Clustering Optimization using Hybrid IFA-PSO and Kernel-Based FCM

Lokesh Gagnani[1], Hitesh Chhinkaniwala[2]

[1] Department of Computer Engineering, C U Shah University, Wadhwan, Gujarat, India
[2] Computer Engineering Department, Adani Institute of Infrastructure Engineering, Ahmedabad, Gujarat, India

[1]gagnani.lokesh@gmail.com, [2]drhiteshrc1@gmail.com

***Abstract*— Clustering of datasets is achieved with a metaheuristic method. We have proposed a method that hybridizes Kernel-based Fuzzy C Means (KFCM), Particle Swarm Optimization (PSO), and Intelligent Firefly Algorithm (IFA). The new method consists of two stages: firstly, a KFCM approach for generation of initial value of object function is developed, and then hybrid IFA-PSO is applied to find the optimal values of object function. The proposed algorithm increases the possibility of finding optimal values of object function (OF) and does not get trapped in local minima. Further IFA has better results due to its global optimization nature. The object function generates optimal cluster partitions in KFCM iterative process. Experimental results indicate that the proposed KFCM-IFA-PSO model achieves better optimization of object function than KFCM-BFO-PSO. Bacteria Foraging Optimization(BFO) is inspired by bacteria's food search behavior but gets trapped in local optima. Therefore, we can conclude that our proposed method is very efficient compared to the previously reported algorithm**.

Keywords— clustering, soft computing, object function, KFCM, IFA, PSO, BFO.

I. INTRODUCTION

Clustering or cluster analysis is the widespread classification technique in unsupervised category. Clustering algorithms have practical use in various applications[2][6][7][11], including social network analysis, robotics, crime analysis, image processing, etc. In learning ,clustering can be expressed as distribution of n objects if the number of clusters is known in advance. Researchers have applied artificial intelligence and computer technology to develop optimization systems, which improve the efficiency of clustering algorithm and in specific the object function. In other words clustering is an unsupervised classification recognition method based on objective function, which is used to partition the objects into different clusters (groups). There is a high degree of similarity among objects in same cluster and high degree of dissimilarity among objects in different clusters.

K-means is one of the well known hard clustering algorithms which takes the input *k clusters*, and divides a collection of *n* objects into *k* clusters. Here intra-cluster distance is high but the inter-cluster distance is low. Cluster likeness can be calculated by the cluster center or by the mean value of the objects.

Fuzzy C-Means (FCM) is the most well known method in cluster analysis, developed by Bezdek [1]. FCM gets trapped in local minimum as there is searching by hill climbing and random initialization of centroids. However FCM gets stuck due to overlapping of clusters. Degree of membership is associated with closeness factor of data to the centers.

Kernel methods transform the patterns from the input space to a new space (called Kernel space), in a way that, the patterns will became more linearly separable. Kernel functions are functions that are able to calculate the dot product between the patterns in kernel space by using their values in input space[4].

Kernel-based FCM resolves the overlapping of clusters of FCM. KFCM algorithm has some advantage in comparison to FCM; firstly, it is more robust to noise and less sensitive to shape of clusters[3]. Secondly, it has the same computational complexity as FCM with a more rapid convergence[4]. Since it uses alternative optimization method for optimizing the object function, it can only assure of finding local minima.

PSO, a swarm intelligence, was invented by Russell and James. Based on the flocking and schooling patters of birds and fish.

IFA is a new swarm intelligent optimization algorithm. It is enhanced version of Firefly Algorithm (FA) developed by Yang[9]. The idea behind IFA is the use of the ranking information for attractiveness. IFA algorithm is able to provide a new method for the global parameter optimization. This paper proposes a clustering model for global optimization of object function based on KFCM, PSO, and IFA.

Firstly introduction to relevant methods and concise literature survey of above mentioned methods, and then presents the novel model, KFCM-IFA-PSO. Section 3 gives results of different methods in six different datasets (binary and multi class) from University of California Irvine Machine Learning Repository. Finally, conclusions are presented in Section 4.

II. TECHNIQUES AND DATASET

A. Techniques

1) KFCM. Clustering based on kernel function is firstly proposed by Girolami[5]. After that Chen and Zhang[3] proposed a modification of FCM algorithm using kernel function and called it KFCM.

Similar to FCM algorithm, KFCM clusters the input date by optimizing the object function defined as

$$J = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} ||\emptyset(x_j) - \emptyset(c_i)||^2 \quad (1)$$

where

$$||\emptyset(x_j) - \emptyset(c_i)||^2 = 2 - 2K(x_j, c_i) \quad (2)$$

The Gaussian Radial Basis (RBF) Kernel is defined as:

$$K(x_j, c_i) = e^{\frac{-||x_j - c_i||^2}{2\sigma^2}} \quad (3)$$

Thus objective function can be modified using Eq (2) as :

$$J = 2\sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} (1 - K(x_j, c_i)) \quad (4)$$

The membership u_{ij}^{m} is calculated as:

$$u_{ij} = \frac{(1-K(x_j,c_i))^{\frac{-1}{(m-1)}}}{\sum_{i=1}^{c}(1-K(x_j-c_i))^{\frac{-1}{(m-1)}}} \quad (5)$$

The center c_i is calculated as:

$$c_i = \frac{\sum_{j=1}^{n} K(x_j,c_i)x_j}{\sum_{j=1}^{n} u_{ij}^{m} K(x_j,c_i)} \quad (6)$$

Here, m is the fuzziness factor and m=1 means simple FCM. x_j is the data from dataset. Initially random cluster centers are chosen and then recalculated with respect to minimizing the Object function. The advantage of KFCM over traditional FCM is that KFCM has the ability to link to kernel domain. σ is a constant to establish a balance between used dimensionality and clustering accuracy. Minimizing objective function means increasing similarity among all the components within an object and reducing similarity between components of one object with others.

2) PSO. PSO is a swarm intelligence technique mainly for optimization purpose. A particle is represented as a single solution in PSO search space. Fitness function is indicated by fitness values of all particles and velocities by the flying of particles. In every step, a particle encounters with two values. First is the local best solution and another is the global best solution.

PSO is widely used optimization technique in various applications due to less parameters and simplicity than other optimization techniques in its category. Particles fitness value can be evaluated by the input values which are particle positional coordinate. Further Vi and Xi are the velocities and position values of particles respectively. Then the particles updates their velocity and position using the equations below:

$$V_i(i+1) = \omega.V_i(i) + C1.\emptyset_1.(P_{best} - X_i(i)) + C2.\emptyset_2.(G_{best} - X_i(i)) \quad (7)$$

$$X_i(i+1) = X_i(i) + V_i(i+1) \quad (8)$$

3) IFA. In classical FA (Firefly Algorithm)[12], the value of objective (fitness) function is determined by the brightness of a firefly. The brightness of firefly is equal to the attractiveness of a firefly which is give by I(x) = f(x). Let the r be the distance at x_i & y_j as Cartesian metric.

Light intensity is represented as:

$$\beta = \beta_{min} + (\beta_0 - \beta_{min})e^{-\gamma r^2} \quad (9)$$

The Cartesian metric:

$$r_{ij} = x_i - x_j = \sqrt{\sum_{k=1}^{d}(x_{i,k} - x_{j,k})^2} \quad (10)$$

Update movement of a firefly consists of 2 parts:(1) Approach better solution (β) and (2) move randomly (α)

$$x_i = x_i + \beta(x_j - x_i) + \alpha \in_i \quad (11)$$

The main logic of our Intelligent Firefly Algorithm (IFA)[9] is to make use of the ranking information such that every firefly is moved by the attractiveness of a fraction of fireflies only and not by all of them. Ø is a highest portion of the fireflies based on their rank. Hence it moves actively not just by attractiveness but also on the top ranking move position of fireflies.

Ø is the parameter added to classical algorithm and is used in determination of the move of fireflies. Further we get traditional FA when Ø=1. Here we get optimized result. Its correct value in algorithm avoids getting stuck at local optima and finds global minimum solution.

4) Hybrid BFO-PSO. W. M. Korani[10] proposed a modified BFO called as hybrid BFO-PSO which collaborate the techniques of PSO and BFO.

BFO is an inspired by the bacteria's searching food behavior present in human body. The solution towards optimality is achieved by processes such as Chemotaxis (cost is calculated), and other three.

The hybrid approach of BFO and PSO, in which PSO exchange the social information and BFO finds solution by dispersal events. The best position which is global and local for each bacterium will finalize the direction. The tumble direction is updated as:

$$\emptyset(j+1) = V.\emptyset(j) + C_1.R_1.(P_{Lbest} - P_{current}) + C_2.R_2.(P_{Gbest} - P_{current}) \quad (12)$$

where P_{Lbest} and P_{Gbest} indicate the local best poistion and global best position.

It is found that the BFO-PSO gets trapped in local minima so that global optimal minima or solution cannot be achieved, which is solved using the hybrid IFA-PSO for global optimization.

2.1.5 KFCM-IFA-PSO. KFCM-IFA-PSO consists of 2 stages: (1) Finding the Initial Objective Function by KFCM and (2) Finding the optimised value of it using hybrid IFA-PSO.

The proposed hybrid IFA-PSO modifies the IFA with respect to PSO. This avoids the problem of trapped in local minima and find global minimum solution. The hybrid IFA-PSO have some difference in position vector. In this hybrid approach[8], Cartesian distance/metric is calculated using Eq (13) and Eq (14) which is the between the position and local/global solution. The Cartesian metric of x and pbest is represented as:

$$r_{px} = \sqrt{\sum_{k=1}^{d}(pbest_{i,j} - x_{i,j})^2} \quad (13)$$

Similarly the Cartesian metric of x and gbest is represented as:

$$r_{gx} = \sqrt{\sum_{k=1}^{d}(gbest_{i,j} - x_{i,j})^2} \quad (14)$$

The x in IFA-PSO which indicate value of position has been arbitrarily evolved from tumble direction equation.

$$x_i(t+1) = w.x_i(t) + c_1 e^{-r_{px}^2}\left(pbest_i - x_i(t)\right) + c_2 e^{-r_{gx}^2}\left(gbest_i - x_i(t)\right) + \alpha(\gamma - 0.5) \quad (15)$$

The algorithm in detail is:

Step 1: Initialize KFCM with initial random centers, m, σ. Initialize IFA-PSO with population size, inertia weight, Ø, and the maximum no cycles to be run.

Step 2: Applying KFCM to find the initial OF with Eq (1).

Step 3: Comparison of value of OF. If minimal go to Step 6.

Step 4: Calculate local and global values of particle after evaluating values of fitness.

Step 5: Fireflies position is updated using Eq (15) till OF converges.

Step 6: KFCM gives optimised OF, center and membership matrix.

Step 7: Evaluate Accuracy by confusion matrix and membership matrix.

The flowchart of proposed algorithm (KFCM-IFA-PSO) is shown in Fig 1 in detail. The initial input from KFCM goes to hybrid IFA-PSO which gives optimized result for center. This acts as input to KFCM which gives final optimized Object function. Datasets are numerical, and assumption is made that there are no missing values.

B. Datasets.

In this study, numerical experiments were conducted on six datasets(4 binary and 2 multiclass), diabetes dataset, liver disorder dataset, heart disease dataset, bank loan dataset, iris dataset, and car dataset from UCI Machine Learning Repository [13].

The details of the datasets considered for experimental results are given in Table I.

TABLE I
DATASET DESCRIPTION

Dataset	# of Tuples	# of Attributes	Class Attribute
Heart	270	13	{1,2}
Liver	345	6	{1,2}
Diabetes	11,51	19	{1,2}
Bank Loan	45,211	16	{1,2}
Iris	150	4	{1,2,3}
Car	1,728	6	{1,2,3,4}

There are no missing values in the datasets considered. However datasets with less than 5% of missing values can be considered.

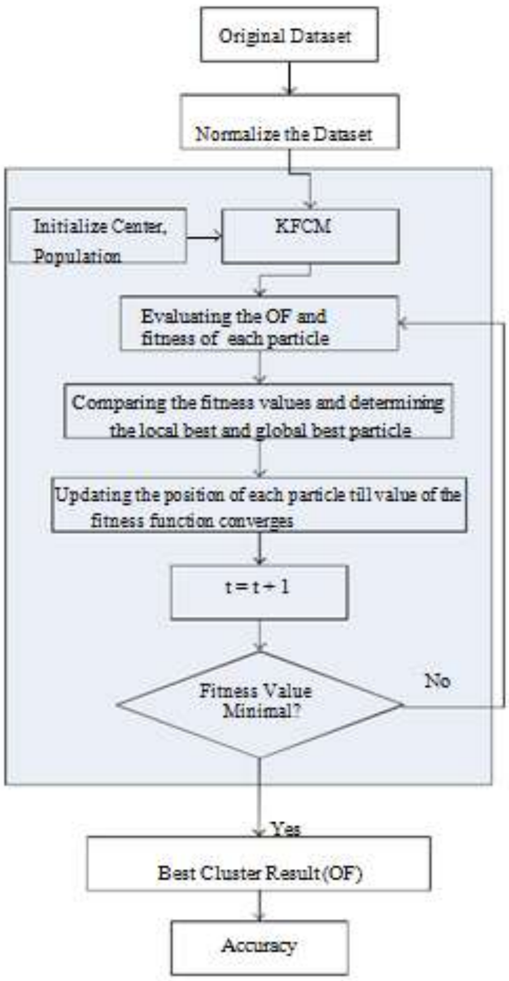

Figure 1: Flowchart for KFCM-IFA-PSO

III. RESULTS AND ANALYSIS

The hybrid KFCM-IFA-PSO was programmed in Java in Netbeans 8.0 platform and run on machine with core i3 processor, 2.20 Ghz, 6GB RAM.

The performance is compared between existing KFCM-BFO-PSO and proposed KFCM-IFA-PSO by running models for several iterations. The six datasets used are normalized to numerical values by manual substitution. Accuracy is calculated by the confusion matrix obtained from membership matrix.

KFCM-BFO-PSO is used for evaluating the performance of KFCM-IFA-PSO. The results of object function are shown in Table II. The new algorithm gave improved results in local as well global optima without trapping. Further results of object function obtained with FCM and KFCM separately were not desirable.

The results were obtained by several runs and finding maximum accuracy. This gave optimized values for object function. The table shows that the accuracy is overall good in proposed KFCM with Hybrid IFA-PSO approach than the existing KFCM with Hybrid BFO-PSO except in multi class dataset of car. However taking the maximum accuracy from runs the object function was minimal in all cases in our proposed KFCM-IFA-PSO.

The results indicate that value of fuzziness component m in KFCM differs for each dataset along with the width of kernel (σ). The width is also represented in terms of gamma(γ). The relationship is denoted by:

$$\gamma = \frac{1}{2\sigma^2} \qquad (16)$$

The values of m and corresponding rho(σ) in KFCM for each dataset are shown in Table III.

The parameters chosen for hybrid BFO-PSO and hybrid IFA-PSO are:
Parameters for Hybrid BFO-PSO:
Inertia (w)=0.9, c1=c2=0.49, r1=0.5, r2=0.7,
limit =0.1, Iterations=5

Parameters for Hybrid IFA-PSO:
w=0.9, c1=c2=0.49, limit=0.1, Iterations=5, α=0.7, Ø=0.49

Both models were run for various iterations in range of [1,20] and limit in range of [0.1,0.01,0.001] but optimal value of Object function and high accuracy was obtained with above parameters.

Accuracy for both the approaches was calculated by following equation:

$$Accuracy\ (\%) = \frac{(TP+TN)}{total\ \ tuples} \times 100 \qquad (17)$$

where TP=summation of all cases that are true positive and TN=summation of all cases that are true negative.

IV. CONCLUSIONS

The proposed method KFCM with hybrid IFA-PSO is able to cluster the datasets and results indicated that it has better performance than KFCM with hybrid BFO-PSO. It also performed well than the results obtained in only FCM and KFCM.

The objective was to minimize the object function without compromising the accuracy. The object function obtained was minimum in about 100% cases in KFCM with hybrid IFA-PSO as compared with KFCM with hybrid BFO-PSO. However here the accuracy as compromised.

ACKNOWLEDGMENT

The authors are highly obliged to Dean, Vice Chancellor and faculties of Research Development & Innovation Council at C U Shah University, Wadhwan, India.

TABLE II
COMPARISON OF RESULTS

Dataset description	**Existing Approach** (KFCM with Hybrid BFO-PSO)		**Proposed Approach** (KFCM with Hybrid IFA-PSO)		**Improvements (%) ***
Dataset	OF	Accuracy (%)	OF	Accuracy (%)	OF (%) **
Heart	521	66	**402**	69	**23**
Liver	610	49	**427**	50	**30**
Diabetes	5,573	56	**3,700**	57	**34**
Bank Loan	8,78,118	79	**1,03,359**	80	**88**
Iris	0.40	89	**0.214**	90	**46**
Car	19.5	35	**9.54**	35	**51**

*without compromising with accuracy, the proposed approach KFCM-IFA-PSO has better results of OF.
** Indicates the percentage change of OF in proposed and existing approach. Low value of OF is desirable.

TABLE II
VALUE OF M AND RHO FOR VARIOUS DATASETS

Class	**Dataset**	**m**	**σ**
Binary	Heart	2.0	600
	Liver	1.8	600
	Diabetes	1.8	600
	Bank Loan	1.2	0.1
Multi Class	Iris	2.0	600
	Car	2.0	600

REFERENCES

[1] Bezdek, J.C.: Pattern recognition with fuzzy objective function algorithms, pp. 95–107. Plenum Press, New York (1981).

[2] C. Pizzuti and D. Talia, ‘‘P-AutoClass: scalable parallel clustering for mining large data sets’’, in IEEE transaction on Knowledge and data engineering, Vol. 15, pp. 629-641, May 2003.

[3] Chen, S. C. and D. Q. Zhang (2004), "Robust Image Segmentation Using FCM With Spatial Constraints Based on New Kernel-Induced Distance Measure." IEEE transactions on systems, man, and cybernetics—part b: Cybernetics 34(4): 1907-1916.

[4] D. Graves and W. Pedrycz, "Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study", Fuzzy Sets and Systems, pp.522-543, 2010.

[5] Girolami, M. (2002), "Mercer kernel-based clustering in feature space." Neural Networks, IEEE Transactions on 13(3): 780-784.

[6] J. Marr, ‘‘Comparison Of Several Clustering Algorithms for Data Rate Compression of LPC Parameters’’, in IEEE International Conference on Acoustics Speech, and Signal Processing, Vol. 6, pp. 964-966, January 2003.

[7] K. C. Wong and G. C. L. Li, “Simultaneous Pattern and Data Clustering for Pattern Cluster Analysis’’, in IEEE Transaction on Knowledge and Data Engineering, Vol. 20, pp. 911-923, Los Angeles, USA, June 2008.

[8] P. Kora and S. R. Kalva, “Hybrid Bacterial Foraging and Particle Swarm Optimization for detecting Bundle Branch Block,” SpringerPlus, vol. 4, no. 1, Sep. 2015.

[9] S.-E. K. Fateen and A. Bonilla-Petriciolet, “Intelligent Firefly Algorithm for Global Optimization,” Studies in Computational Intelligence, pp. 315–330, Nov. 2013.

[10] W. M. Korani, “Bacterial foraging oriented by particle swarm optimization strategy for PID tuning,” Proceedings of the 2008 GECCO conference companion on Genetic and evolutionary computation - GECCO ’08, 2008

[11] X. L. Yang, Q. Song and W. B. Zhang, ‘‘Kernel-based

[12] Deterministic Annealing Algorithm For Data Clustering’’, in IEEE Proceedings on Vision, Image and Signal Processing, Vol. 153, pp. 557-568, March 2007.

[13] X.-S. Yang, “Firefly Algorithms,” Nature-Inspired Optimization Algorithms, pp. 111–127, 2014.

[14] UCI Machine Learning Repository, http://archive.ics.uci.edu/ml.