International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.3 (2016), pp.179-188
http://dx.doi.org/10.14257/ijsip.2016.9.3.16

A Novel Data Clustering Algorithm based on Modified Adaptive Particle Swarm Optimization

Ganglong Duan, Wenxiu Hu, Zhiguang Zhang

Xi'an University of Technology, Shaanxi 710054, China
duanganglong2@126.com

Abstract

Fuzzy clustering is a popular unsupervised learning method used in cluster analysis which allows a point in large data sets belongs to two or more clusters. Prior work suggests that Particle Swarm Optimization based approach could be a powerful tool for solving clustering problems. In this paper, we propose a data clustering algorithm based on modified adaptive particle swarm optimization. We choose to use artificial bee colony algorithm combined with PSO technique to modify the traditional clustering methods due to its fast convergence and the presence of adaptive mechanisms based on the evolutionary factor. On the one hand, Particle Swarm Optimization is proven to be an effective and robust technique for fuzzy clustering. On the other hand, the artificial bee colony algorithm has the capability to generate diversity within the swarm when the guide bees are in the exploration mode. Through numerical analysis and experimental simulation, we verify that our algorithm performs much better compared with other state-of-the-art algorithms. Future research schedule is also discussed in the final part.

***Keywords**: Data Clustering, Artificial Bee Colony, Fuzzy C-Means Algorithm, Adaptive Particle Swarm Optimization, Clustering Validation*

1. Introduction

Clustering algorithms have emerged and rapidly developed as an alternative powerful meta-learning tool to undertake a broad range of applications because it is particularly useful for segmenting large multidimensional data into distinguishable representative clusters. However, Data clustering is one of the most complicated engineering problems and it is considered as an NP-hard problem. An objective of clustering is to identify parts of the data that have high degrees of similarity with other parts, and group the similar parts together into clusters automatically. Modern applications of data clustering are generally divided into the following categories: (1) Data mining. In [1], Maryam proposed a novel efficient approach based on differential evolution algorithm for data mining. Their algorithm uses a differential operator for producing new solutions and exchanges information among population members. The advantage of their algorithm is the ability of keeping sufficient memory, which keeps information of suitable solution in the recent population. Xu *et. al.*, proposed an unsupervised adaptation approach to leveraging feedback loop data for data mining [2]. They proposed to perform a CMLLR-based recognition of an unknown utterance by selecting a set of CMLLR transforms from the most similar cluster, which are pre-trained by using the utterances. (2) Image processing and computer vision. Wang *et. al.*, [3] proposed a novel effective image representation method using kernel classification. Data clustering technique is used to undertake the feature selection work. In [4], Feng *et. al.*, proposed an example image super-resolution algorithm based on modified k-means with hybrid particle swarm optimization. Moreover, Chen et al. introduced an automatic semantic modeling of indoor scenes from data clustering in [5] as an example for computer vision application. (3) Pattern recognition and machine learning. Jun *et. al.*, proposed a document clustering method using dimension reduction and support vector clustering (SVC) to overcome sparseness. Data clustering served

ISSN: 2005-4254 IJSIP

as the selector for the dimension reduction procedure [6]. (4) Medical applications. Gene selection and medical image processing are two major applications for data clustering. More literature reviews for the topic could be found in [7-16].

In this paper, we propose a data clustering algorithm based on modified adaptive particle swarm optimization. The rest of the paper is organized as the follows: Section 2 gives the basic introduction to the related algorithms and methodologies. Section 3 discusses our proposed novel algorithm and in the fourth section, we conduct experimental analysis and simulation to verify the effectiveness and feasibility of our algorithm compared with other state-of-the-art methods. In the final part, we conclude our work and set up our prospect.

2. Overview of Related Work

Fuzzy C-means clustering (FCM) was firstly introduced by Dunn in 1973 [17] and was extended by Bezdek in 1981 [18]. Since then, FCM becomes one of the primary promising fuzzy clustering algorithms in the research community. Moreover, majority of variants for FCM have been introduced. For instance, Optimal-selection-based suppressed fuzzy c-means clustering algorithm [19] is a fuzzy clustering technique which ranks all the data based on their biggest membership degree values, and then the membership degree values of the top r ranked data points are modified while the membership degree values of the other data points are not changed in the iteration procedure. In [20], Fan et al. introduced a revised suppressed fuzzy C-Means clustering algorithm, they points out a method to select the fixed suppressed rate by the structure of the data itself. In [21], Gong et al. proposed the MRF based fuzzy clustering method which adopts the advancement of the Markov random field (MRF) energy function to modify the traditional FCM. All of the current fuzzy clustering algorithms perform better than the original ones.

With regards to particle swarm optimization (PSO) approaches, literature reviews in [22] have brought an overview introduction for pure particle swarm optimization. As for as the PSO applications in the data clustering research area is concerned. Tsai et al. introduced a selective particle regeneration based PSO for data clustering [23]. Moreover, Kuo *et. al.*, combined the PSO and genetic algorithm for dynamic clustering [24]. Their algorithm can automatically cluster data by examining the data without a pre-specified number of clusters.

3. Our Proposed Framework for Fuzzy Data Clustering

3.1. Fuzzy C-means Clustering (FCM)

Clustering approaches based on fuzzy logic, such as FCM and its related approaches have proved to be competitive to conventional clustering algorithms, especially for real-world applications and utilities. The comparative and significant advantages of these approaches is that instead of considering sharp boundaries between the individual clusters, they allow each feature vector to be included to different clusters by a certain degree (We generally name these techniques to be soft-clustering compared with hard-clustering which is adopted by most of tradition lustering algorithms). Eigenvector of membership degree is often thought of as a function of cluster centroid distance or other representative of the vector of the cluster. FCM is a pretty standard least squared error model that generalizes an earlier and very popular non-fuzzy c-means model that produces hard clusters of the data. Through minimizing the weighted within group amount of squared error function shown in the formula 1, the optimal partition denoted as c is iteratively produced.

$$J = \sum_{i=1}^{n} \sum_{j=1}^{c} \left(u_{ij}\right)^{m} d^{2}\left(y_i, c_j\right) \tag{1}$$

In the formula 1, the dataset for processing is in a d-dimensional vector space represented as $Y = \left[y_1, y_2, ..., y_n\right]$. n represents the number of data items and c is the number of clusters defined by the user. In the $j-th$ cluster, the membership degree of y_i is defined as u_{ij}. m is a weighted exponent on each fuzzy membership. c_j represents the cluster center whereas the square distance measure (SDM) between the center c_j and the object y_i is defined as $d^2\left(y_i, c_j\right)$. We could derive the optimal solution iteratively for the issue according to the following steps:

Step 1) Input(c, m, data);

Step 2) Initialize the fuzzy partition matrix $U=\left[u_{ij}\right]$;

Step 3) Start iteration process and set $t=1$;

Step 4) Calculate the c cluster centers with U^t according to the equation 2;

$$c_i = \frac{\sum_{i=1}^{n}\left(u_{ij}\right)^m y_i}{\sum_{i=1}^{n}\left(u_{ij}\right)^m} \tag{2}$$

Step 5) Calculate the membership U^{t+1} according to the equation 3;

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(d_{ij} / d_{kj}\right)^{2/(m-1)}} \tag{3}$$

Step 6) Return to the step 4 if the stopping criteria is not met, set $t=t+1$.

3.2. Artificial Bee Colony Algorithm

The artificial bee colony algorithm was firstly introduced by Karagoba [25] which was inspired by the behavior of bees. In the original version of the algorithm, there are three types of artificial bees: employed, onlookers and scouts. The employed bees are responsible to explore the food sources. The bees that decide to exploit a food source depending on the information shared by the employed bees which are called onlookers bees. The bees that try to look for new food resources are called scout bees. In general, the bees could be separated into two types: the guide bees and the guided. The guide bees have two different behaviors: exploration and exploitation. The guided bees are the onlooker bees. In general, the number of the guide bees and guided bees is the same. Besides, there is only one employed bee for every food source. We use SN to represent the total number of food resources. In each resource, the update procedure will be based on the equation 4.

$$v_i = x_i + r_i\left(x_i - x_k\right) \tag{4}$$

Where r_i represents the vector of random numbers generated within the range $\left[-1,1\right]$. k represents one of the food sources. The updated solution v_i is compared with the original one and the better on remains. Besides, each guided bee needs to select one of the available food sources to explore. The probability to select a food source can be calculated by the formula 5.

$$p_i = \frac{f_i t_i}{\sum_{j=1}^{SN} f_i t_j} \tag{5}$$

Each food source is evaluated in each iterations. If a food source does not improve its fitness after a pre-determined number of steps, this food source is discarded and the guide bee is responsible for this food source changes its behavior from exploitation mode to exploration mode. Later, the food source is replaced by the novel one according to the equation 6.

$$x_{id} = x_d^{\min} + r\left(x_d^{\max} - x_d^{\min}\right) \tag{6}$$

3.3. Modified Artificial Bee Colony based Particle Swarm Optimization

The core algorithm of this paper is the artificial bee colony based particle swarm optimization (ABCPSO). We chose the ABCPSO due to its fast convergence and the presence of adaptive mechanisms based on the evolutionary factor. On the other hand, the artificial bee colony algorithm has the capability to generate diversity within the swarm when the guide bees are in the exploration mode. Actually, we propose here to include the exploration ability of the artificial bee colony algorithm instead of using the traditional learning Strategy used in traditional particle swarm optimization technique. The traditional learning Strategy is not enough to ensure the required diversity for optimization in high dimensional spaces. In our model, the $i-th$ guide bee will be optimized based on the expression 7. The B represents the barycenter of the guide bees, the $step$ is the vector of dispersion steps of the swarm, these two terms are defined in the formula 8 and 9.

$$v_i = x_i + step \times r_i \left[x_i - B / Dis\left(x_i, B\right) \right] \tag{7}$$

$$B = \sum_{i=1}^{N/2} x_i\, fit_i \Big/ \sum_{i=1}^{N/2} fit_i \tag{8}$$

$$step_d = \left\{1 - \left[1 / \left(1 + e^{-\alpha\left(f_{evol} + displacement\right)}\right)\right]\right\}\left(x_d^{\max} - x_d^{\min}\right) \tag{9}$$

In the formula 7, the $Dis\left(x_i, B\right)$ denotes the Euclidian distance between B and current position x_i. When the fitness value changes, we will update the value of dispersion based on the formula 10. If the new position performs better than the current one, the step value will decrease and the vice versa.

$$step = step \pm f_{evol}\, step \tag{10}$$

After this, the guided bees need to select one of the food sources, where the probability for choosing the food source is evaluated by using the equation presented in formula 5. Later, every bee will search for new source by equation 11.

$$v_i = x_i + r_i\left(fs_i - B\right) \tag{11}$$

The next step is to estimate the evolutionary state, such as in the traditional adaptive PSO algorithm. However, this estimation just considers the food sources, which are the current potential solutions of the problem. In the Figure 1, we present the membership function adopted in our algorithm.

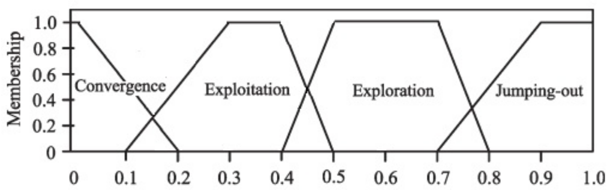

Figure 1.The Membership Function Adopted by the ABCPSO Algorithm

3.4. Detailed Steps of the Proposed ABCPSO Algorithm

3.4.1. Particle and Velocity Encoding/Decoding: We define a $2\times k$ matrix to be a particle in the formula 12. In the first row, we pre-define the centers and the values in the second row take in control of the activation of each center.

$$X_i = \begin{pmatrix} x_{1,1}^i & x_{1,2}^i & \cdots & x_{1,\mathrm{k}}^i \\ t_{2,1}^i & t_{2,2}^i & \cdots & t_{2,\mathrm{k}}^i \end{pmatrix} \tag{12}$$

The velocity matrix should have the same dimension as the position matrix with a range. As an example, we set the range to be $[v_{\min}, v_{\max}]$ which means all values of the velocity should fall into the scope between $v_{\min}$ and $v_{\max}$. Therefore, $i-th$ velocity is expressed as:

$$V_i = \begin{pmatrix} v_{x1,1}^i & v_{x1,2}^i & \cdots & v_{x1,\mathrm{k}}^i \\ v_{t2,1}^i & v_{t2,2}^i & \cdots & v_{t2,\mathrm{k}}^i \end{pmatrix} \tag{13}$$

The first row is the velocity of the centers, and the second row is the velocity of the threshold values. As for the decoding step, $Y = (y_1, \mathrm{y}_2, ..., y_n)$ represents the d-dimensional data set. We could decode the cluster centers as $C = (c_1, \mathrm{c}_2, ..., \mathrm{c}_k)$ using the formula 2.

3.4.2. Clustering Validation Techniques: The aim of clustering validation is to evaluate the clustering results to find the best partition that fits the underlying data. Therefore, cluster validity is used to quantitatively evaluate the results of the clustering algorithm. Compactness and separated as two widely used standard measuring the quality of partitioned data set to a number of clusters. Traditional operation algorithms iterative use different input values and select the best effective measure to determine the "best" clustering. We introduce some of the state-of-the-art validity indices below. (1) Modified Partition Coefficient Index (MOC): Modification of the PC index, which can reduce the monotonic tendency, is proposed by Dave in 1996 [26], an optimal cluster number is found by maximizing MPC to produce a best clustering performance for a data set. The equation 14 defines the index. (2) Fukuyama and Sugeno Index (FS): In the formula 15, we define this index which is firstly proposed by Fukuyama in 1989 [27]. The best clustering performance for a data set is found by maximizing the value of Fukuyama and Sugeno Index. (3) Weighted Inter-Intra Index (Wint): The weighted interintra measure is firstly introduced by Strehl [28] in 2002. It compares the compactness of the data to its separation, the calculation of Wint is shown in the formula 16. Weighted Inter-Intra index comes to its maximum value when the cluster structure is optimal.

$$MPC = 1 - \frac{c}{c-1}(1 - PC) \tag{14}$$

$$FS = \sum_{i=1}^{n}\sum_{j=1}^{c} \mu_{ij}^m \left\| x_i - c_j \right\| - \sum_{i=1}^{n}\sum_{j=1}^{c} \mu_{ij}^m \left\| c_j - \bar{c} \right\| \tag{15}$$

$$Win.t = \left(1 - \frac{2c}{n}\right) \cdot \left(1 - \frac{\sum_i \frac{1}{n - |c_i|} \sum_{j \neq i} in.ter(c_i, c_j)}{\sum_i \frac{2}{|c_i| - 1} in.tra(c_i)}\right) \tag{16}$$

3.4.3. Flow-chart Description of Our Methodology

The overview of our methodology is shown in the Table 1 below.

Table 1.The Flow-Chart Description of Our Algorithm

Algorithm 1. Our Proposed Methodology
1. **Input**: The dataset $Y=[y_1,y_2,...,y_n]$, number of cluster c and coefficient m.
2. Initialize the swarm randomly.
3. Start the iteration and set t=1.
4. Update the velocity and position of each particle.
5. Update the personal best and global best.
6. Calculate the partition matrix U.
7. If not meet the stopping criterion, t=t+1 and jump to step 4.
8. Construct the data using partition matrix U.
9. Calculate the reconstruction error through the following equation.
$y_i=\sum_{j=1}^{c}u_{ij}^{m}c_j/\sum_{j=1}^{c}u_{ij}^{m}$
10. Select the partition matrix and centers.
11. **Output**: A partition matrix U and corresponding centers.

4. Experimental Analysis and Simulation

In order to verify the effectiveness and feasibility of our proposed methodology, we conduct numerical and experimental simulation in this section. We firstly introduce the set-up of the experimental environment then we conduct simulation with numerical analysis.

4.1. Set-up of the Experiment

The simulation environment is initialized as the follows. Six physical machines equipped with 2 TB hard disk and 8 GB of RAM, and the simulation software is installed on Windows Win8 platform and Intel core 4 quad core 3.2 GHz and 4 GB of RAM. The experiments are conducted on a number of datasets taken from the UCI repository shown in the Table 2, we generate a novel synthetic data through Matlab shown in the Figure 2.

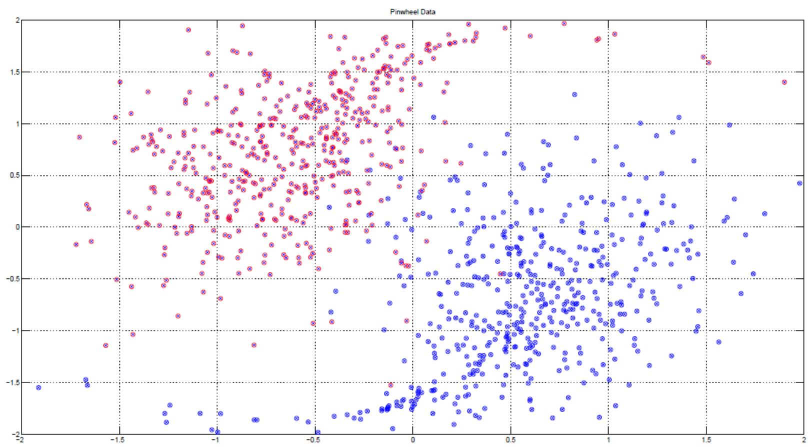

Figure 2.The Synthetic Data Set Generated by Matlab

Table 2.The Data Sets for the Experiment

Data-Set	Dimensions	Instances	Classes
Pinwheel	2	1000	2
Transfusion	4	748	2
Jain	2	373	2
Thyroid	5	215	2
Breast-W	9	699	2

4.2. Experiment and Simulation

In this section, we firstly test the proposed algorithm individually. Later, we compare our method with other state-of-the-art algorithms. In the Table 3 and Figure 3 we show the reconstruction errors of the transfusion data set, where c ranges from 2 to 9, have been calculated using the proposed algorithm. We could easily draw the conclusion that c=2 is the optimal number of clusters. In the Table 4 and Figure 4, we show the clustering accuracy of our algorithm compare with other methods.

	C=2	C=3	C=4	C=5	C=6	C=7	C=8	C=9
Wint	23.7	28.9	24.9	42.3	50.7	63.7	41.6	37.5
PCAES	36.6	81.9	62.8	47.9	66.7	56.9	52.9	77.3
XB	21.3	42.6	26.3	36.1	27.5	61.3	68.3	62.9
FS	22.6	23.6	24.3	46.3	29.7	36.7	42.5	27.6
MPC	17.3	22.5	42.6	43.3	12.8	22.8	43.6	44.9
SE	23.6	38.2	27.6	44.6	49.8	26.5	33.4	36.8

Table 3.The Reconstruction Error

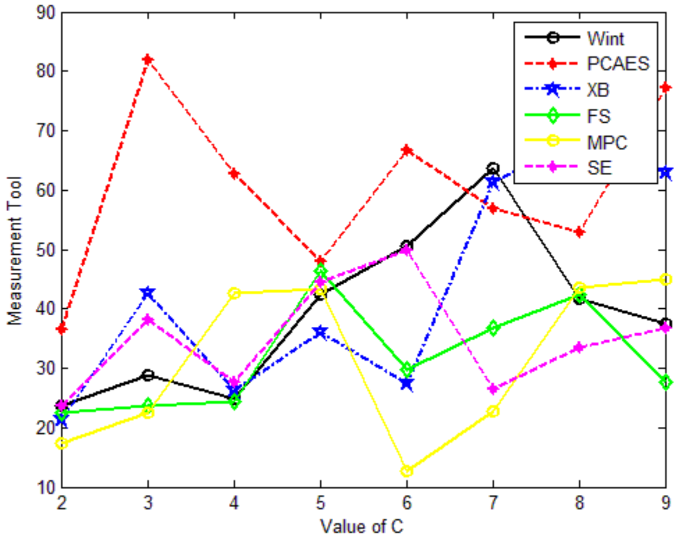

Figure 3.The Reconstruction Error Graph

Table 4.The Clustering Accuracy Experimental Result

	Test1	Test2	Test3	Test4	Test5	Test6	Test7	Test8
Ours	85.7	88.4	90.5	82.3	88.2	86.5	80.4	87.5
DEN	72.3	71.3	74.0	75.6	79.3	70.9	69.3	77.3
FCM	74.9	79.3	81.6	74.3	74.3	80.6	74.5	72.9
EM	80.3	86.2	86.9	81.9	81.7	87.5	78.3	87.6

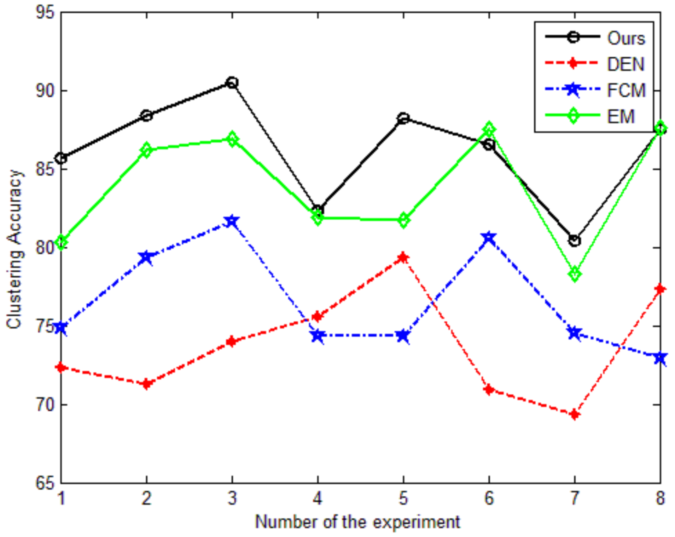

Figure 4.The Clustering Accuracy Graph

5. Conclusion and Summary

Clustering algorithms have emerged and rapidly developed as an alternative powerful meta-learning tool to undertake a broad range of applications because it is particularly useful for segmenting large multidimensional data into distinguishable representative clusters. In this paper, we propose a data clustering algorithm based on modified adaptive particle swarm optimization. We chose to take combination ABC algorithm and PSO technique due to its fast convergence and the presence of adaptive mechanisms based on the evolutionary factor. On the other hand, the artificial bee colony algorithm has the capability to generate diversity within the swarm when the guide bees are in the exploration mode. Actually, we propose in the previous sections to include the exploration ability of the artificial bee colony algorithm instead of using the traditional learning strategy used in traditional particle swarm optimization technique. The traditional learning Strategy is not enough to ensure the required diversity for optimization in high dimensional spaces. Through numerical analysis and experimental simulation, we could conclude that our algorithms perform much better compare with other popular algorithms. In the future research, we have scheduled to conduct more research focused on the optimization part of PSO to enhance the accuracy.

Acknowledgements

This research is supported by the youth fund of national natural science fund project: information technology based on the dynamic adjustment speed economic value measurement and validation studies (NO. 71403206) and the project supported by the natural science foundation of the Shaanxi province education department: flexible recommendation system based on hybrid strategy research and development (NO. 14JK1522).

References

[1] M. Hosseini, M. Sadeghzade and R. Nourmandi-Pour, "An efficient approach based on differential evolution algorithm for data clustering", Decision Science Letters, vol. 3, no. 3, **(2014)**, pp. 319-324.
[2] X. Wu, X. Zhu, G.-Q. Wu and W. Ding, "Data mining with big data", Knowledge and Data Engineering, IEEE Transactions on, vol. 26, no. 1, **(2014)**, pp. 97-107.
[3] H. Wang and J. Wang, "An effective image representation method using kernel classification", in Tools with Artificial Intelligence (ICTAI), 2014 IEEE 26th International Conference on, **(2014)** Nov, pp. 853–858.
[4] S. Chen, Z. Xu and Y. Tang, "A hybrid clustering algorithm based on fuzzy c -means and improved particle swarm optimization", Arabian Journal for Science and Engineering, vol. 39, no. 12, **(2014)**, pp. 8875-8887.
[5] B. Liu and H. Fan, (2013), "Semantic labeling of indoor scenes from rgb-d images with discriminative learning", Sixth International Conference on Machine Vision (ICMV **2013**).
[6] S. Jun, S. Park and D. Jang, "Document clustering method using dimension reduction and support vector clustering to overcome sparseness", Expert Systems with Applications, vol. 41, no. 7, **(2014)**, pp. 3204–3212.
[7] A. Hatamlou, "Black hole: a new heuristic optimization approach for data clustering. Information Sciences", vol. 222, no. 3, **(2013)**, pp. 175–184.
[8] V. Bhatnagar, S. Kaur and S. Chakravarthy, "Clustering data streams using grid-based synopsis", Knowledge and Information Systems, vol. 41, no. 1, **(2014)**, pp. 127-152.
[9] Y. Cheung and H. Jia, "Categorical-and-numerical-attribute data clustering based on a unified similarity metric without knowing cluster number", Pattern Recognition, vol. 46, no. 8, **(2013)**, pp. 2228–2238.
[10] A. George, "Efficient high dimension data clustering using constraint-partitioning k-means algorithm", International Journal Arab Journal of Information technology, vol. 10, no. 5, **(2013)**, pp. 467-476.
[11] G. Krishnasamy, A. J. Kulkarni and R. Paramesran, "A hybrid approach for data clustering based on modified cohort intelligence and K-means", Expert Systems with Applications, vol. 41, no. 13, **(2014)**, pp. 6009-6016.
[12] P.-L. Lin, P.-W. Huang, C.-H. Kuo and Y. H. Lai, "A size-insensitive integrity-based fuzzy c-means method for data clustering", Pattern Recognition, vol. 47, no. 5, **(2014)**, pp. 2042-2056.
[13] H. Qin, X. Ma, T. Herawan and J. M. Zain, "MGR: An information theory based hierarchical divisive clustering algorithm for categorical data", Knowledge-Based Systems, vol. 67, **(2014)**, pp. 401-411.
[14] M. Yuwono, S. W. Su, B. D. Moulton and H. T. Nguyen, "Data clustering using variants of rapid centroid estimation", Evolutionary Computation, IEEE Transactions on, vol. 18, no. 3, **(2014)**, pp. 366-377.
[15] R. Maronna, "Charu C. Aggarwal and Chandan K. Reddy (eds.): Data clustering: algorithms and applications", Statistical Papers, **(2015)**, pp. 1-2.
[16] G. Yan, "Finding common ground among experts' opinions on data clustering: With applications in malware analysis", Data Engineering, IEEE International Conference on, IEEE, **(2014)**, pp. 15-27.
[17] J. C. Dunn, "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters", Journal of Cybernetics, vol. 3, **(1973)**, pp. 32-57.
[18] J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms", ISBN 0-306-40671-3, **(1981)**.
[19] F. Zhao, J. Fan and H. Liu, "Optimal-selection-based suppressed fuzzy c-means clustering algorithm with self-tuning non local spatial information for image segmentation", Expert Systems with Applications, vol. 41, no. 9, **(2014)**, pp. 4083–4093.
[20] W. Hung, M. Yang and D. Chen, "Parameter selection for suppressed fuzzy c-means with an application to mri segmentation", Pattern Recognition Letters, vol. 27, no. 5, **(2006)**, pp. 424–438.
[21] M. Gong, L. Su, M. Jia, and W. Chen, "Fuzzy clustering with a modified mrf energy function for change detection in synthetic aperture radar images", Fzzy Ym Ranaon on, vol. 22, no. 1, **(2014)**, pp. 98 - 109.
[22] B. Xin, J. Chen, J. Zhang, H. Fang and Z. Peng, "Hybridizing differential evolution and particle swarm optimization to design powerful optimizers: a review and taxonomy", IEEE Transactions on Systems Man and Ctbernetics Part C-Applications and Reviews, vol. 42, no. 5, **(2012)**, pp. 744-767.
[23] C. Tsai and I. Kao, "Particle swarm optimization with selective particle regeneration for data clustering", Expert Systems with Applications, vol. 38, no. 6, **(2011)**, pp. 6565–6576.
[24] R. J. Kuo, Y. J. Syu, Z. Chen and F. C. Tien, "Integration of particle swarm optimization and genetic algorithm for dynamic clustering. Information Sciences", vol. 195, no. 13, **(2012)**, pp. 124–140.
[25] D. Karaboga, B. Gorkemli, C. Ozturk and N. Karaboga, "A comprehensive survey: artificial bee colony (abc) algorithm and applications", Artificial Intelligence Review, vol. 42, no. 1, **(2014)**, pp. 21-57.
[26] J. Fan, C. Wu and Y. Ma, "A modified partition coefficient", Signal Processing Proceedings, 2000. WCCC-ICSP 2000. 5th International Conference on, IEEE, vol. 3, **(2000)**, pp. 1496-1499.
[27] S. Sengupta, S. De, A. Konar and R. Janarthanan, "An improved fuzzy clustering method using modified Fukuyama-Sugeno cluster validity index", Recent Trends in Information Systems (ReTIS), 2011 International Conference on, IEEE, **(2011)**, pp. 269-274.
[28] D. Wodarz, E. Aescht and W. Foissner, "A weighted coenotic index (wci): description and application to soil animal assemblages", Biology and Fertility of Soils, vol. 14, no. 1, **(1992)**, pp. 5-13.