Journal of Computer Science 8 (2): 188-194, 2012
ISSN 1549-3636

A New Cooperative Algorithm Based on PSO and K-Means for Data Clustering

Mehdi Neshat, Shima Farshchian Yazdi,
Daneyal Yazdani and Mehdi Sargolzaei
Department of Computer Engineering,
Shirvan branch, Islamic Azad University, Shirvan, Iran

Abstract: Problem statement: Data clustering has been applied in multiple fields such as machine learning, data mining, wireless sensor networks and pattern recognition. One of the most famous clustering approaches is K-means which effectively has been used in many clustering problems, but this algorithm has some drawbacks such as local optimal convergence and sensitivity to initial points. **Approach:** Particle Swarm Optimization (PSO) algorithm is one of the swarm intelligence algorithms, which is applied in determining the optimal cluster centers. In this study, a cooperative algorithm based on PSO and k-means is presented. **Result:** The proposed algorithm utilizes both global search ability of PSO and local search ability of k-means. The proposed algorithm and also PSO, PSO with Contraction Factor (CF-PSO), k-means algorithms and KPSO hybrid algorithm have been used for clustering six datasets and their efficiencies are compared with each other. **Conclusion:** Experimental results show that the proposed algorithm has an acceptable efficiency and robustness.

Key words: Particle Swarm Optimization (PSO), Contraction Factor (CF-PSO), Sum of Intra cluster Distances (SISD), difference between Gbest fitness, local optimum, clustering algorithm

INTRODUCTION

Clustering is an unsupervised classification technique in which datasets that are often vectors in multi dimensional space, based on a similarity criterion, are divided into some clusters. Data clustering has vast application in data categorization (Memarsadeghi and Leary, 2003),(Velmuruqan and Santhanam, 2010), data compression (Celebi, 2011), data mining (Pizzuti and Talia, 2003), pattern recognition (Wong and Li, 2008), compacting (Marr, 2003), machine learning (Yang *et al*., 2007), image segmentation (Vannoorenberghe and Flouzat, 2006) and Data clustering importance in various sciences causes the introduction of various methods of data clustering (Hartigan, 1975). When used on a set of objects, which have attributes that characterize them, usually represented as vectors in a multi-dimensional space, are grouped into some clusters. When the predefined clusters number is K and there are N m-dimensional data, clustering algorithm would assign each of these data to one of the clusters, such that assigned data to a cluster with respect to a specific criterion are more similar to each other than data in other clusters.

The k-means clustering algorithm was developed by Hartigan (1975) which is one of the earliest and simplest clustering approaches that has been ever widely used. K-means method starts with K cluster centers and divides a set of objects into K subsets. This is one of the most famous and applied clustering techniques since it can be easily understood and implemented and its time complexity is linear. But k-means method has major weaknesses. One of these weaknesses is extra sensitivity to initial values of cluster centers. Objective function of k-means has multiple local optimums and k-means method is not capable to guarantee to pass local optimums. Therefore, if initial position of cluster centers in problem space was chosen inappropriately, this could converge to a local optimum.

Data clustering is of NP problems. One of the most applied methods for finding suitable solution for these kinds of NP problems belongs to swarm intelligence algorithms. Particle Swarm Optimization (PSO) is one of the most famous swarm intelligence algorithms, which was presented by Kennedy and Eberhart (1995). This algorithm is an effective technique for solving optimization problems that works based on probability rules and population. So far, different PSO-based methods for solving data clustering problem have been presented (Esmin *et al*., 2008; Kao and Lee, 2009; Tsai and Kao, 2010). Presented a hybridized algorithm based on k-means methods and PSO, called KPSO in (Merwe

Corresponding Author: Mehdi Neshat, Department of Computer Engineering, Shirvan Branch, Islamic Azad University, Shirvan, Iran

and Engelbrecht, 2003). In KPSO, first, k-means method is executed and then, outcome of k-means is used as one of the particles in initial solution of PSO. Therefore, first in this method, high convergence rate of k-means is used and after k-means converges, PSO is applied for exiting from local minimums and improving the result of k-means. In this study, a cooperative algorithm is proposed based on PSO and k-means. In the proposed algorithm, first, particles perform optimization process in PSO. After particle swarm convergence, obtained cluster centers by particles are used as initial cluster centers of k-means algorithm. After forwarding PSO's output to k-means, particles are reinitialized and performs clustering again. In fact, in the proposed algorithm, PSO is used for a global search and k-means is used for a local search. The proposed algorithm and also k-means, PSO, CF-PSO (Eberhart and Shi, 2000) and KPSO algorithms are applied for clustering 6 real datasets iris, glass, wine, sonar, pima and WDBC. Comparing obtained results from experiments shows an acceptable efficiency of the proposed algorithm.

MATERIALS AND METHODS

K-means algorithm: Clustering in D-dimensional Euclidean space is a process in which a set of N members, based on a similarity criterion, is divided into K groups or clusters. Various clustering methods are represented so far. The base of clustering algorithms is measuring the similarity between data and it is determined how much similar these two data vectors are, by a function. K-means algorithm is one of the oldest and most famous clustering methods. This method sorts data vectors in D-dimensional space in clusters, which their number was determined before, this clustering is based on Euclidean distance between data and cluster center which are considered as similarity criterion.

Euclidean distance between data vectors of a cluster with the center of that cluster is less than their Euclidean distance with other cluster centers. Standard k-means algorithm is as below:

˜Initial positions of K cluster centers are determined randomly. Following phases are repeated:

For each data vector: the vector is allocated to a cluster which its Euclidean distance from its center is less than the other cluster centers. The distance to cluster center is calculated by Eq. 1:

$$\mathrm{Dis}\left(X_p.Z_j\right) = \sqrt{\sum_{i=1}^{D}\left(X_{pi} - Z_{ji}\right)^2} \qquad (1)$$

In Eq. (1), X_p is pth data vector, Z_j is jth cluster center and D is the dimension of data and cluster center.

b) Cluster center are updated by Eq. 2:

$$Z_j = \frac{1}{n_j}\left[\sum_{\forall Xp \in Cj} X_p\right] \qquad (2)$$

In Eq. 2, n_j is the number of data vectors corresponding to jth cluster and C_j is a subset of the total data vectors which constitute jth cluster and are in it.

Phases (a) and (b) are repeated until stop criterion is satisfied (Hartigan, 1975).

Particle swarm optimization algorithm: PSO is one of the swarm intelligence methods and evolutionary optimization techniques, which was proposed by Kennedy and Eberhart (1995). PSO was presented according to animals social interactions such as bird folk and fish swarm. In this method, there is a swarm of particles that each of particles shows a feasible solution for optimization problem. Every particle tries to move toward final solution by adjusting its path and moving toward the best personal experience and also the best swarm experience.

Suppose that the population size is N. For particle i ($1 \le i \le N$) in D-dimension space, current position is $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ and velocity is $x_i = (v_{i1}, v_{i2}, \dots, v_{iD})$. During optimization process, velocity and position of each particle at each step is updated by Eq.3 and 4:

$$\begin{aligned} v_{i,j}(t+1) &= wv_{i,j}(t) + c_1R^1_{i,j}\left(\mathrm{Pbest}_{i,j}(t) - x_{i,j}(t)\right) \\ &+ c_2R^2_{i,j}\left(\mathrm{Gbest}_j(t) - x_{i,j}(t)\right) \end{aligned} \qquad (3)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \qquad (4)$$

where, $x_{i,j}$ is the component j of particle i, c_1 and c_2 are acceleration coefficients and w is inertia weight that can be a constant number or a positive function (Shi and Eberhart, 1998). R is a random number with uniform distribution in interval [0, 1]. Pbest_i (t) is the best position that is found by particle i until time t (the best individual experience of particle i) and Gbest (t) is the best position that until time t is found by whole swarm's members (the best swarm experience). At each iteration, the best individual experience of particle i is given by Eq. 5:

$$\mathrm{Pbest}_i(t+1) = \begin{cases} \mathrm{Pbest}_i(t), & \text{if } f\left(x_i(t+1)\right) \ge f\left(\mathrm{Pbest}_i(t)\right) \\ x_i(t+1), & \text{if } f\left(x_i(t+1)\right) < f\left(\mathrm{Pbest}_i(t)\right) \end{cases} \qquad (5)$$

where, f(x) is the fitness value of vector x. The best swarm experience is given by Eq. 6:

$$\text{Gbest}(t+1)=\arg\min_{P_i} f\left(\text{Pbest}_i(t+1)\right), 1\le i\le N \qquad (6)$$

Clerc presented another version of PSO in which by using construction factor (CF-PSO), PSO convergence rate has been improved. In this version of PSO, particles velocity is updated by Eq. 7:

$$v_{i,j}(t+1)=\chi\left(\begin{array}{c} v_{i,j}(t)+c_1R^1_{i,j}\left(\text{Pbest}_{i,j}(t)-x_{i,j}(t)\right) \\ +c_2R^2_{i,j}\left(\text{Gbest}_j(t)-x_{i,j}(t)\right)\end{array}\right) \qquad (7)$$

Eberhart showed that the appropriate value of χ is 0.729843788 and $c_1=c_2=2.05$ (Eberhart and Shi, 2000). According to how particles move in PSO, particles may leave search space, which leads to decrease efficiency and algorithm convergence rate. To remove this problem, some constraints are considered for velocity components' values. For this reason, in each of iterations, after computing velocity by Eq. 3, all of its components' values would be considered in various dimensions. The value of each velocity vector component can be clamped to the range $[-V_{max},V_{max}]$ to reduce the likelihood of particles leaving the search space. The value of V_{max} is usually chosen to be $K\times X_{max}$ (Here, X_{max} is the length of changes interval in search space dimensions), with $0.1\le K\le 1$ (Bergh and Engelbrecht, 2004). To find the optimal cluster centers, PSO algorithm applies Eq. 8 as the fitness function (Tsai and Kao, 2010). Eq. 8 shows generating function of Sum of Intra Cluster Distances (SICD) which is one of the most known evaluating criteria for clustering data. Less value of SICD is higher quality the clustering is performed. Therefore, for data clustering, PSO algorithm should minimize the fitness function in Eq. 8:

$$J(C_1,C_2,\dots,C_K)=\sum_{i=1}^{K}\left(\sum_{X_j\in C_i}\left\| Z_i-X_j\right\|\right) \qquad (8)$$

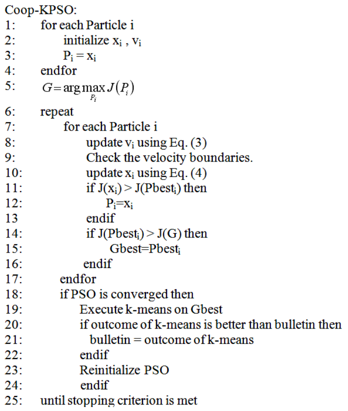
In Eq. (8), the Euclidean distance between each data vector in a cluster and the centroid of that cluster is calculated and summed up. Here, we have K clusters C_i $(1\le i\le K)$ that each of N data vectors X_j) are clustered on the basis of distance from each of these cluster centers Z_i $(1\le i\le K)$. Data vectors belong to a cluster that their Euclidean distance from its cluster center is less than their Euclidean distance from other cluster centers. Thus, PSO's objective is to determine cluster centers that are minimizing Eq. 8. Since data vectors and cluster center vectors are d-dimensional and there are K clusters, eachparticle should represent K cluster centers in d-dimensional space, consequently it has K×d �FQUOTE k×d components in its vector. Fig. 1 shows a vector of a particle that contains K d-dimensional cluster centers.

Proposed algorithm: In this section, a new cooperative algorithm based on PSO and k-means algorithms is described. The purpose of designing the proposed algorithm is to take advantages of both algorithms and remove their weaknesses. K-means is of high convergence rate, but it's very sensitive to initializing the cluster centers and in the case of selecting inappropriate initial cluster centers, it could converge to a local optimum. PSO can pass local optima to some extent but cannot guarantee reaching to global optima. However, PSO's computational complexity for data clustering is much more than k-means. How the proposed algorithm functions remove weaknesses of these two algorithms and apply their advantages is as following:

In the proposed algorithm, first, the particles are initialized in PSO. Each of particles contains K cluster centers which are displaced in the problem space by performing PSO algorithm. PSO continues to perform until the particles converge. After convergence of PSO, Gbest position including the best cluster centers which have found by particles so far is considered as the input of k-means. Then, k-means algorithm starts working and while it is not converged, it continues working. Therefore, PSO searches globally and as far as it can, it passes local optima. After convergence of PSO's particles, PSO's output would have an appropriate initial cluster centers for k-means.

$$\boxed{\left(Z_{1,1},Z_{1,2},\dots,Z_{1,d},Z_{2,1},Z_{2,2},\dots,Z_{2,d},\dots,Z_{K,1},Z_{K,2},\dots,Z_{K,d}\right)}$$

Fig. 1: Structure of a particle position in data clustering problem space



```
Coop-KPSO:
for each Particle i
    initialize $x_i$ , $v_i$
    $P_i = x_i$
endfor
$G=\arg\max_{P_i} J(P_i)$
repeat
    for each Particle i
        update $v_i$ using Eq. (3)
        Check the velocity boundaries.
        update $x_i$ using Eq. (4)
        if $J(x_i) > J(\text{Pbest}_i)$ then
            $P_i=x_i$
        endif
        if $J(\text{Pbest}_i) > J(G)$ then
            $\text{Gbest}=\text{Pbest}_i$
        endif
    endfor
    if PSO is converged then
        Execute k-means on Gbest
        if outcome of k-means is better than bulletin then
            bulletin = outcome of k-means
        endif
        Reinitialize PSO
    endif
until stopping criterion is met
```

Fig. 2: Pseudo code of proposed algorithm

Hence, after sending PSO' outcome to k-means, this algorithm starts searching locally. Consequently, in the proposed algorithm, global search ability of PSO has been used and after converging, a great part of optimization process will be given to k-means to utilize high capability of local search of this algorithm and its high convergence rate. Since initial cluster centers for k-means are obtained by PSO and k-means is used for local search, k-means weakness of sensitivity to initial cluster centers is removed. But, PSO capability may not be enough for preventing from being trapped in local optima. If this algorithm is trapped in local optima, it cannot present proper initial cluster values to k-means. Thereafter, according to low ability of k-means in passing local optima, the obtained result cannot be acceptable. To raise this problem, after convergence of PSO, the output of this algorithm is sent to k-means. Simultaneously with starting of k-means, PSO's particles are initialized and start global search again. In fact, in one time of executing the proposed algorithm, PSO has many times of chance to perform an acceptable global search. It should be noted that in the proposed algorithm, in each time of executing PSO, particles just search globally and converge after a short time and k-means undertakes the remaining of optimization process which is local search. Therefore, with respect to low computational complexity of k-means, huge amount of computations for local search is prevented. In the proposed algorithm, it has been tried to utilize this conserved computation load for giving new opportunities to PSO in order to perform an acceptable global search in at least one of given opportunities to it. Hence, for each execution of global search by PSO, k-means is also performed once. In the proposed algorithm, to determine the convergence of particle swarm, the difference of obtained results in consecutive iterations of performing the algorithm is used. When particles converge, the obtained results difference in consecutive iterations decreases, so by considering a threshold for the difference between Gbest fitness values in iterations i and j, it can determine their convergence. In the proposed algorithm, because PSO and k-means algorithms are performed multiple times, always, it has to save the best found cluster centers by algorithm so far. For this purpose, a bulletin is applied that each time k-means finishes after convergence of PSO, the obtained result of that will be compared with saved result in bulletin. If obtained cluster centers are better than saved result in bulletin, saved value in bulletin is updated. K-means execution finishes when after two consecutive iterations of its execution, cluster centers wouldn't be displaced. Pseudo code of the proposed algorithm is represented in Fig. 2.

RESULTS

Experiments: Experiments were performed on 6 datasets and efficiency of k-means, PSO, hybridized algorithm of PSO and k-means called KPSO (Merwe and Engelbrecht, 2003) and Proposed method were compared on these datasets. In all the methods, the objective function is Eq. 8, which calculates sum of intra cluster distances. In this study 6 datasets were used that all of them were selected among the standard real dataset of UCI (http://archive.ics.uci.edu/ml/) which include datasets of Iris, WDBC, Sonar, Glass, Wine and Pima. Brief specifications of datasets including name, size, number of attributes, number of classes and number of available data in every class are given in Table 1.

The performance of the six algorithms is evaluated and compared using the following criteria:

Sum of Intra cluster Distances (SISD): The distance between each data vector within a cluster and the cluster center of that cluster is calculated and summed up. Eq. 8 is used for calculating the SICD which has to be minimized.

Error rate: It is defined as the number of misplaced points over the total number of the points in the dataset which is given by Eq. 9:

$$\mathrm{Err}=\left(\frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\left(\mathrm{if}\ \left(\mathrm{Class}(\mathrm{i})=\mathrm{Cluster}(\mathrm{i})\right)\mathrm{then}\,0\,\mathrm{else}\,1\right)\right)\times 100 \qquad (9)$$

N is the total number of the points in dataset and Class(i) represents the class number which point i belongs to and Cluster (i) represents the cluster number which point i was assigned to. Eq. 9 shows the number of misplaced points divided by the total number of points.

In the proposed algorithm, since it is convenient that PSO performs fast global search and converges, it should use a version of PSO which is of high convergence rate. Among PSO's versions, PSO with contraction coefficient (Shi and Eberhart, 1998) has higher convergence rate, therefore, this version is applied in the proposed algorithm. Parameters adjustment in algorithms is as following:

Table 1: Characteristics of the Data Sets Considered

Name	No. of classes	No. of attribute	(size of classes in parentheses) size
Iris	3	4	150 (50, 50, 50)
Pima	2	8	768 (500, 268)
Wine	3	13	178 (48, 71, 59)
Glass	6	9	214 (70, 17, 76, 13, 9, 29)
Sonar	2	60	208 (111, 97)
WDBC	2	30	569 (357, 212)

Initial positions of cluster centers of all algorithms are selected among data vectors randomly. Population size in PSO, PSO with contraction coefficient, KPSO and the proposed algorithm is considered 5 times of problem space dimensions according to (Kao *et al*., 2008). In the proposed algorithm and PSO with contraction coefficient, $c_1 = c_2 = 2.05$ and $\chi = 0.729843788$ are considered. $c_1 = c_2 = 2$ are considered in PSO and KPSO and inertia weight value is obtained by "w = 0.5 + rand/2" at each iteration (Kao *et al*., 2008). With respect to various experiments, if SICD relating to Gbest is less than 0.1 in 5 iterations, it means that particle swarm has converged. In (Kao *et al*., 2008) the number of iterations of algorithms execution based on PSO is equal to 10 times of the problem space dimensions.

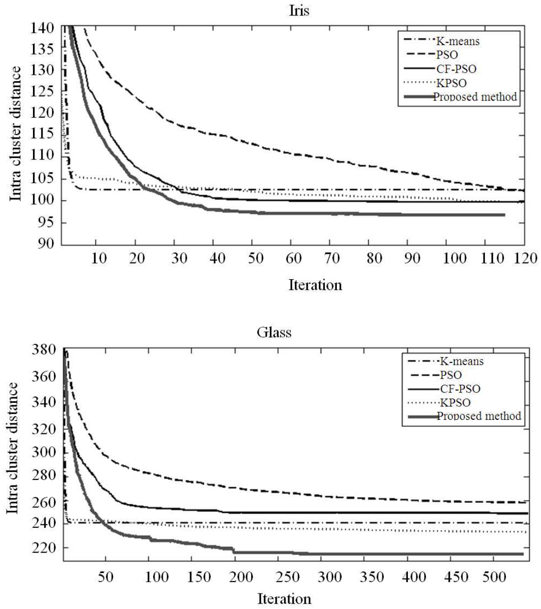


Fig. 3: Convergence behaviors of the five algorithms for Iris and Glass dataset

For instance, for clustering Iris dataset which has 3 data classes of four dimensions, the problem space would be 12 dimensional as Fig. 1 (D = 12). Therefore, the algorithms have to be run 120 times. One of the other stop criteria for performing the algorithms is based on the number of fitness evaluations. In this study, for fairness of comparisons, the number of times which PSO executes fitness evaluations in 10×D iterations of its performance, i.e. SICD is calculated, is considered as stop criterion of algorithms. Hereon, all the algorithms can do the same number of fitness evaluations until finishing their work. The algorithms are performed 50 times for data clustering and the best, mean and standard deviation of SICD obtained from algorithms are presented in Table 2 for clustering of 6 datasets.

Fig. 3; illustrate the convergence behaviors of the five algorithms for iris and glass dataset.
As it is observed in Fig. 3, the proposed algorithm is executed in less iteration in comparison with other algorithms because during execution of the proposed algorithm, these iterations are used for k-means execution. In fact, each iteration of PSO execution with N particles is equal to N iterations of k-means execution based on fitness evaluation number.

DISCUSSION

As results show in Table 2, the proposed algorithm is of better efficiency than other tested algorithms. Nevertheless, it has to be mentioned that the best obtained result from CF-PSO is better than the proposed algorithm in three cases because this algorithm is of greater local search ability than k-means. Indeed, although CF-PSO has higher computational load than k-means for local search, the accuracy of its obtained results is better than k-means. As a result, the best obtained results from it are sometimes better than the proposed algorithm because in the proposed algorithm local search is done by k-means.

Table 2: Comparison of sum of intra cluster distances for the five clustering algorithms

Data set	Criteria	K-means	PSO	CF-PSO	KPSO	Proposed method
Irish	Best	97.320000	97.1000000	96.6500000	96.7800000	96.7300000
	Mean	102.5700000	102.2600000	99.7600000	99.6100000	96.9100000
	Std. Dev	11.3400000	5.8100000	10.9300000	7.2100000	0.1700000
Pima	Best	52072.2400000	47627.7300000	47562.1200000	47606.0200000	47561.3500000
	Mean	55076.5200000	48153.1800000	47613.0400000	47876.1200000	47580.4300000
	Std. Dev	7790.3700000	523.1500000	175.4900000	284.1200000	59.9700000
Wine	Best	16555.6800000	16307.1600000	16294.1800000	16298.9200000	16292.6800000
	Mean	17662.7300000	16320.6700000	16303.8800000	16307.5800000	16293.7500000
	Std. Dev	1878.0700000	9.5300000	5.6000000	7.2300000	0.8800000
Glass	Best	213.4200000	230.6400000	210.4400000	212.0300000	211.0400000
	Mean	241.0300000	258.0200000	248.8000000	233.2800000	214.8300000
	Std. Dev	25.3200000	12.2400000	18.7800000	14.0500000	4.0900000
WDBC	Best	152647.2500000	149537.7300000	149473.8500000	149480.9300000	149473.1500000
	Mean	179794.2500000	149830.8700000	149659.0400000	149594.0500000	149475.0800000
	Std. Dev	55222.1700000	364.7300000	511.6100000	198.3100000	1.3700000
Sonar	Best	234.7700000	271.8300000	233.7500000	234.6500000	234.6300000
	Mean	235.0600000	276.6800000	245.0200000	234.9200000	234.7800000
	Std. Dev	0.1500000	3.7900000	15.6200000	0.2200000	0.1600000

Table 3: Comparison of error rate for the five clustering algorithms

Data Set	K-means	PSO	CF-PSO	KPSO	Proposed method
Irish	16.05±10.10	10.64±4.50000	12.57±6.730000	12.58±7.67000	10.24 ± 0.5900
Pima	34.10±0.310	39.40±0.83000	39.24±0.260000	39.49±0.82000	39.19±0.02000
Wine	34.38±6.080	28.74±0.39000	28.53±0.370000	28.59±0.47000	28.46±0.30900
Glass	48.30±3.140	48.72±1.34000	47.28±1.570000	47.80±1.98000	45.82±0.81000
WDBC	19.12±9.220	13.18±1.80e-15	13.18±1.80e-15	13.18±1.81e-15	13.18±1.81e-15
Sonar	44.95±0.970	46.60±0.420000	47.03±1.000000	44.89±0.840000	44.00±1.470000

While the averages of results of the proposed algorithm, in all cases, are better than the other tested algorithms in this study. The reason is usage of strategies which have been used for global search in this algorithm. In fact, the proposed algorithm is successful in finding the global optima in most runs and can prevent final result from being trapped in local optima, whereas, this ability is observed less in other algorithms and they cannot guarantee passing local optima. This weakness causes that other algorithms to be of less strength and not to be able to reach to almost the same results in their various implementations. For instance, CF-PSO algorithm that has gotten better results in some cases couldn't reach to these results in various runs again. This weakness in k-means is observed more than other algorithms and the reason is its sensitivity to initial centers positions and low capability of passing local optimums. In KPSO, also since PSO is performed after k-means, in some cases, it may be possible that the improper obtained result from k-means even causes PSO's particles to be trapped in that local optimum.

Therefore, KPSO has less strength in clustering some datasets either. Obtained results from CF-PSO in all conditions is better than PSO with random inertia weight. It shows that applying CF-PSO instead of PSO with random inertia weight can have better results. Totally, results of Table 2 show that the proposed algorithm is of very high strength in comparison to other tested algorithms which lower values of standard deviation of the proposed algorithm's results confirm it.

Table 3 shows mean value of error rate with standard deviation of clustering of 5 algorithms on datasets of Table 1 over 50 runs. It is seen that, in 4 cases, mean value of error rate of proposed algorithm is less than the other algorithms and in Pima doesn't have the least error rate, while in this one case, mean value of SICD of proposed algorithm is better than the other algorithms'. This is there is no absolute correlation between the SICD and the error rate.

CONCLUSION

In this study, a new cooperative algorithm based on k-means and PSO is presented. In the proposed algorithm, PSO performs global search and k-means is responsible for local search. The process of the proposed algorithm is such that the strength and ability of preventing from being trapped in local optimums is improved. The proposed algorithm along with four other algorithms is used for clustering 6 standard datasets and obtained results are compared with each other. Experimental results show that the proposed algorithm is of higher robustness and better efficiency to other tested algorithms. To improve the obtained results of the proposed algorithm, it can increase local search ability around the best found position by the algorithm. This is issue that merits further research.

REFERENCES

Bergh, F.V.D. and A.P. Engelbrecht, 2004. A cooperative approach to particle swarm optimization. IEEE Trans. Evolutionary Comput., 8: 225-239. DOI: 10.1109/TEVC.2004.826069

Celebi, M.E. 2011. Improving the performance of k-means for color quantization. J. Image Vision Comput. 29: 260-271. DOI: 10.1016/j.imavis.2010.10.002

Eberhart, R.C. and Y. Shi, 2000. Comparing inertia weights and constriction factors in particle swarm optimization. Proceedings of the Congress Evolutionary Computation, July 16-19, IEEE Xplore Press, La Jolla, CA , USA., pp: 84-88. DOI: 10.1109/CEC.2000.870279

Esmin, A.A.A., D.L. Pereira and F. de Araujo, 2008. Study of different approach to clustering data by using the particle swarm optimization algorithm. Proceedings of the IEEE Congress on Evolutionary Computation, June 1-6, IEEE Xplore Press, Hong Kong, pp: 1817-1822. DOI: 10.1109/CEC.2008.4631035

Hartigan, J.A., 1975. Clustering Algorithms. 1st Edn., Wiley, New York, ISBN: 047135645X, pp: 351.

Kao, Y. and S.Y. Lee, 2009. Combining K-means and particle swarm optimization for dynamic data clustering problems. Proceedings of the IEEE International Conference on Intellgent Computing and Intelligent System, Nov. 20-22, IEEE Xplore Press, Shanghai, pp: 757-761.DOI: 10.1109/ICICISYS.2009.5358020

Kao, Y.T., E. Zahara and I.W. Kao, 2008. A hybridized approach to data clustering. Expert Syst. Appl., 34: 1754-1762. DOI: 10.1016/j.eswa.2007.01.028

Kennedy, J. and R. Eberhart, 1995. Particle swarm optimization. Proceedings of the IEEE International Conference on Neural Networks, Nov. 27-Dec. 1, IEEE Xplore Press, Perth, WA, Australia, pp: 1942-1948. DOI: 10.1109/ICNN.1995.488968

Marr, J., 2003. Comparison of several clustering algorithms for data rate compression of lpc parameters. Proceedings of the IEEE International Conference on Acoustics Speech Signal Processing, (ASSP'03), IEEE Xplore Press, Alabama, pp: 964-966. DOI: 10.1109/ICASSP.1981.1171119

Memarsadeghi, N. and D.P. Leary, 2003. Classified information: The data clustering problem. IEEE Tran. Comput. Sci. Eng., 5: 54-60. DOI: 10.1109/MCISE.2003.1225861

Merwe, D.W.V.D. and A.P. Engelbrecht, 2003. Data clustering using particle swarm optimization. Proceedings of the Congress on Evolutionary Computation, Dec. 8-12, IEEE Xplore Press, South Africa, pp: 215-220. DOI: 10.1109/CEC.2003.1299577

Pizzuti, C. and D. Talia, 2003. P-AutoClass: Scalable parallel clustering for mining large data sets. IEEE Trans. Knowledge Data Eng., 15: 629-641. DOI: 10.1109/TKDE.2003.1198395

Shi, Y. and R. Eberhart, 1998. A modified particle swarm optimizer. Proceedings of the IEEE International Conference Evolutionary Computation, May, 4-9, IEEE Xplore Press, Anchorage, AK , USA., pp: 69-73. DOI: 10.1109/ICEC.1998.699146

Tsai, C.Y. and I.W. Kao, 2010. Particle Swarm Optimization with Selective Particle Regeneration for Data Clustering. Expert Syst. Appl. 38: 6565-6576. DOI: 10.1016/j.eswa.2010.11.082

Velmuruqan, T. and T. Santhanam, 2010. Computational complexity between K-Means and K-medoids clustering algorithms for normal and uniform distributions of data points. J. Comput. Sci., 6: 363-368. DOI:10.3844/jcssp.2010.363.368

Vannoorenberghe, P. and G. Flouzat, 2006. A belief-based pixel labeling strategy for medical and satellite image segmentation. Proceedings of the IEEE International Conference on Fuzzy System, Juley 16-21, IEEE Xplore Press, Vancouver, BC., pp: 1093-1098. DOI: 10.1109/FUZZY.2006.1681846

Wong, A.K.C. and G.C.L. Li, 2008. Simultaneous pattern and data clustering for pattern cluster analysis. IEEE Trans. Knowledge Data Eng., 20: 911-923. DOI: 10.1109/TKDE.2008.38

Yang, X.L., Q. Song and W.B. Zhang, 2007. Kernel-based deterministic annealing algorithm for data clustering. IEEE Proc. Vision, Image Signal Proces., 153: 557-568. DOI: 10.1049/ip-vis:20050366