ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMPROVED FUZZY CLUSTERING METHOD BASED ON INTUITIONISTIC FUZZY PARTICLE SWARM OPTIMIZATION

[1]V.KUMUTHA, [2]S. PALANIAMMAL

[1]D.J. Academy For Managerial Excellence, Department of Computer Science, Coimbatore, 032

[2]Sri Krishna College Of Technology, Department of Science & Humanities, Coimbatore, 042

E-mail: [1]kumuthav@gmail.com, [2]splvlb@yahoo.com

ABSTRACT

Recent advances in technology has led to huge growth in generating high dimensional data sets by capturing millions of facts in various fields, time phases, localities and brands. Microarray data contains gene expression from thousands of genes (features) from only tens of hundreds of samples. The rich source of information generated from microarray experiments often consist of incomplete and/or inconsistent data. Data mining is a powerful technology that automates the process of discovering hidden patterns. Traditional fuzzy clustering approaches are available which lacks to process efficiently in case of incomplete or inconsistent data. It has high influence over the resulting partitions. In this proposed approach, the degree of membership to indeterminacy is extended by adopting the concept of generalization of fuzzy logic, which is known as intuitionistic fuzzy logic. This paper proposes a hybrid approach for clustering high dimensional data set using FCM and Intuitionistic Fuzzy Particle Swarm Optimization (IFPSO) to overcome the local convergence problem. To find similarity among objects and cluster centers intuitionistic based similarity measure is used. Intuitionistic fuzzy particle swarm optimization optimizes the working of the Fuzzy c-means algorithm. Experimental results of proposed approach shows better results when compared with the existing methods.

Keywords: *Fuzzy Clustering; Intuitionistic Fuzzy; Particle Swarm Optimization; Gene Expression Data; Yeast Data; Degree Of Membership*

1. INTRODUCTION

The most common popular Data mining techniques discussed are clustering and classification[4]. Classification is a supervised learning method, whereas clustering is an unsupervised learning method. The goal of clustering is to ascertain new set of categories[12]. Clustering technique groups objects of similar pattern into one partition. Clustering techniques are broadly classified into hard and soft partition [7]. The traditional hard partitioning methods allow one object to lie in only one cluster at a time. The hard partition gives undesirable results, i) while fixing an object that almost lie between two clusters and ii) placing an outlier. This adverse situation can be fixed by fuzzy clustering. Fuzzy clustering allows one data item to belong to several clusters concurrently with different membership degrees. The assigning to a partition is determined by the membership degree that lies between 0 and 1[16].

Fuzzy c-means (FCM) is the most common fuzzy clustering algorithm. The algorithm uses objective function to measure the desirability of partitions. Fuzzy c-means is an effective algorithm, whereas the random selections of center point make iterative process falling into local optima solution hence different initializations may lead to different results.

Uncertainty is one of the major challenges posed by real-world clustering applications in the localization of the feature vectors. In microarray data, the number of samples is very limited while the volume of genes is very large; such data sets are very sparse in high-dimensional gene space. Moreover most of the genes collected may not necessarily be of interest. Uncertainty about which genes are relevant makes it difficult to select informative genes[5]. To handle this problem intuitionistic fuzzy approach is used. Intuitionistic Fuzzy Sets (IFSs)[1] are generalized fuzzy sets, which are useful in coping with the hesitancy originating from imperfect or imprecise information. Membership and non-membership

Journal of Theoretical and Applied Information Technology
10th April 2014. Vol. 62 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

value are elements involved in this sets. The degree of membership denotes the validity or trueness of the element to the set, whereas the non-validity of falseness of the element to the set denotes the non-membership value. Apart from validity and non-validity of the element, another element named hesitancy or indeterminacy or uncertainty poses difficulty in determining the validness of the membership of the element to the group. Recent research indicates that applying intuitionistic fuzzy sets to high dimensional data provides optimal clustering results.

Particle Swarm Optimization (PSO) is a population based search technique that share similar characteristics to Genetic Algorithm. It is an adaptive algorithm based on social-psychological metaphor; a population of individuals adapts by returning stochastically toward previously successful regions [8]. In this paper the FCM algorithm is combined with intuitionistic fuzzy particle swarm optimization taking the merits of both to give efficient results.

The paper is organized as follows. Section 2 briefs about the related works. Section 3 describes the materials and methods. Section 4 elaborates about the proposed work. Experimental results on data sets are given in section 5 and section 6 concludes the work.

2. RELATED WORK

There are many PSO based Fuzzy clustering methods available in the literature. [15] in their proposed work used the distance between the sample and cluster centers to distribute the membership in order to meet the constraints of FCM. The optimum particle has been directed to close the group in an optimized way. A chaotic particle swarm fuzzy clustering algorithm based on chaotic particle swarm and gradient method was proposed in [3]. Adaptive inertia weight factor and iterative chaotic map with infinite collapses are used. The method uses chaotic PSO to search the fuzzy clustering model, using the search capability of fuzzy c-means and thereby avoided the local convergence problem. The gradient operator is superior over the FCM algorithm.

An efficient hybrid method based on fuzzy particle swam optimization (FPSO) and Fuzzy C-Means (FCM) algorithms, to solve the fuzzy clustering problem, especially for large data sets was presented in [9]. The performance was improved by seeding the initial swarm with the result of the c-means algorithm. The experiment indicates that the computation times and solution quality of FPSO for large datasets was better than FCM.

In [10], PSO algorithm and fuzzy methods were combined to avoid local peaks and find global optimal solution. This approach uses global search capacity to overcome FCM deficits. It finds optimal location of clusters' centers for input dataset and finally finds the member components of each cluster.

A hybrid approach, in which fuzzy c-means clustering method and artificial neural networks were used in fuzzy time series to get more accurate forecasts were presented in [6]. Fuzzification step in FCM removes problems caused by partition of discourse of universe and fuzzy relationships defined by artificial neural networks avoids use of difficult matrix operations.

Two methods for minimizing the reformulated objective functions of the fuzzy c-means clustering model by particle swarm optimization: PSO–V and PSO-U. 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 were presented in [13]. The approach was compared with alternating optimization and ant colony optimization methods.

3. MATERIALS AND METHODS

3.1 Fuzzy c-means algorithm

Instead of assigning an object to a single cluster, the iterative method, Fuzzy C-Means algorithm (FCM) uses the concept of fuzzy membership, where each object will have different membership values on each cluster. It partitions set of n objects in R^d dimensional [2] space into c (1 < c < n) $O=\{o_1,o_2,\ldots o_n\}$ fuzzy clusters with $Z=\{z_1,z_2,\ldots .z_n\}$ 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, $\mu_{ij,}$ the element in the ith row and jth column in μ, point out the degree of association or membership function of the ith object with the jth cluster. The characters of μ are as follows:

$$\mu_{ij} \varepsilon\, [0,1]\ \forall j = 1,2,\ldots.,c \quad (1)$$

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

$$0 < \sum_{i=1}^{n} \mu_{ij} < 1 \forall_{ij} = 1,2,\ldots.c \quad (3)$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

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

$$J_m = \sum_{i=1}^{n} \sum_{j=1}^{c} \mu_{ij}^m d_{ij}^2 \quad 1 \leq m < \alpha \qquad (4)$$

where

$d_{ij} = |o_i - z_j|$ (5)

where m $(m>1)$ is a scalar termed the weighting exponent and controls the fuzziness of the resulting clusters and d_{ij} is the Euclidean distance from object o_i to the cluster center z_j. The z_j, centroid of the jth cluster, is obtained using Eq. (6).

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

Algorithm 1. Fuzzy c-means

1. Select m (m>1) and initialize the membership function values , μ_{ij} i=1,2,…n, j=1,2…c
2. Compute the cluster centers Z_j, j = 1,2,..., c by using Eq. (6)
3. Compute Euclidian distance, d_{ij}, i = 1,2,..., n; j=1,2,..., c
4. Update the membership function, μ_{ij} i = 1, 2... n; j=1,2,..., c by using below equation μ_{ij}

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

If not converged, go to step 2.

3.2. PSO

Particle swarm optimization (PSO) is a population-based stochastic optimization technique inspired by bird flocking and fish schooling [8] which 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. 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.

$V(t+1)=wV(t)+c_1r_1(pbest(t)-X(t))+c_2r_2(gbest(t)-X(t))$; (8)

$X(t+1)=X(t)+V(t+1)$ (9)

where, X and V are position and velocity of particle respectively. w is inertia weight, c1 and c2 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].

3.2.1 Fuzzy PSO Algorithm

A particle swarm optimization with fuzzy set theory is called fuzzy particle swarm optimization (FPSO) [11]. Using fuzzy relation between variables, FPSO redefines the position and velocity of particles and its also applied for clustering problem. In this method X is the position of particle, the fuzzy relation for the set of data objects $O=\{o_1,o_2,\ldots o_n\}$, to set of clusters centers $Z=\{z_1,z_2,\ldots z_n\}$can be expressed as follows

$$X = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1c} \\ \cdots & \cdots & .. \\ \mu_{n1} & \cdots & \mu_{nc} \end{bmatrix} \qquad (10)$$

Here, μ_{ij} is the membership function of the ith object with the jth cluster with constraints

$$\forall_i [0,1] \forall_i = 1,2\ldots n \forall_j = 1,2\ldots c \qquad (11)$$

$$\sum_{j=1}^{c} \mu_{ij} = 1 \forall_i = 1,2\ldots n \qquad (12)$$

therefore it is known 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 between -1 and 1.

The equations (13) and (14) are used for updating the positions and velocities of the particles based on the matrix.

$$V(t+1) = wV(t) + c_1 r_1 \times pbest(t) - X(t) + c_2 r_2 \times gbest(t) - X(t) \qquad (13)$$

$$X(t+1) = X(t) \oplus V(t+1) \qquad (14)$$

After updating the position matrix, it may violate the constraints given in (11) and (12) since it is compulsory to normalize the position matrix. First all the negative elements in matrix are set to zero. If all elements in a row of the matrix are zero, they need to be reevaluated using series of random numbers within the interval between 0 and 1, and then the matrix undergoes the following transformation without violating the following constraints:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$Xnormal = \begin{bmatrix} \mu_{11} \Big/ \sum_{j=1}^{c} \mu_{1j} & \cdots & \mu_{1c} \Big/ \sum_{j=1}^{c} \mu_{1j} \\ \cdots & \cdots & \cdots \\ \mu_{n1} \Big/ \sum_{j=1}^{c} \mu_{1j} & \cdots & \mu_{nc} \Big/ \sum_{j=1}^{c} \mu_{1j} \end{bmatrix} \quad (15)$$

This technique uses the following equation as fitness function for evaluating the solutions.

$$f(X) = \frac{K}{J_m} \quad (16)$$

Here, K is a constant and J_m is the objective function of FCM algorithm. The smaller is J_m, the better is the clustering effect and the higher is the individual fitness $f(X)$. The termination condition in this method is the maximum number of iterations or no improvement in gbest in a number of iterations. The FCM algorithm is quicker than the FPSO algorithm because it need not incur as much of function evaluations, but it normally go down into local optima. FCM algorithm incorporated with FPSO algorithm to form a hybrid clustering algorithm called FCM-FPSO which maintains the merits of both FCM and FPSO algorithms.

Algorithm 2. Fuzzy PSO

Input : Dataset
Output : Objective Values
Step 1. Initialize the parameters including population size P, c1, c2, w and the maximum iterative count.
Step 2. Create a swarm with P particles (X, pbest, gbest and V are n × c matrices).
Step 3.Initialize X, V, pbest for each particle and gbest for the swarm.
Step 4. Calculate the cluster centers for each particle using

$$z_j = \frac{\sum_{i=1}^{n} \mu_{ij}^{m} o_i}{\sum_{i=1}^{n} \mu_{ij}^{m}} \quad (17)$$

Step 5. Calculate the fitness value of each particle using Eq. (16)
Step 6. Calculate pbest for each particle.
Step 7. Calculate gbest for the swarm.
Step 8. Update the velocity matrix for each particle using Eq. (13)
Step 9. Update the position matrix for each particle usingEq. (14)
Step 10. If terminating condition is not met, go to step 4.

3.2.2Intuitionistic Fuzzy Set

The Intuitionistic Fuzzy Set (IFS) was defined as an extension of the ordinary Fuzzy Set [2] [14]. As opposed to a fuzzy set in X, given by:

$A = \{(x, \boldsymbol{\mu_A} (x)) \mid x \in \boldsymbol{X} \}$ (18)

where$\mu_A(x) \rightarrow [0,1]$ is the membership function of the fuzzy set A, an intuitionistic fuzzy set B is given by:

$B = \{ x, \boldsymbol{\mu_B} (x)), \nu_B\boldsymbol{(x)} \mid x \in \boldsymbol{X}\}$ (19)

where$\mu_B(x) \rightarrow [0,1]$ and $\nu_B (x) \rightarrow [0,1]$ are such that:

$0 \leq \boldsymbol{\mu_B} (x)), \nu_{B(x)} \leq 1$ (20)

and$\mu_B(x)$,$\nu_B(x)$ $\in [0,1]$ denote degrees of membership and non-membership of x ∈ B, respectively.

For each intuitionistic fuzzy set B in X, ‘‘hesitation margin’’ (or ‘‘intuitionistic fuzzy index’’) of $x \in B$ is given by:

$\pi_B (x) = 1 - \boldsymbol{\mu}(x) - \nu_B(x)$ (21)

which expresses a hesitation degree of whether x belongs to B or not. It is obvious that $0 \leq \pi_B(x) \leq 1$, for each x ∈ X. To describe an intuitionistic fuzzy set completely, it is necessary to use any two functions from the triplet: membership function; non-membership function; and hesitation margin.

4. PROPOSED APPROACH

4.1 Intuitionistic Fuzzy Particle Swarm Optimization (IFPSO)

Existing approaches works fine for the data-sets which are not corrupted with noise but if the dataset is noisy or distorted then it wrongly classifies noisy pixels because of its abnormal feature data and results in an incorrect membership and improper clustering.

The above said problem was also faced by the fuzzy particle swarm optimization approach. To overcome the problem of abnormal features that exist among the particle clustering can be overwhelmed by introducing the concept of intuitionistic fuzzy based particle swarm optimization which is the generalization of fuzzy based particle swarm optimization. In this approach each particle is concerned not only with the membership function but the in deterministic degree is also taken into consideration for handling the abnormality problem. The abnormality problem arises due to the inconsistency of the particles position information.

Journal of Theoretical and Applied Information Technology
10th April 2014. Vol. 62 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

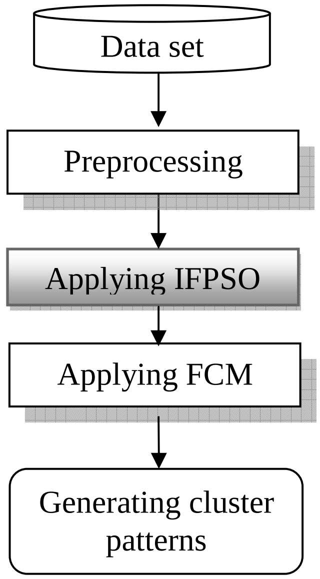

Figure 1: Framework of Proposed Work

This proposed approach IFPSO introduces another degree into the consideration for handling the uncertainty problem among the particles using hesitation degree which is also known as in deterministic degree.

$$X = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1c} \\ \cdots & \cdots & \cdots \\ \mu_{n1} & \cdots & \mu_{nc} \end{bmatrix} \quad (22)$$

Algorithm 3. IFPSO algorithm

Input : Dataset
Output : Objective Values
Step 1. Initialize the parameters including population size P, c1, c2, w and the maximum iterative count.
Step 2. Create a swarm with P particles (X, pbest, gbest and V are n × c matrices).
Step 3.Initialize X, V, pbest for each particle and gbest for the swarm.
Step 4. Calculate the cluster centers for each particle using

$$Z_j = \sum_{k=1}^{n} \frac{\mu_{ik}^{*} x_{ik}}{\sum_{k=1}^{n} \mu_{ik}^{*}} \quad (23)$$

Step 5. Calculate the fitness value of each particle using Eq. (16)
Step 6. Calculate pbest for each particle.
Step 7. Calculate gbest for the swarm.
Step 8. Update the velocity matrix for each particle using Eq. (13)
Step 9. Update the position matrix for each particle using Eq. (14)
Step 10. If terminating condition is not met, go to step 4.

4.2. Hybrid Fuzzy C Means and Intuitionistic Fuzzy Particle Swarm Optimization

The FCM algorithm is quicker than the IFPSO algorithms because it uses few function evaluations, but it normally go down into local optima. Rather, FCM algorithm incorporated with IFPSO algorithm to form a hybrid clustering algorithm called FCM-IFPSO which maintains the merits of both FCM and IFPSO algorithms. In FCM-IFPSO algorithm, FCM is applied to the particles in the swarm every number of iterations/generations such that the fitness value of each particle is improved. The algorithm 4 illustrate hybrid FCM-IFPSO.

Algorithm 4. FCM-IFPSO algorithm

Input : Dataset
Output : Objective Values
Step 1. Initialize the parameters of IFPSO and FCM including population size P, c1, c2, w, and m.
Step 2. Create a swarm with P particles (X, pbest, gbest and V are n× c matrices).
Step 3. Initialize X, V, pbest for each particle and gbest for the swarm
Step 4. IFPSO algorithm
4.1 Calculate the cluster centers for each particle using by (23)
4.2 Calculate the fitness value of each particle using by (16)
4.3 Calculate pbest for each particle.
4.4 Calculate gbest for the swarm.
4.5 Update the velocity matrix for each particle using by (13)
4.6 Update the position matrix for each particle using by (14)
4.7 If terminating condition is not met, go to step 4
Step 5. FCM algorithm
5.1 Compute the cluster centers *j z* , j = 1,2,..., c , by using (13)
5.2 Compute Euclidean distance, d_{ij}, i = 1,2,..., n; j=1,2,..., c
5.3 Update the membership function μ*ij*, i = 1,2…n

j=1,2,...,c $$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{ik}} \right)^{\frac{2}{m-1}}} \quad (24)$$

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.

Journal of Theoretical and Applied Information Technology
10th April 2014. Vol. 62 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Step 6. If FCM-IFPSO terminating condition is not met, go to Step 4.

5. EXPERIMENTAL RESULTS

The algorithms discussed in the previous section have been implemented using MATLAB. For evaluating the performance of the proposed work, four different benchmark datasets are taken into consideration.

5.1 Parameter settings

The optimized performance of the IFPSO and FCM-IFPSO, fine tuning has been executed and best values for their parameters are chosen. The experimental results based on these algorithms achieve best under the following settings: c1,c2, the value is3.0 - population is 12, and weight values are: the minimum of weight value is 0.2 and maximum of value is 0.8 and weighting exponent component m value is 2 which is common to all the algorithms. The FPSO and IFPSO terminating condition is till it reaches the maximum iteration value when the algorithm cannot improve the gbest in 500 consecutive iterations. Also the FCM–IFPSO terminating condition is met, when the algorithm cannot improve the gbest in 2 consecutive iterations.

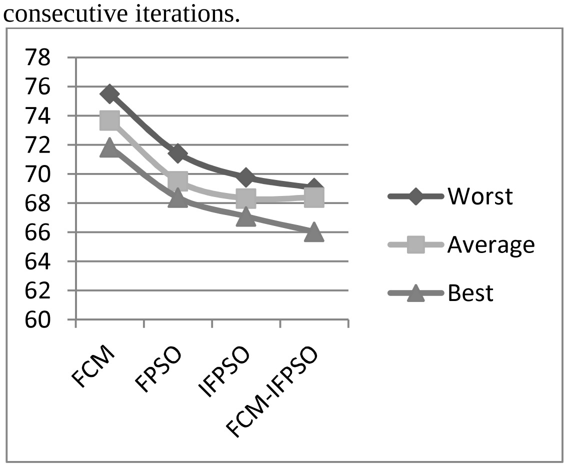

Figure 2:Results for Iris data set

The experimental results of over 100 independent runs for FCM and 10 independent runs for IFPSO and FCM–IFPSO are shown in the figures 2,3,4 & 5. Figure 2 shows the best, average and worst cases of objective function values obtained applying over iris data set for FCM, FPSO, IFPSO and FCM-IFPSO algorithms. The proposed hybrid approach FCM-IFPSO method depicts better results over the other existing methods.

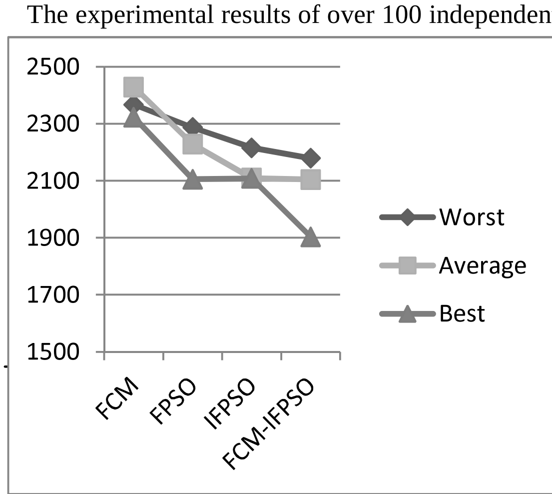

Figure 3:Results for yeast data set

Figure 3 shows the best, average and worst cases of objective function values obtained applying over yeast data set for FCM, FPSO, IFPSO and FCM-IFPSO algorithms.

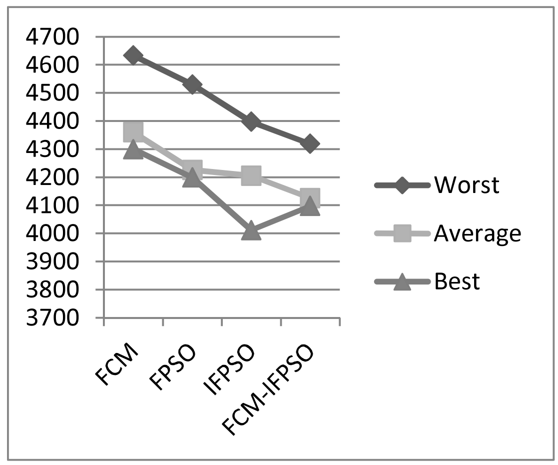

Figure 4:Results for colon cancer data set

Figure 4 shows the best, average and worst cases of objective function values obtained applying over colon cancer data set for FCM, FPSO, IFPSO and FCM-IFPSO algorithms. The proposed approach depicts better results when compared to the existing approaches and it can flee from local optima.

Figure 5shows the best, average and worst cases of objective function values obtained applying over leukemia data set for FCM, FPSO, IFPSO and FCM-IFPSO algorithms. The proposed approach depicts better results when compared to the existing approaches.

Journal of Theoretical and Applied Information Technology
10th April 2014. Vol. 62 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

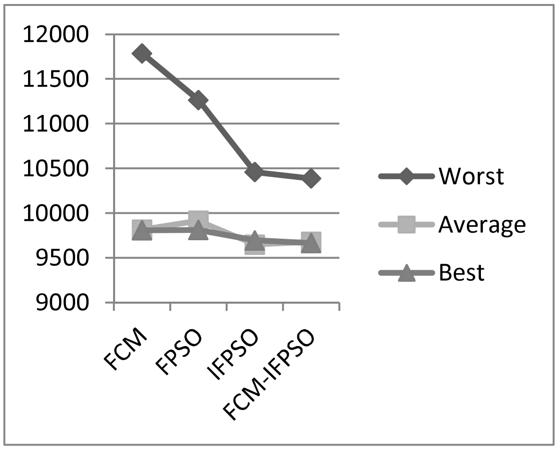

Figure 5:Results for leukemia data set

Table 1 describes the number of samples and genes of the microarray dataset namely, yeast, colon cancer and leukemia datasets. Table 2 describes the number of attributes and instances of the benchmark iris dataset.

Table 1. Gene expression dataset

Data set	No. of samples	No. of genes
Yeast	79	2467
Colon cancer	62	2000
Leukemia	72	7129

Table 2. Benchmark dataset

Data set	No. of attributes	No. of instances
Iris	4	150

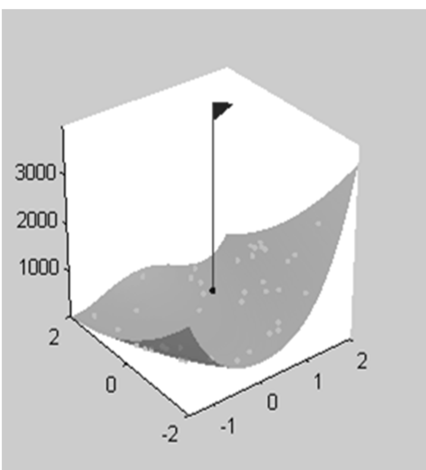

Figure 6: Average cumulative change for 129 generations

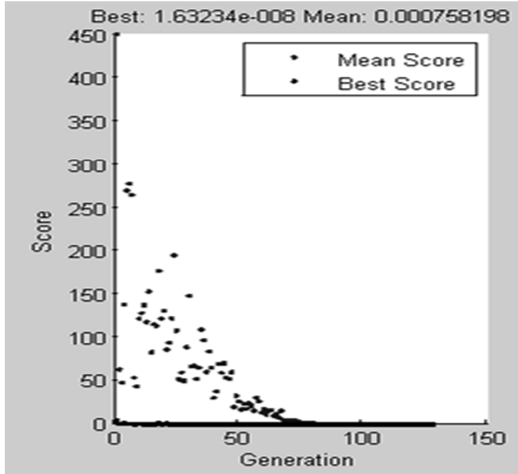

Figure 7:Mean and best score value for 129 generations

The above figure shows the various generations runs of proposed FCM-IFPSO algorithm with the mean and best score value.

The figure 6 shows Average cumulative change in value of the fitness function over 50 generations less than 1e-006 and constraint violation less than 1e-006, after 129 generations. Figure 7 shows the mean and best score over 129 generations (Final best point: [0.99999 0.99998]).

6. CONCLUSION

Data mining poses a vital issue in clustering of high dimensional gene expression data. Several algorithms have been addressed in the literature for handling such data. The fuzzy c-means algorithm is sensitive to initialization and is easily trapped in local optima. On the other hand, the fuzzy particle swarm algorithm is a global stochastic tool which could be implemented and applied easily to solve various function optimization problems. Both of them fail to handle the uncertainty condition and they left the concept of indeterminacy when there is a presence of vagueness or incompleteness in clustering dataset.

In this paper, an optimization approach is put forward in order to overcome the shortcomings of the fuzzy c-means and fuzzy particle swarm optimization. The proposed method removes the indeterminacy thereby a good performance of the desired clusters is obtained. Instead of just considering the membership value of each object in the cluster, the proposed approach takes into account the in- deterministic value as an important factor in the case of incompleteness in the dataset clustering. Experimental results over well-known data sets, Iris, Yeast, Colon cancer, and Leukemia,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

show that the proposed hybrid method is proficient and can reveal very encouraging results in term of quality of solution found. A new hybrid method combining fuzzy c-means and Intuitionistic fuzzy Particle swarm optimization algorithm have been applied successfully for real world datasets. The computational results show that the performance of the proposed algorithm is better than the other existing algorithms.

Our future work would be directed towards Intuitionistic fuzzy clustering combined with intuitionistic fuzzy particle swarm optimization in order to achieve better results for microarray data.

REFRENCES:

[1] Atanassov, K., "Intuitionistic fuzzy sets: past, present and future", Proceedings of the 3rd Conference of the European Society for Fuzzy Logic and Technology, 2003, pp. 12–19.

[2] Atanassov, K.T., "Intuitionistic Fuzzy Sets Theory and Applications", *Studies in Fuzziness and Soft Computing*, 1999, Phisica-Verlag.

[3] Chaoshun, L., Jianzhong, Z., Pangao, K., Jian, X., "A novel chaotic particle swarm optimization based fuzzy clustering algorithm", *Neuro computing*, Vol. 83, No.15, 2012, pp. 98–109.

[4] Chen, M.S., Han, J., Yu, P.S., "Data mining: An Overview from Database Perspective", *IEEE Trans. Knowl. Data Eng.*, 1996.

[5] Daxin, J. C., Tang A. Z., "Cluster Analysis for Gene Expression Data: A Survey", *IEEE Trans. Knowl. Data Eng.*, Vol. 16, No. 11, 2004, pp. 1370-1386.

[6] Erol, E., Cagdas, H. A., Ufuk, Y., "Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks", *Expert Systems with Applications*, Vol. 40, 2013, pp. 54.

[7] Jain, A.K., Mutry, M.N., Flynn, P.J., "Data Clustering: A Review", *ACM Computing Surveys*, Vol. 31, No. 3, 1999, pp. 264-323.

[8] Kennedy, J., Eberhart, R., "Swarm Intelligence", *Morgan Kaufmann Publishers, Inc.*, 2001, San Francisco, CA.

[9] Mehdizadeh, E., Sadi-Nezhad, S., Tavakkoli-Moghaddam, R., "Optimization Of Fuzzy Clustering Criteria By A Hybrid PSO And Fuzzy C-Means Clustering Algorithm", *Iranian Journal of Fuzzy Systems*, Vol. 5, No. 3, 2008, pp. 1-14.

[10] Mir, M., Tadayon Tabrizi, G., "Improving Data Clustering Using Fuzzy Logic and PSO Algorithm", *20thIranian Conference on Electrical Engineering*, 2012, Tehran, Iran

[11] Pang, W., Wang, K., Zhou, C., Dong, L. "Fuzzy discrete particle swarm optimization for solving traveling salesman problem", *Proceedings of the fourth international conference on computer and information technology*, IEEE CS Press 2004, pp.796–800.

[12] Rokach. L., Maimon, O., "Clustering methods", *Data Mining and Knowledge Discovery Handbook*, Springer, New York, 2005, pp. 321–352.

[13] Runkler, T. A., Katz, C., "Fuzzy clustering by particle swarm optimization", *IEEE international conference on fuzzy systems*, 2006, pp. 601–608.

[14]Shanthi, S., Murali Bhaskaran, V. "Intuitionistic Fuzzy C-Means and Decision Tree Approach for Breast Cancer Detection and Classification", *European Journal of Scientific Research*, Vol.66 No.3, 2011, pp. 345-351.

[15] Qiang, N., Xinjian, H., "An Improved Fuzzy C-means Clustering Algorithm based on PSO", *Journal of Software*, Vol. 6, No. 5, 2011.

[16] Zadeh, L.A. "Fuzzy sets", *Information Control*, Vol. 8, 1965, pp. 338–356.