Fuzzy C-means and fuzzy swarm for fuzzy clustering problem

Hesam Izakian [a,*], Ajith Abraham [b]

[a] Department of Computer Engineering, University Of Isfahan, Iran
[b] Machine Intelligence Research Labs, MIR-Labs, Washington, USA

ARTICLE INFO

ABSTRACT

Fuzzy clustering is an important problem which is the subject of active research in several real-world applications. Fuzzy c-means (FCM) algorithm is one of the most popular fuzzy clustering techniques because it is efficient, straightforward, and easy to implement. However, FCM is sensitive to initialization and is easily trapped in local optima. Particle swarm optimization (PSO) is a stochastic global optimization tool which is used in many optimization problems. In this paper, a hybrid fuzzy clustering method based on FCM and fuzzy PSO (FPSO) is proposed which make use of the merits of both algorithms. Experimental results show that our proposed method is efficient and can reveal encouraging results.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Clustering is the process of assigning data objects into a set of disjoint groups called clusters so that objects in each cluster are more similar to each other than objects from different clusters. Clustering techniques are applied in many application areas such as pattern recognition (Webb, 2002), data mining (Tan, Steinbach, & Kumar, 2005), machine learning (Alpaydin, 2004), etc. Clustering algorithms can be broadly classified as Hard, Fuzzy, Possibilistic, and Probabilistic (Hathway & Bezdek, 1995).

K-means is one of the most popular hard clustering algorithms which partitions data objects into k clusters where the number of clusters, k, is decided in advance according to application purposes. This model is inappropriate for real data sets in which there are no definite boundaries between the clusters.

After the fuzzy theory introduced by Lotfi Zadeh, the researchers put the fuzzy theory into clustering. Fuzzy algorithms can assign data object partially to multiple clusters. The degree of membership in the fuzzy clusters depends on the closeness of the data object to the cluster centers. The most popular fuzzy clustering algorithm is fuzzy c-means (FCM) which introduced by Bezdek (1974) and now it is widely used.

Fuzzy c-means clustering is an effective algorithm, but the random selection in center points makes iterative process falling into the local optimal solution easily. For solving this problem, recently evolutionary algorithms such as genetic algorithm (GA), simulated annealing (SA), ant colony optimization (ACO), and particle swarm optimization (PSO) have been successfully applied.

Particle swarm optimization (PSO) is a population-based optimization tool, which could be implemented and applied easily to solve various function optimization problems, or the problems that can be transformed to function optimization problems (Kennedy & Eberhart, 2001). Pang, Wang, Zhou, and Dong (2004), proposed a version of particle swarm optimization for TSP called fuzzy particle swarm optimization (FPSO). In this paper, a hybrid fuzzy clustering algorithm based on FCM and FPSO called FCM–FPSO is proposed. The experimental results over six real-life data sets indicate the FCM–FPSO algorithm is superior to the FCM algorithm and FPSO algorithm.

The rest of the paper is organized in the following manner. In Section 2, we investigate the related works and Section 3 introduces fuzzy c-means clustering, in Section 4 fuzzy PSO algorithm for clustering is discussed; Section 5 presents our hybrid clustering method, and Section 6 reports the experimental results. Finally, Section 7 concludes this work.

2. Related works

Runkler and Katz (2006), introduced two new methods for minimizing the reformulated objective functions of the fuzzy c-means clustering model by particle swarm optimization: PSO–V and PSOU. In PSO–V each particle represents a component of a cluster center, and in PSO–U each particle represents an unscaled and unnormalized membership value. Also they compared the proposed methods with alternating optimization and ant colony optimization. ZHAO (2007), presented an ant colony clustering algorithm for optimally clustering N objects into K clusters. The algorithm employs the global pheromone updating and the heuristic information to construct clustering solutions and uniform crossover operator to further improves solutions discovered by ants.

---

* Corresponding author.
E-mail addresses: hesam.izakian@gmail.com (H. Izakian), ajith.abraham@ieee.org (A. Abraham).

In Li, Liu and Xu (2007), in order to overcome the shortcomings of Fuzzy c-means, a PSO-based fuzzy clustering algorithm is discussed. The proposed algorithm uses the capacity of global search in PSO algorithm to overcome the shortcomings of FCM. Gan, Wu and Yang (2009), proposed the genetic fuzzy K–modes algorithm for clustering categorical data sets. They treated the fuzzy K–modes clustering as an optimization problem and used genetic algorithm to solve the problem in order to obtain globally optimal solution. To speed up the convergence process of the algorithm, they used the one-step fuzzy K–modes algorithm in the crossover process instead of the traditional crossover operator.

In Yang, Sun and Zhang (2009), a hybrid data clustering algorithm based on PSO and KHM is proposed, which makes use of the merits of both algorithms. The proposed method not only helps the KHM clustering escape from local optima but also overcomes the shortcoming of the slow convergence speed of the PSO algorithm. In Liu, Yih, Wu, and Liu (2008), authors used a Fuzzy c-mean algorithm based on Picard iteration and PSO (PPSO–FCM), to overcome the shortcomings of FCM.

3. Fuzzy c-means algorithm

Fuzzy c-means partitions set of n objects $o = \{o_1, o_2, \ldots, o_n\}$ in R^d dimensional space into c $(1 < c < n)$ fuzzy clusters with $Z = \{z_1, z_2, \ldots, z_c\}$ cluster centers or centroids. The fuzzy clustering of objects 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 clusters. μ_{ij}, the element in the ith row and jth column in μ, indicates the degree of association or membership function of the ith object with the jth cluster. The characters of μ are as follows:

$$\mu_{ij} \in [0,1] \quad \forall i = 1, 2, \ldots, n; \quad \forall j = 1, 2, \ldots, c \tag{1}$$

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

$$0 < \sum_{i=1}^{n} \mu_{ij} < n \quad \forall j = 1, 2, \ldots, c \tag{3}$$

The objective function of FCM algorithm is to minimize the Eq. (4):

$$J_m = \sum_{j=1}^{c} \sum_{i=1}^{n} \mu_{ij}^m d_{ij} \tag{4}$$

where

$$d_{ij} = \|o_i - z_j\| \tag{5}$$

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 o_i to the cluster center z_j. The z_j, centroid of the jth cluster, is obtained using Eq. (6).

$$z_j = \frac{\sum_{i=1}^{n} \mu_{ij}^m o_i}{\sum_{i=1}^{n} \mu_{ij}^m} \tag{6}$$

The FCM algorithm is iterative and can be stated as follows (Bezdek, 1974):

Algorithm 1. *Fuzzy c-means:*

1. Select m $(m > 1)$; initialize the membership function values μ_{ij}, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, c$.
2. Compute the cluster centers z_j, $j = 1, 2, \ldots, c$, according to Eq. (6).
3. Compute Euclidian distance d_{ij}, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, c$.
4. Update the membership function μ_{ij}, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, c$ according to Eq. (7).

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{ik}}\right)^{\frac{2}{m-1}}} \tag{7}$$

5. If not converged, go to step 2.

Several stopping rules can be used. One is to terminate the algorithm when the relative change in the centroid values becomes small or when the objective function, Eq. (4), cannot be minimized more. The FCM algorithm is sensitive to initial values and it is likely to fall into local optima.

4. Particle swarm optimization

Particle swarm optimization (PSO) is a population-based stochastic optimization technique inspired by bird flocking and fish schooling originally designed and introduced by Kennedy and Eberhart (1995) and is based on iterations/generations. The algorithmic flow in PSO starts with a population of particles whose positions represent the potential solutions for the studied problem, and velocities are randomly initialized in the search space. In each iteration, the search for optimal position is performed by updating the particle velocities and positions. Also in each iteration, the fitness value of each particle's position is determined using a fitness function. The velocity of each particle is updated using two best positions, personal best position and global best position. The personal best position, *pbest*, is the best position the particle has visited and *gbest* is the best position the swarm has visited since the first time step. A particle's velocity and position are updated as follows.

$$V(t+1) = w \cdot V(t) + c_1 r_1 (pbest(t) - X(t))$$
$$+ c_2 r_2 (gbest(t) - X(t)); \quad k = 1, 2, \ldots P \tag{8}$$

$$X(t+1) = X(t) + V(t+1) \tag{9}$$

where, X and V are position and velocity of particle respectively. w is inertia weight, c_1 and c_2 are positive constants, called acceleration coefficients which control the influence of *pbest* and *gbest* on the search process, P is the number of particles in the swarm, r_1 and r_2 are random values in range [0, 1].

4.1. Fuzzy particle swarm optimization for fuzzy clustering

Pang et al. (2004) proposed a modified particle swarm optimization for TSP called fuzzy particle swarm optimization (FPSO). In their proposed method the position and velocity of particles redefined to represent the fuzzy relation between variables. In this sub-section, we describe this method for fuzzy clustering problem.

In FPSO algorithm X, the position of particle, shows the fuzzy relation from set of data objects, $o = \{o_1, o_2, \ldots, o_n\}$, to set of cluster centers, $Z = \{z_1, z_2, \ldots, z_c\}$. X can be expressed as follows:

$$X = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1c} \\ \vdots & \ddots & \vdots \\ \mu_{n1} & \cdots & \mu_{nc} \end{bmatrix} \tag{10}$$

In which μ_{ij} is the membership function of the ith object with the jth cluster with constraints stated in (1) and (2). Therefore, we can see that the position matrix of each particle is the same as fuzzy matrix μ in FCM algorithm. Also, the velocity of each particle is stated using a matrix with the size n rows and c columns the elements of which are in range $[-1, 1]$. We get the Eqs. (11) and (12) for updating the positions and velocities of the particles based on matrix operations.

$$V(t+1) = w \times V(t) + (c_1 r_1) \times (pbest(t) - X(t)) + (c_2 r_2) \times (gbest(t) - X(t)) \tag{11}$$

$$X(t+1) = X(t) \oplus V(t+1) \tag{12}$$

After updating the position matrix, it may violate the constraints given in (1) and (2). So, it is necessary to normalize the position matrix. First, we make all the negative elements in matrix to

Table 1
Results of FCM, FPSO and FCM–FPSO methods on six real data sets.

Instances (n, c, d)	FCM			FPSO			FCM–FPSO		
	Worst	Average	Best	Worst	Average	Best	Worst	Average	Best
Iris (150, 3, 4)	71.58	70.43	67.92	69.72	67.39	66.26	**62.96**	**62.55**	**62.19**
Glass (214, 6, 9)	73.37	72.87	72.26	87.37	86.97	86.26	**73.11**	**72.64**	**72.23**
Cancer (683, 2, 9)	2235.8	2213.3	2196.8	2750.1	2724.4	2704.6	**2218.7**	**2190.5**	**2181.9**
Wine (178, 3, 13)	12192.5	11989.7	11682.7	12250.1	11528.8	11173.2	**11218.0**	**10603.9**	**10411.7**
CMC (1473, 3, 9)	3548.5	3534.7	3517.1	4190.1	4095.6	4025.2	**3531.2**	**3485.6**	**3416.5**
Vowel (871, 6, 3)	73390.8	71504.7	69069.1	100021.5	99394.0	98834.2	**68332.5**	**67872.6**	**67411.3**

Bold figures show the best results among others.

become zero. If all elements in a row of the matrix are zero, they need to be re-evaluated using series of random numbers within the interval [0, 1] and then the matrix undergoes the following transformation without violating the constraints:

$$Xnormal = \begin{bmatrix} \mu_{11}/\sum_{j=1}^{c}\mu_{1j} & \cdots & \mu_{1c}/\sum_{j=1}^{c}\mu_{1j} \\ \vdots & \ddots & \vdots \\ \mu_{n1}/\sum_{j=1}^{c}\mu_{nj} & \cdots & \mu_{nc}/\sum_{j=1}^{c}\mu_{nj} \end{bmatrix} \quad (13)$$

In FPSO algorithm the same as other evolutionary algorithms, we need a function for evaluating the generalized solutions called fitness function. In this paper, Eq. (14) is used for evaluating the solutions.

$$f(X) = \frac{K}{J_m} \quad (14)$$

therein K is a constant and J_m is the objective function of FCM algorithm (Eq. (4)). The smaller is J_m, the better is the clustering effect and the higher is the individual fitness $f(X)$. The FPSO algorithm for fuzzy clustering problem can be stated as follows:

Algorithm 2. *Fuzzy PSO for fuzzy clustering:*

1. Initialize the parameters including population size P, c_1, c_2, w, and the maximum iterative count.
2. Create a swarm with P particles (X, *pbest*, *gbest* and V are $n \times c$ matrices).
3. Initialize X, V, *pbest* for each particle and *gbest* for the swarm.
4. Calculate the cluster centers for each particle using Eq. (6).
5. Calculate the fitness value of each particle using Eq. (14).
6. Calculate *pbest* for each particle.
7. Calculate *gbest* for the swarm.
8. Update the velocity matrix for each particle using Eq. (11).
9. Update the position matrix for each particle using Eq. (12).
10. If terminating condition is not met, go to step 4.

The termination condition in proposed method is the maximum number of iterations or no improvement in *gbest* in a number of iterations.

5. Hybrid fuzzy c-means and fuzzy particle swarm optimization for clustering problem

The FCM algorithm is faster than the FPSO algorithm because it requires fewer function evaluations, but it usually falls into local optima. In this paper, the FCM algorithm integrated with FPSO algorithm to form a hybrid clustering algorithm called FCM–FPSO which maintains the merits of both FCM and PSO algorithms. FCM–FPSO algorithm applies FCM to the particles in the swarm every number of iterations/generations such that the fitness value of each particle is improved. The FCM–FPSO algorithm for fuzzy clustering problem can be stated as follows:

Algorithm 3. *Hybrid FCM–PSO for fuzzy clustering:*

1. Initialize the parameters of FPSO and FCM including population size P, c_1, c_2, w, and m.
2. Create a swarm with P particles (X, *pbest*, *gbest* and V are $n \times c$ matrices).
3. Initialize X, V, *pbest* for each particle and *gbest* for the swarm.
4. FPSO algorithm:
   4.1 Calculate the cluster centers for each particle using Eq. (6).
   4.2 Calculate the fitness value of each particle using Eq. (14).
   4.3 Calculate *pbest* for each particle.
   4.4 Calculate *gbest* for the swarm.
   4.5 Update the velocity matrix of each particle using Eq. (11).
   4.6 Update the position matrix of each particle using Eq. (12).
   4.7 If FPSO terminating condition is not met, go to step 4.
5. FCM algorithm
   5.1 Calculate the cluster centers for each particle using Eq. (6).
   5.2 Compute Euclidian distance d_{ij}, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, c$; for each particle using Eq. (5)
   5.3 Update the membership function μ_{ij}, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, c$; for each particle using Eq. (7).
   5.4 Calculate *pbest* for each particle.
   5.5 Calculate *gbest* for the swarm.
   5.6 If FCM terminating condition is not met, go to step 5.
6. If FCM–FPSO terminating condition is not met, go to step 4.

6. Experimental results

6.1. Parameter settings

In order to optimize the performance of the FPSO and FCM–FPSO, fine tuning has been performed and best values for their parameters are selected. Based on experimental results these algorithms perform best under the following settings: $c_1 = c_2 = 2.0$, $P = 10$, $w = 0.9 \rightarrow 0.1$. The FCM terminating condition in Algorithm 1 is when the algorithm cannot improve the generated solution, the FPSO terminating condition in Algorithm 2 is the maximum when the algorithm cannot improve the *gbest* in 1000 consecutive iterations, the FPSO terminating condition in Algorithm 3 is the maximum number of iterations 1000 or no changes in *gbest* in 200 consecutive iterations, and the FCM terminating condition in Algorithm 3 is the number of iterations 5. Also the FCM–FPSO terminating condition in Algorithm 3 is when the algorithm cannot improve the *gbest* in 2 consecutive iterations. In all of algorithms m, the weighting exponent, is set to 2.

6.2. Experimental results

For evaluating the FCM, FPSO, and FCM–FPSO methods, six well-known real-world data sets have been considered:

– Fisher's iris data set, which consists of three different species of iris flower. For each species, 50 samples with four features were collected;
– Glass, which consists of 214 objects and 6 different types of glasses. Each type has 9 features;
– Wisconsin breast cancer data set, which consists of 683 objects and 2 categories characterized by 9 features;
– Wine, which consists of 178 objects and 3 different types characterized by 13 features;
– Contraceptive Method Choice (CMC), which consists of 1473 objects and 3 different types characterized by 9 features and
– Vowel data set, which consists of 871 Indian Telugu vowel sounds, the data set has three features and six overlapping clusters.

These data sets cover examples of data of low, medium and high dimensions. These algorithms are implemented using VC++. The experimental results of over 100 independent runs for FCM and 10 independent runs for FPSO and FCM–FPSO are summarized in Table 1. The figures in this table are the objective function values (Eq. (4)).

As shown in this table, the hybrid FCM–FPSO obtained superior results than others in all of data sets and it can escape from local optima. Also, the experimental results show that when the size of data set (number of objects or clusters) is small, the FPSO surpasses FCM but with increasing the size of data set, FCM obtains better results than FPSO.

7. Conclusion

The fuzzy c-means algorithm is sensitive to initialization and is easily trapped in local optima. On the other hand, the particle swarm algorithm is a global stochastic tool which could be implemented and applied easily to solve various function optimization problems, or the problems that can be transformed to function optimization problems. In this paper, in order to overcome the shortcomings of the fuzzy c-means we integrate it with fuzzy particle swarm algorithm. Experimental results over six well-known data sets, Iris, Glass, Cancer, and Wine, show that the proposed hybrid method is efficient and can reveal very encouraging results in term of quality of solution found.

References

Alpaydin, E. (2004). *Introduction to machine learning.* Cambridge: MIT Press.
Bezdek, J. (1974). *Fuzzy mathematics in pattern classification.* Ph.D. thesis. Ithaca, NY: Cornell University.
Gan, G., Wu, J., & Yang, Z. (2009). A genetic fuzzy k-modes algorithm for clustering categorical data. *Expert Systems with Applications*(36), 1615–1620.
Hathway, R. J., & Bezdek, J. (1995). Optimization of clustering criteria by reformulation. *IEEE transactions on Fuzzy Systems*, 241–245.
Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of the IEEE international conference on neural networks* (pp. 1942–1948).
Kennedy, J., & Eberhart, R. (2001). *Swarm intelligence.* Morgan Kaufmann.
Li, L., Liu, X., & Xu, M. (2007). A novel fuzzy clustering based on particle swarm optimization. In *First IEEE international symposium on information technologies and applications in education* (pp. 88–90).
Liu, H.C., Yih, J.M., Wu, D. B., & Liu, S.W. (2008). Fuzzy C-mean clustering algorithms based on Picard iteration and particle swarm optimization. In *2008 international workshop on education technology and training & 2008 international workshop on geoscience and remote sensing* (pp. 838–842).
Pang, W., Wang, K., Zhou, C., & Dong, L. (2004). Fuzzy discrete particle swarm optimization for solving traveling salesman problem. In *Proceedings of the fourth international conference on computer and information technology* (pp. 796–800). IEEE CS Press.
Runkler, T. A., & Katz, C. (2006). Fuzzy clustering by particle swarm optimization. In *2006 IEEE international conference on fuzzy systems* (pp. 601–608). Canada.
Tan, P. N., Steinbach, M., & Kumar, V. (2005). *Introduction to data mining.* Boston: Addison-Wesley.
Webb, A. (2002). *Statistical pattern recognition.* New Jersey: John Wiley & Sons.
Yang, F., Sun, T., & Zhang, C. (2009). An efficient hybrid data clustering method based on K-harmonic means, and particle swarm optimization. *Expert Systems with Applications*(36), 9847–9852.
ZHAO, B. (2007). An ant colony clustering algorithm. In *Proceedings of the sixth international conference on machine learning and cybernetics* (pp. 3933–3938). Hong Kong.