IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 06-10

A Review on Optimization Techniques for Fuzzy Clustering

K.Vijayakumari[1], V. Babydeepa[2]
[1]*Research Scholar, Department of Computer Science Government Arts College, Karur, Tamilnadu, India*
[2] *Assistant Professor, Department of Computer Science Government Arts College, Karur, Tamilnadu, India*

Abstract: *Fuzzy clustering is a one of the most fashionable problems in the data mining because of it is an NP hard problem. However, the conventional fuzzy clustering algorithms have many drawbacks, including the poor convergence rate, least classification accuracy. Hence, to solve the drawbacks of conventional methods, many research works have been developed using optimization techniques. This paper discussed the wide-range of studies focused on the fuzzy clustering using optimization algorithms and it helps the researcher how to enhance the efficiency of the clustering accuracy in the futures.*

Keywords: *Data mining, Fuzzy Clustering, Optimization Techniques*

I. Introduction

The theory of information is naturally related with the concept of uncertainty. The primary characteristic of an uncertainty is an involved in the situation of problems solving and it is a result of any information insufficiency such as incomplete, imprecise, fragmentary, undependable, formless, and conflicting. Uncertainty can be observed as an attribute of information. The fuzzy reasoning permits to handle much of this uncertainty which signifies uncertainty by numbers in the range between 0 and 1 [1].

On the other hand, clustering analysis is an essential tool for data analysis which is used to convert to the small groups from the unlabeled data based on their similarity[2]. Fuzzy clustering is an expansion compared with conventional clustering and it is an enlargement of the classical set. In the conventional clustering, each data points fall into exactly only one class. However, fuzzy clustering allows the data points into several classes at the same time. Each data points have a membership degree which conveys to what degree this object belongs to the various classes. In the fuzzy clustering, the fuzzy c-means (FCM) is well famous partitioned clustering algorithm which was developed by Bezdek,and etc., [3].

II. Problem Description

Let $O = \{o_1, o_2, ...o_n\}$ in R^d dimensional space into $c\ (1 < c < n)$ fuzzy clusters with $Z = \{z_1, z_2, ..., z_c\}$ cluster center. In the fuzzy clustering, each data points are described by fuzzy set μ with n rows and c column. μ indicates the degree of association or membership function of the i^{th} data points with the j^{th} cluster.

The fuzzy clustering is described by a fuzzy matrix μ with n rows and c columns in which n is the number of data objects and c is the number of features. μ_{ij}, the element in the i[th] row and j[th] column in μ, indicates the degree of association or membership function of the i[th] object with the j[th] cluster. The uniqueness of μ is as follows:

$$\forall_i\ [0, 1]\quad \forall_i = 1, 2\ldots n\ \forall_j = 1, 2\ldots c \tag{2}$$

$$\sum_{j=1}^{c} \mu_{ij} = 1 \quad \forall_i = 1, 2\ldots n \tag{3}$$

$$o < \sum_{j=1}^{c} \mu_{ij} < n \quad \forall_j = 1, 2\ldots c \tag{4}$$

The minimization of the objective function of FCM algorithms is defined as follows,

$$J_m = \sum_{j=1}^{c} \sum_{i=1}^{n} \mu_{ij} d_{ij} \tag{5}$$

Where

$$d_{ij} = \left| o_i - z_j \right| \tag{6}$$

In which, m (m>1) is a scalar termed the weighting exponent and controls the fuzziness of the resulting clusters and d_{ij} is the Euclidian distance from object $z = \{z_1, z_2, ..., z_n\}$ to the cluster center z_j. The z_j centroids of the j_{th} cluster, is obtained using below equation

$$z_j \frac{\sum_{i=1}^{n} \mu_{ij}^{m} o_i}{\sum_{i=1}^{n} \mu_{ij}^{m}} \tag{7}$$

Finally, the membership values may be updated as follows,

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} (\frac{d_{ij}}{d_{ik}})^{\frac{2}{m-1}}} \tag{8}$$

III. Fuzzy C Means

A fuzzy c-means is a famous partitioned clustering method which is developed by Bezdek et al. [4]. FCM is an iterative algorithm that finds clusters in data sets which use the concept of fuzzy membership instead of assigning a pixel to a single cluster and each pixel will have different membership values on each cluster. Given an unlabeled data set $O = \{O_1, O_2, ...O_n\}$ in R^d dimensional space, the FCM is to separate the many partition from given data set into *c (*$1 < c < n$*)* fuzzy clusters with $z = \{z_1, z_2, ..., z_n\}$ cluster centers. However, the conventional fuzzy clustering algorithms have many drawbacks, including the sensitivity to initialization, poor convergence rate. Many research works have been carried out to enhance the performance of the clustering efficiency and minimum computation cost. In [5] , the objective function of fuzzy clustering is to manage the groups and calculable parts of the clustering. Hence, the main aim of the fuzzy clustering algorithm is to reduce the objective function and usually it has number of calculation parts. The proposed new efficient fuzzy objective functions with efficient fuzzy parameters that can help to reduce the computation time and obtains more efficient groups.

There are many enhancement method has been applied to solve above mentioned defects exists in conventional FCM. The data reduction based FCM has applied to solve the clustering problem with higher accuracy in [6], the proposed model is able to decrease the number of individual patterns in order to enhance the partition quality. The proposed algorithm is applied to the problem of image segmenting on 32 magnetic resonance images. The noise sensitivity fault of FCM was eliminated with help of the probabilistic FCM [7]. In the image clustering, spatial information in the image is does not totally used in conventional FCM. Hence, utilize the image spatial information, a spatial functions is used in FCM [8]. However, the conventional and enhanced fuzzy c–means has many shortcoming including falling into local minima and an initialization of the cluster center by randomly. A new objective function of the proposed methods is developed to achieve the optimal clustering result by minimizing the distribution within clusters and exploiting the entropy of the weights of the attribute simultaneously[9]. The proposed approximate fuzzy c-means is an implemented based upon replacing the necessary terms in the FCM equation. The approximation facilitates AFCM to develop a search for table advance for computing and that has reduced the CPU time during each iteration [10]. Fuzzy kernel c means is dealing with some shortcoming in the conventional FCM[11], suppressed fuzzy c-means [12] and etc.,

The genetic algorithm combined with fuzzy c-means for solving document clustering [13], image clustering [14]. fuzzy c-means hybrid with fuzzy particle swarm optimization for clustering [15]. Such optimization algorithms are failing to get local minimum and global optima. The Bee Colony Optimization is population-based stochastic optimization algorithm and it has been attractive to develop due to its simple and effective [16, 17]. Possibilistic fuzzy c-means is a hybrid of both possibilistic c-means and conventional fuzzy c-means that frequently keep away from a variety of problems compared with PCM, FCM and FPCM. The proposed methods can solve the defect of noise sensitivity defect in the conventional FCM[18].

Feature weighting assignment can be allocated to a feature for demonstrating the significance of the feature. The proposed method an appropriate assignment of weight to the features and it can enhance the performance of conventional fuzzy c-means clustering algorithm. The assignment of weight is given by the gradient descent technique [19]. However, the problem of local optima could be present in the Fuzzy C-Means clustering algorithms. To attempt this shortcoming, the swarm intelligent algorithms have been employed to solve local optima problem and achieves global optima. Such as Genetic Algorithm [20], Simulated Annealing [21], Ant Colony Optimization [22], Particle Swarm Optimization [23, 24]. Social spider optimization [25].

IV. Particle Swarm Optimization

Hybridization of the Particle Swarm Optimization and fuzzy set theory is called fuzzy particle swarm optimization (FPSO) [26]. The position and velocity of particles of the fuzzy particle swarm optimization using the fuzzy membership function and it also applied for clustering problems [27]. In this method, x is the position of a particle, the fuzzy relation for the set of data objects, $O = \{o_1, o_2, \ldots o_n\}$ to the set of cluster centers $z = \{z_1, z_2, \ldots z_n\}$. In which technique, the position and velocity of particles redefined to represent the fuzzy relation between variables. In this part, the methods are described the system for fuzzy clustering problem[15]. In FPSO algorithm X, the position of a particle, shows the fuzzy relation from set of data objects, $o = \{o_1, o_2, \ldots., o_n\}$ to set of cluster centers, $z = \{z_1, z_2, \ldots., z_n\}$.

A novel chaotic fuzzy particle swarm optimization algorithm based on chaotic particle swarm (CPSO) and gradient method has proposed to solve a fuzzy clustering problem. An adaptive inertia weight factor (AIWF) and iterative chaotic map with infinite collapses (ICMIC) are introduced, The chaotic fuzzy particle swarm optimization clustering algorithm employs the CPSO to investigate the fuzzy clustering model, take advantage of the searching ability of fuzzy c-means (FCM) and prevent from the problem of local optima. Meanwhile, the accelerating the convergence speed of the proposed algorithm is used the gradient operator [5] .

Two latest techniques are applied to reduce the recreated objective functions of the FCM clustering methods with the help of particle swarm optimization (PSO) [28]. The balanced fuzzy particle swarm optimization (BFPSO) was applied to a combinatorial optimization problem [29]. A new particle swarm optimization based fuzzy clustering algorithm is discussed in order to solve the problems of FCM [30]. In [31], a fuzzy decision-making method is to contract with a clustering supplier problem in a system of supply chain. The proposed approach is to cluster the suppliers under fuzzy environments.

V. Bee Colony Optimization

A bee colony optimization algorithm is a meta-heuristic approach to solve the optimization problem in order to achieve global optimum[32]. In this model, it leads to the surfacing of cooperative behaviors intelligence of honey bee swarms and it consists of three fundamental components such as employed bees, onlookers and scouts. The first half of the colony consists of the employed artificial bees and the second half includes the onlookers. There is just one employ bee for every food sources. Around the hive, the quantity of employed bees are the same to the quantity of food sources[33].

The occupation of the employee bee is to search the food location in the food source and pass the information to the onlooker bee after collecting the location of food source. The most important role of the onlooker bees is obtaining the information from employee bee and measure the amount of nectar or food source quality. A food source is exploited by the onlooker bees depending on its amount of nectar. The role of scout bees is to investigate about the new food source in D-Dimensional search spaces. The fuzzy artificial bee algorithm (FABC) has been developed to solve the problem of clustering that the proposed method is to derive the behavior bee colonies to locate food in the local and global search to find the best clustering centers [34].

VI. Social Spider Optimization

The population based techniques have been proposed for solving optimization techniques based on the cooperative behavior of the social- spider by E. Cuevas[35]. The social spider optimization algorithm guesses that the entire search space is a communal web, whereas the social spiders are interacting with each others. Every social spider receives a weight according to the fitness value of the given solutions that is denoted by the social spiders. In the social spider optimization algorithms, there are two search agents, namely male and female.

Each individual is to carry out by a set of different evolutionary operators which mimic different cooperative behaviors that commonly assumed within the colony depending upon the gender. The attractive feature of the social-spider is the highly female-biased population. The social spider optimization algorithm is established by defining the number of male and female spiders that will be differentiated as individual in the search space. The traditional social spider optimization clustering algorithms is an integrated with fuzzy theory called fuzzy social spider optimization algorithm for solving the clustering problem as a proposal in [25].

VII. Hybrid Clustering Method

Hybridization of the clustering algorithm can be enhanced their quality of the clustering and can solve the demerits of both clustering algorithms. Major issues of the clustering problem are to solve local optima and need to find the global minimum. In some cases, the local optimum has also considered issues the clustering. Hence, both local and global optimum values are very significant task to obtain optimum clustering efficiency.

The following section discussed many research works that have been focused on hybridization methods to solve the fuzzy clustering problem.

Table 1: Summery of the reviews

S.No	Authors	Years	Methods	Objectives
1.	Robert l. [10]	1986	Approximating FCM	To enhance CPU time
2.	S. Eschrich et al., [6]	2003	Data Reduction based FCM	To reduce the computational complexity with higher accuracy
3.	Zhong-dong Wu [11]	2003	Kernal FCM	To enhance the performance of the clustering analysis
4.	Jiu-Lun Fan [12]		Suppressed FCM	To enhance the clustering efficiency
5.	X. Wang et al., [19]	2004	Improved FCM	To enhance the classification efficiency
6.	S. Eschrich et al., [18]	2005	Possibilistic FCM	To improve the classification accuracy
7.	Thomas A. Runkler et al., [28]	2006	PSO	To accelerate the performance of the FCM
8.	L. Li et al., [30]	2007	Fuzzy PSO	To avoid local optima
9.	Esmaeil Mehdizadeh [31]	2009	Fuzzy PSO	To enhance of clustering performance
10.	S.R. Kannan et al.,[42]	2011	FCM	To enhance the clustering efficiency
11.	K Vijayakumari [43]	2014	FBCO	To enhance clustering accuracy
12.	Jin Zhou et al., [9]	2016	Entropy-regularized weighted FCM (EWFCM)	To accomplishing the optimal clustering outcome
13.	Robati et al., [29]	2012	Balanced FPSO	To enhance the efficiency
14.	D. Karaboga [44]	2010	ABC	To enhance the clustering quality
15.	H. Izakian et al., [36]	2011	FCM+FPSO	To enhance the efficiency of the clustering and solve demerits of the both FCM and FPSO
16.	T. M. Silva Filho et al., [38]	2015	FCM + Improved FPSO	To achieve global optima
17.	A. Jain et al., [37]	2014	FCM+IFPSO	To achieve similar genes
18.	S. Chen [39]	2014	FCM + Improved FPSO	-
19.	P. Padmavathi et al [25]	2018	FSSO	To enhance the clustering accuracy

The FCM clustering algorithm is hybrid with fuzzy particle swarm optimization clustering algorithm used to enhance clustering quality[36]. Particle swarm optimization based FCM algorithm were proposed to solve gene clustering [37]. The hybridization of FCM with self-adaptive PSO is proposed as an improved method to solve clustering process and the proposed methods is to initialize one particle to achieve the best outcome with less iterations [38]. A hybrid clustering algorithm based on an enhanced PSO and FCM has proposed and that makes use of the merits of both algorithms [39]. The operator of the fuzzy c means comes at the phase of scout bee phase of the ABC algorithm to enhance the cluster quality [40]. Maximizing the classification and reduce the computation complexity are the two main goals of a fuzzy expert system based microarray data classification. The proposed methods employs artificial bee colony (ABC) algorithm to develop the points of membership functions. The proposed methods produced a precise fuzzy system with extremely interpretable and compressed rules for all the data sets [41]

VIII. Conclusion

This paper has commenced with the fundamental idea about the clustering and their explanations. The main intention of the paper is to introduce the fundamental and foundational idea of each commonly used clustering algorithm such as identify the source of each one, and analyze their advantages. Especially, the optimization based fuzzy clustering methods are very interesting to solve the fuzzy clustering problem. Hence, in this survey paper can help to beginners to do their research with clear manner.

References

[1]. P. Melin and O. J. A. s. c. Castillo, "A review on type-2 fuzzy logic applications in clustering, classification and pattern recognition," vol. 21, pp. 568-577, 2014.
[2]. A. A. Esmin, R. A. Coelho, and S. J. A. I. R. Matwin, "A review on particle swarm optimization algorithm and its variants to clustering high-dimensional data," vol. 44, pp. 23-45, 2015.
[3]. J. C. Bezdek, R. Ehrlich, W. J. C. Full, and Geosciences, "FCM: The fuzzy c-means clustering algorithm," vol. 10, pp. 191-203, 1984.
[4]. J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, pp. 191-203, 1984.
[5]. C. Li, J. Zhou, P. Kou, and J. J. N. Xiao, "A novel chaotic particle swarm optimization based fuzzy clustering algorithm," vol. 83, pp. 98-109, 2012.
[6]. S. Eschrich, J. Ke, L. O. Hall, and D. B. J. I. t. o. f. s. Goldgof, "Fast accurate fuzzy clustering through data reduction," vol. 11, pp. 262-270, 2003.

[7]. N. R. Pal, K. Pal, J. M. Keller, and J. C. Bezdek, "A possibilistic fuzzy c-means clustering algorithm," *IEEE transactions on fuzzy systems*, vol. 13, pp. 517-530, 2005.

[8]. K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, and T.-J. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," *computerized medical imaging and graphics*, vol. 30, pp. 9-15, 2006.

[9]. J. Zhou, L. Chen, C. P. Chen, Y. Zhang, and H.-X. J. N. Li, "Fuzzy clustering with the entropy of attribute weights," vol. 198, pp. 125-134, 2016.

[10]. R. L. Cannon, J. V. Dave, J. C. J. I. t. o. p. a. Bezdek, and m. intelligence, "Efficient implementation of the fuzzy c-means clustering algorithms," pp. 248-255, 1986.

[11]. Z.-d. Wu, W.-x. Xie, and J.-p. Yu, "Fuzzy c-means clustering algorithm based on kernel method," in *Computational Intelligence and Multimedia Applications, 2003. ICCIMA 2003. Proceedings. Fifth International Conference on*, 2003, pp. 49-54.

[12]. J.-L. Fan, W.-Z. Zhen, and W.-X. J. P. R. L. Xie, "Suppressed fuzzy c-means clustering algorithm," vol. 24, pp. 1607-1612, 2003.

[13]. Y. Gao, S. Wang, and S. Liu, "Automatic clustering based on GA-FCM for pattern recognition," in *Computational Intelligence and Design, 2009. ISCID'09. Second International Symposium on*, 2009, pp. 146-149.

[14]. F. T. Mathew, B. R. Seyi, and O. C. Akinwale, "Image Clustering using a Hybrid GA-FCM Algorithm 1," 2013.

[15]. H. Izakian and A. Abraham, "Fuzzy C-means and fuzzy swarm for fuzzy clustering problem," *Expert Systems with Applications*, vol. 38, pp. 1835-1838, 2011.

[16]. Z. N. Alqattan and R. Abdullah, "A comparison between artificial bee colony and particle swarm optimization algorithms for protein structure prediction problem," in *International Conference on Neural Information Processing*, 2013, pp. 331-340.

[17]. D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm," *Journal of global optimization*, vol. 39, pp. 459-471, 2007.

[18]. N. R. Pal, K. Pal, J. M. Keller, and J. C. J. I. t. o. f. s. Bezdek, "A possibilistic fuzzy c-means clustering algorithm," vol. 13, pp. 517-530, 2005.

[19]. X. Wang, Y. Wang, and L. J. P. r. l. Wang, "Improving fuzzy c-means clustering based on feature-weight learning," vol. 25, pp. 1123-1132, 2004.

[20]. U. Maulik and S. Bandyopadhyay, "Genetic algorithm-based clustering technique," *Pattern recognition*, vol. 33, pp. 1455-1465, 2000.

[21]. S. Z. Selim and K. Alsultan, "A simulated annealing algorithm for the clustering problem," *Pattern recognition*, vol. 24, pp. 1003-1008, 1991.

[22]. P. Shelokar, V. K. Jayaraman, and B. D. Kulkarni, "An ant colony approach for clustering," *Analytica Chimica Acta*, vol. 509, pp. 187-195, 2004.

[23]. S. C. Cohen and L. N. de Castro, "Data clustering with particle swarms," in *Evolutionary Computation, 2006. CEC 2006. IEEE Congress on*, 2006, pp. 1792-1798.

[24]. D. Saravanan and V. Eswaramurthy, "COMBINING INDEPENDENT COMPONENT ANALYSIS AND FUZZY PARTICLE SWARM OPTIMIZATION FOR FUZZY CLUSTERING."

[25]. P. Padmavathi, V. Eswaramurthy, and J. Revathi, "Fuzzy Social Spider Optimization Algorithm for Fuzzy Clustering Analysis," in *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)*, 2018, pp. 1-6.

[26]. W. Pang, K.-p. Wang, C.-g. Zhou, and L.-j. Dong, "Fuzzy discrete particle swarm optimization for solving traveling salesman problem," in *Computer and Information Technology, 2004. CIT'04. The Fourth International Conference on*, 2004, pp. 796-800.

[27]. L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, pp. 338-353, 1965.

[28]. T. A. Runkler and C. Katz, "Fuzzy clustering by particle swarm optimization," in *Fuzzy Systems, 2006 IEEE International Conference on*, 2006, pp. 601-608.

[29]. A. Robati, G. A. Barani, H. N. A. Pour, M. J. Fadaee, and J. R. P. J. A. M. M. Anaraki, "Balanced fuzzy particle swarm optimization," vol. 36, pp. 2169-2177, 2012.

[30]. L. Li, X. Liu, and M. Xu, "A novel fuzzy clustering based on particle swarm optimization," in *Information Technologies and Applications in Education, 2007. ISITAE'07. First IEEE International Symposium on*, 2007, pp. 88-90.

[31]. E. J. I. J. o. M. S. Mehdizadeh and E. Management, "A fuzzy clustering PSO algorithm for supplier base management," vol. 4, pp. 311-320, 2009.

[32]. D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical report-tr06, Erciyes university, engineering faculty, computer engineering department2005.

[33]. D. Karaboga, B. Gorkemli, C. Ozturk, and N. Karaboga, "A comprehensive survey: artificial bee colony (ABC) algorithm and applications," *Artificial Intelligence Review*, vol. 42, pp. 21-57, 2014.

[34]. D. Karaboga and C. Ozturk, "Fuzzy clustering with artificial bee colony algorithm," *Scientific research and Essays*, vol. 5, pp. 1899-1902, 2010.

[35]. E. Cuevas, M. Cienfuegos, D. Zaldívar, and M. Pérez-Cisneros, "A swarm optimization algorithm inspired in the behavior of the social-spider," *Expert Systems with Applications*, vol. 40, pp. 6374-6384, 2013.

[36]. H. Izakian and A. J. E. S. w. A. Abraham, "Fuzzy C-means and fuzzy swarm for fuzzy clustering problem," vol. 38, pp. 1835-1838, 2011.

[37]. A. Jain, S. Agrawal, J. Agrawal, and S. Sharma, "F-FDRPSO: A Novel Approach Based on Hybridization of Fuzzy C-means and FDRPSO for Gene Clustering," in *Proceedings of the Third International Conference on Soft Computing for Problem Solving*, 2014, pp. 709-719.

[38]. T. M. Silva Filho, B. A. Pimentel, R. M. Souza, and A. L. J. E. S. w. A. Oliveira, "Hybrid methods for fuzzy clustering based on fuzzy c-means and improved particle swarm optimization," vol. 42, pp. 6315-6328, 2015.

[39]. S. Chen, Z. Xu, Y. J. A. J. f. S. Tang, and Engineering, "A hybrid clustering algorithm based on fuzzy C-means and improved particle swarm optimization," vol. 39, pp. 8875-8887, 2014.

[40]. M. Krishnamoorthi and A. Natarajan, "Artificial bee colony algorithm integrated with fuzzy c-mean operator for data clustering," *Journal of Computer Science*, vol. 9, p. 404, 2013.

[41]. P. GaneshKumar, C. Rani, D. Devaraj, and T. A. A. Victoire, "Hybrid ant bee algorithm for fuzzy expert system based sample classification," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 11, pp. 347-360, 2014.

[42]. S. Kannan, R. Devi, S. Ramathilagam, and A. J. A. M. M. Sathya, "Some robust objectives of FCM for data analyzing," vol. 35, pp. 2571-2583, 2011.

[43]. K. Vijayakumari, M. Preetha, and K. Velusamy, "Performance Analysis of Clustering Based on Fuzzy System."

[44]. D. Karaboga, C. J. S. r. Ozturk, and Essays, "Fuzzy clustering with artificial bee colony algorithm," vol. 5, pp. 1899-1902, 2010.