global optimization capabilities.

II. PFCM AND ABC

A. PFCM

FCM is sensitive to the noise and PCM can easily cause the consistent clustering. To avoid these defects, Pal and Bezdek combined two methods and proposed PFCM [5].

$$J(U, T, V, X) = \sum_{i=1}^C \sum_{j=1}^N (au_{ij}^m + bt_{ij}^\eta) d_{ij} + \sum_{i=1}^C \gamma_i \sum_{j=1}^N (1-t_{ij})^\eta \quad (1)$$

$$s.t. \quad \sum_{i=1}^C u_{ij} = 1, u_{ij} \geq 0, t_{ij} \leq 1, 1 \leq i \leq C, 1 \leq j \leq N$$

where $a > 0$, $b > 0$, $m > 1$, $\eta > 1$ and γ_i are user defined constants. The constants a and b define the relative importance of fuzzy membership and possibility values in the objective function. Lagrangian multiplier method is employed to optimize the above equation. To minimize equation, we obtain the following equations:

$$u_{ij} = \left(\sum_{k=1}^C \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)} \right)^{-1}, 1 \leq i \leq C, 1 \leq j \leq N \quad (2)$$

$$t_{ij} = \frac{1}{1 + (b/\gamma_i d_{ij}^2)^{1/(\eta-1)}}, \quad 1 \leq i \leq C, 1 \leq j \leq N \quad (3)$$

$$v_i = \frac{\sum_{j=1}^N (au_{ij}^m + bt_{ij}^\eta)x_j}{\sum_{j=1}^N (au_{ij}^m + bt_{ij}^\eta)}, \quad 1 \leq i \leq C \quad (4)$$

It is obvious that PFCM carries forward the merits of FCM and PCM. It has strong robust to noise, and will not lead to the coincide clustering. However, too many parameters need be artificially specified, which will cause the instability to the algorithm, especially γ_i need lots of training.

B. ABC

FCM ABC is a new swarm intelligence algorithm proposed by Turkish scholars Karaboga in 2005, which is aimed at solving the function optimization problems by simulating the foraging behavior of honey bees [8]. The algorithm has successfully applied to many fields since it has the virtues of simple operation, less parameters, good robustness and excellent global optimization capacity.

There are three types of bees in the ABC algorithm: employed bees, onlooker bees, and scout bees. The employed bees search locally optimal solutions around the neighborhood of its present position; then, they share the information, from which the onlooker bees tend to select good optimal solutions; The scout bees are translated to employed bees and abandon their solutions to search new one once the local optimization keep certain long time.

Employed bees and onlooker bees operate in local searching as below:

$$v_{ij} = x_{ij} + rand(x_{ij} - x_{kj}) \quad (5)$$

Selective probability for onlooker bees:

$$P_i = \frac{fitness(i)}{\sum_i^{SN} fitness(i)} \quad (6)$$

where SN is the number of employed bees, fitness(i) is the fitness value of the i^{th} solution of the employed bees. The fitness is described as equation below:

$$fitness(i) = \frac{1}{1 + objfun(i)}. \quad (7)$$

III. ABC-PFECM

A. PFECM

In 2006, Yang and Wu proposed a new possibilistic clustering algorithm called unsupervised possibilistic clustering (UPC) [6]. Compared with PCM, UPC solves the dependency of γ_i on experience. Inspired by this, we introduce covariance matrix [9] to set γ_i and eliminate too much user-defined parameter, which can reflect compact and discrete degree of the data. To some extent, covariance matrix is benefit to decrease computational complexity. Besides, to weaken the instability of the algorithm caused by incognizance in the structure, dimension and class number of data, we put forward PFECM algorithm [10] by introducing partition entropy (PE) [11] into PECM.

$$J(U,T,V) = \sum_{i=1}^c \sum_{j=1}^N (u_{ij}^m + t_{ij}) d_{ij}^2 + \frac{\sigma^2}{m^2 c} \sum_{i=1}^c \sum_{j=1}^N (t_{ij} \log t_{ij} - t_{ij}) + \lambda \sum_{i=1}^c \sum_{j=1}^N t_{ij} \log t_{ij} \quad (8)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^c (d_{ij}^2)^{2/(m-1)} d_{kj}}, \quad \forall i, j \quad (9)$$

$$t_{ij} = \exp\left(-\frac{m^2 c (d_{ij}^2 + \lambda)}{\sigma^2 + m^2 c \lambda}\right), \quad \forall i, j \quad (10)$$

$$v_i = \frac{\sum_{j=1}^N (u_{ij}^m + t_{ij}) x_j}{\sum_{j=1}^N (u_{ij}^m + t_{ij})}, \quad \forall i, 0 \leq t_{ij} \leq 1 \quad (11)$$

B. ABC-PFECM

Aiming to solve the defects of PFCM in poor global optimization capabilities, sensitivity to initial values and dependence on parameter [12], we proposed ABC-PFECM by combining ABC and PFECM.

In ABC-PFECM, candidate solutions correspond to cluster centers $V = [v_1, v_2, \dots, v_c]$. The fitness function and local searching function are shown in equation (12) and (13).

$$fitness = \frac{1}{1 + J(U, T, V)} \quad (12)$$

$$v_{ij} = x_{ij} + rand(x_{ij} - x_{kj}) + \beta(x_{best} - x_{ij}) \quad (13)$$

Algorithm ABC-PFECM

Initialize the parameters of ABC, population number NP, the number of onlooker bees SN, limit, iterations maxcycle, initial cluster centers V, cluster number C and λ ;

Generate new solution sol by (13), then Update U, T by (9), (10);

- ① Update $fitness(i)$ according to the greedy criterion;
- ② The onlooker bees select their employed bee to follow, based on the probability described as equation (6);
- ③ If the $fitness(i)$ is still unchanged after $limit$ times, the onlooker bee become scout bee, then a new searching position is needed;
- ④ Repeat ③-⑥ until iterations or accuracy is satisfied;
- ⑤ Get cluster centers V_{best} , the corresponding U and cluster result as $k = \arg \max(u_{ij})$.

IV. NUMERICAL EXPERIMENTS

In order to validate the effectiveness and feasibility of ABC-PFECM, We conduct several experiments to study the performance of the algorithm. Parameters of ABC-PFECM are set as:

$m=2$, NP=50, limit=30, max cycle=300, $\lambda=0.4$, which are used in the following experiments.

We conduct numerical experiments on a two-dimensional 12 points data set X12 [13]. Ten points of the data set X12 constitute two diamonds except for the points X6 and X12, symmetrically distributed on both sides of the y-axis. X6 and X12 are noise or outlier equidistant from the two clustering centers. Running FCM, PFECM algorithm respectively, the results are shown in Table I.

As the experimental results show in Table I, the memberships of the points X6 and X12 are 0.5 both in FCM and ABC-PFECM algorithm. However, because the points x6 and x12 are noise or outlier, and far away from the clustering center, the possibility values in ABC-PFECM algorithm are very small. Therefore, ABC-PFECM algorithm carries on the virtues of PFCM algorithm, and has the robustness to noise and outliers.

We change the value of parameter λ from 0.1 to 1, and the clustering accuracy of ABC-PFECM on IRIS data set are illustrated in the table II. From the Table II, we acquire parameter λ has crucial influence on the clustering accuracy. And we get the value of parameter λ when the clustering accuracy of ABC-PFECM is the highest. In addition, the value will be used as the basis for the following experiment.

According to the experimental result above, we take $\lambda=0.4$, testing the Clustering performance of FCM, PCM, IPFCM($\lambda=0$), PFECM,ABC-PFECM algorithms on the UCI standard data sets Iris and wine, respectively. The results were listed in Table III.

TABLE I. MEMBERSHIPS AND POSSIBILITIES VALUE OF DATA SET X12 ON FCM AND ABC—PFECM

method	U	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
FCM	U1	0.94	0.97	0.99	0.90	0.92	0.50	0.08	0.03	0.01	0.10	0.06	0.50
	U2	0.06	0.03	0.01	0.10	0.08	0.50	0.92	0.97	0.99	0.90	0.93	0.50
ABC-PFECM	U1	0.94	0.96	0.99	0.91	0.92	0.50	0.08	0.04	0.01	0.09	0.06	0.50
	U2	0.06	0.04	0.01	0.09	0.08	0.50	0.92	0.96	0.99	0.91	0.94	0.50
	T1	0.19	0.49	0.77	0.16	0.41	0.03	0.00	0.00	0.00	0.00	0.00	0.00
	T2	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.41	0.49	0.77	0.16	0.19

TABLE II. THE VALUE OF PARAMETER λ WITH CLUSTERING ACCURACY OF ABC-PFECM ON IRIS

λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
accuracy	0.90	0.90	0.90	0.9067	0.9067	0.9067	0.90	0.90	0.90	0.90

TABLE III. ACCURACY OF DIFFERENT METHODS ON IRIS AND WINE DATA SET

Database	FCM	PCM	IPFCM	PFECM	ABC-PFECM
Iris	89.3%	66.7%	90%	90%	90.67%
wine	68.54%	64.51%	70.22%	71.35%	73.03%

Iris data set is a four-dimensional data set containing 150 samples with three clusters, well wine is a thirteen-dimensional data set containing 178 samples with three clusters, We carry out experiments on these two data sets. The results in table III demonstrate that PFECM carries forward the benefits of PFCM algorithm, eliminate the problem of generating coincident

clusters by adding partition entropy function to objective function of PFCM algorithm. Furthermore, PFECM algorithm receives an illustrious property in high-dimensional and multi-cluster situation. Introduced ABC algorithm to PFECM, ABC-PFECM avoid the problem of IPFCM sensitive to initial value

and poor global optimization property of global, indeed improve the clustering accuracy and convergence.

V. SUMMARY

In this paper, we propose PFECM algorithm by effectively combining PFCM, UPC, and partition entropy. In the PFECM, covariance matrix reflects compact and discrete degree of the data, and is used to eliminate the over-dependence of parameters. By introducing partition entropy, the stability of the algorithm is improved in unfamiliar to the structure of data, high-dimensional and multi-cluster situation. Besides, to figure out the problems about poor global optimization capabilities, sensitivity to initial values and dependence on parameter, ABC algorithm is also introduced to optimize the proposed model. The experimental results demonstrate ABC-PFECM algorithm owing splendid properties in decreasing computational complexity, improving clustering accuracy and enhancing global optimization capabilities.

ACKNOWLEDGMENT

This work was financially supported by the Natural Science Foundation of Shandong Province (ZR2014FM039).

REFERENCES

- [1] J. C. Bezdek, Pattern recognition with fuzzy objective function algorithms, plenum pree .New York, 1981.
- [2] Igor Skrjanc and Dejan Dovzan. Evolving Gustafson-Kessel Possibilistic C-Means Clusteing. Procedia Computer Science, Vol 53 2015, pp191-198.
- [3] Krishnapuram. R and Keller J M. A possibilistic approach to clustering[J]. IEEE Trans Fuzzy Systems, 1 (2), 1993,pp98-110.
- [4] Baoju Zhang, Shan Qin, Wei Wang, Dan Wang, Lei Xue. Data stream clustering base on Fuzzy C-Mean algorithm and entropy theory. Signal Processing. 2015, in press.
- [5] Nikhil R. pal, Kuhu Pal, James M. Keller and James C. Bezdek. A possibilistic fuzzy c-Means Clustering Algorithm. IEEE transactions on fuzzy systems, Vol 13(4), 2005, pp517-530.
- [6] Miin Shen Yang and Kuo Lung Wu. Unsupervised possibilistic clustering. Pattern Recognition, vol.39,2006,pp5-21.
- [7] Haiyan Yu, Xiaobin Zhi and Jiulun Fan. Image segmentation based on weak fuzzy partition entropy. Neurocomputing,168,2015, pp994-1010.
- [8] Mingyan Jiang. The Artificial Bee Colony Algorithm and its Applications . Science press, Beijing, 2014.
- [9] Yating Hu, Chuncheng Zuo, Fuheng Qu, Weili Shi. Unsupervised Possibilistic Clustering Based on Kernal Methods. Physics Procedia vol.25,2012,pp1084-1090.
- [10] Haijun Fu, Xiaohong Wu, Hanping Mao, Bin Wu. Fuzzy Entropy Clustering Using Possibilistic Approach. Procedia engineering, vol .15,2011,pp1993-1997.
- [11] M.Zarinbal, M. H. Fazel Zarandi, I.B.Turksen.Relative entropy fuzzy c-means clustering. Information Sciences, vol. 260, 2014, pp74-97.
- [12] Gulsen Aydin Keskin and Coskun Ozkan.Research Article Multiple Criteria ABC Analysis with FCM clustering. Journal of Industrial Engineering, 2013.
- [13] Wu X H, Zhou J J. A novel possibilistic fuzzy c-means clustering. Acta Electronica Sinica, 2008, vol.36 (10),pp1996-2000.