I.J. Intelligent Systems and Applications, **2018, 8, 36-53**
Published Online August 2018 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2018.08.04

Multi-Swarm Whale Optimization Algorithm for Data Clustering Problems using Multiple Cooperative Strategies

Ravi Kumar Saidala
Acharya Nagarjuna University, Nagarjuna Nagar, 522510, India
E-mail: saidalaravikumar@gmail.com

Nagaraju Devarakonda
LBR College of Engineering, Mylavaram, 521230, India
E-mail: dnagaraj_dnr@yahoo.co.in

Received: 28 July 2017; Accepted: 11 September 2017; Published: 08 August 2018

***Abstract*—**Computational Intelligence (CI) is an as of emerging area in addressing complex real world problems. The WOA has taken its root from the collective intelligent foraging behavior of humpback whales (Megaptera Novaeangliae). The standard WOA is suffers from the selection of best agent while whales searching and encircling prey. This research paper deals with the multi-swarm cooperative strategies for finding the best agents which balances the two phase's exploration and exploitation. The performance of invoked Multi-Swarm cooperative strategies into standard WOA i.e, MsWOA is first benchmarked on a set of 23 standard mathematical benchmark function problems which includes 7 Uni-Modal, 6 Multi-modal and 10 fixed dimension multi-modal functions. The obtained graphical and statistical results have been portrayed along with the previously established techniques. The obtained results depicts that the proposed cooperative strategies for WOA outperforms in solving optimization problems of standard benchmark functions. This paper also studies the numerical efficiency of proposed techniques on the problem of data clustering where 10 real data clustering problems have been taken from data repository https://archive.ics.uci.edu.data. Statistical results for the obtained cluster centroids, intra-cluster distances and inter-cluster distances confirms that the cooperative strategies for best whale agent selection improves the performance WOA for function optimization problems as well as data clustering problems.

***Index Terms*—**Nature Inspired Algorithms, Meta-heuristic optimization techniques, WOA, Cooperative Strategies, Data Cluster Problems.

I. Introduction

Computational Intelligence is the fast growing area in tending to this real world and complex problems. The majority of the computational Intelligence algorithms are inspired by nature [1]. Researchers now conceded the way that the most ideal approach to discover answers for this real world and complex problems by studying the nature, this unveiling the micro secrets in nature to produce new optimization algorithms. Population based meta-heuristic algorithms are ending up noticeably more well known in solving complex real world problems with use of no single gradient function. Modern problems are for the most part with various complex variables that are should have been utilized as a part of finding optimal solutions. The last two decades have witnessed many real world complex problems are addressed by the field of meta-heuristics. The main objective of these algorithms is enabling the optimization process in giving the best possible solution in any situation where there are an extensive number of possible solutions.

In Data Mining, cluster analysis identifies the patterns which are recognized from each other through their features, by finding the partitions that have similar characteristics. Partitioning is a collection of objects into clusters all together that items in one cluster ought to have a minimization of intra-cluster similarity and the maximization of inter cluster likeness. Finding optimal clusters or optimal structures of data is a difficult task when it comes to complex data. Many classical algorithms have been developed for discovering clusters of data however finding optimal structures of data is as yet a big challenging task. There are two main problems in clustering. The first is choice of initial number of centroids and the second one is distance function optimization problem. Numerous data researchers have demonstrated that k-Means clustering technique is the absolute best answer for discovering hidden patterns by clustering the entire dataset. Clustering also called unsupervised learning in some traditional fields, for example, machine learning.

Authors [2-6] used nature inspired algorithms to find the optimized structures by applying them to cluster techniques. For instance, in *k*-Means clustering similarity function should be minimizing centroid distance to obtain optimized clusters. *i.e.* minimization of sum of squared Euclidean distance of objects from respective cluster means shown in Eq. (1).

$$d_{min} = \sum_{j=1}^{K} \sum_{Z_i \in C_i} \left\| Z_i - \mu_j \right\|^2 \quad (1)$$

where μ_j is the mean of C_j.

Let's consider an optimization problem. We have given an undertaking to distinguish the natural product sort. Each organic product has recognized three particular components they are contour, color, size. X is a dataset which has the above three components. Initial number of clusters and cluster centroids are chosen. I now need to cluster up the natural products into clusters in which it definitely has a place with. The clustering component is said to be good only if clustering method acquires two conditions, one is intra-cluster minimization and the other is inter-cluster maximization. Numerous classical clustering algorithms have the optimization problem, failed to deal with non-linear dataset and insensitive to the noisy or anomaly data. We assumed that, we have an optimization function that measures how "good" clustering. By clustering up the organic products, and after that calling an optimization function on each group to measure its goodness, and afterward summing up the goodness of each group, at that point we can quantify how "good" a certain set of groups is. At the point when there is an increase in the aggregate number of natural product sort and components, at that point the problem has all the earmarks of being a NP-Complete, so it can be solved with agreeing to a parallel meta-heuristic algorithm. Nature inspired meta-heuristic algorithm can optimize the non-linear functions in an efficient manner.

Table 1.a. List of Uni and Multi modal benchmark functions (F_1 to F_{13}). where V_no = 30 , f_{min} = 0

F_Name	Objective Function	Range
Sphere function	$F_1(x) = \sum_{i=1}^{n} x_i^2$	[-100, 100]
Schwefel 2.22 function	$F_2(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$	[-10, 10]
Schwefel 1.2 function	$F_3(x) = \sum_{i=1}^{n} \left(\sum_{j=1}^{i} x_j\right)^2$	[-100, 100]
Rotated Schwefel 2.21 function	$F_4(x) = \max_i \{\lvert x_i \rvert, \ 1 \le i \le n\}$	[-100, 100]
Rosenbrock's function	$F_5(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - {x_i}^2)^2 + (x_i - 1)^2]$	[-30, 30]
	$F_6(x) = \sum_{i=1}^{n} ([x_i - 0.5])^2$	[-100, 100]
Quartic function	$F_7(x) = \sum_{i=1}^{n} i{x_i}^4 + random[0,1)$	[-1.28, 1.28]
Schwefel function	$F_8(x) = \sum_{i=1}^{n} -x_i^2 \sin(\sqrt{\lvert x_i \rvert})$ (Note: f_{min} =-418.9829*5)	[-500, 500]
Rotated Rastrigin's function	$F_9(x) = \sum_{i=1}^{n} [x_i^2 - 10\cos(2\pi \ x_i) + 10]$	[-5.12, 5.12]
Ackley function	$F_{10}(x) = -20e^{\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right)} - e^{\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right)} + 20 + e$	[-32, 32]
Griewank function	$F_{11}(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	[-600, 600]
Penalized 1 function	$F_{12}(x) = \frac{\pi}{n}\{10\sin(\pi y_1) + \sum_{i=1}^{n-1}(y_i - 1)^2[1 + 10\sin \pi {y_{i+1}}^2] + (y_n - 1)^2\} + \sum_{i-1}^{n} u(x_i, 10, 100, 4)$ $y_i = 1 + \frac{x_i + 1}{4} u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m & x_i > a \\ 0 & -a < x_i < a \\ k(-x_i - a)^m & x_i < -a \end{cases}$	[-50, 50]
Penalized 2 function	$F_{13}(x) = 0.1\{\sin(3\pi x_1)^2 + \sum_{i=1}^{n} (x_i - 1)^2[1 + \sin(3\pi x_i + 1)^2] + (x_n - 1)^2[1 + \sin 2\pi x_n)^2]\} + \sum_{i-1}^{n} u(x_i, 5, 100, 4)$	[-50, 50]

Problem statement the problem of optimized structures of a dataset for clustering is formalized as takes after. Let consider a dataset D_T, where T denotes the total number of data instances and N be the dataset dimensions (variables) that are considered. Let the input N-dimensional dataset. Consider a dataset D featured by P attributes:

$$D_T = [\{x_{1T}, x_{2T}, \ldots, x_{PT}\}]^T \quad (2)$$

Assume dataset D_T with N dimensions is partitioned into C_i clusters, where i = 1 to k, so that for each D_T Similarity measured by distance d then,

$$d\big(D_T | C_i(\{x_{1T}, x_{2T}, \ldots\ldots\ldots, x_{PT}\})\big) = max\left\{d\left(D_T | C_l(x_{1T}, x_{2T}, \ldots\ldots\ldots, x_{PT})\right)\right\} \quad (3)$$

d is the distance between D_T and C_i. C_i Should meet the following conditions:

$$C_i \neq \emptyset, C_i \cap C_s = \emptyset \; and \; \bigcup_{i=1}^{k} C_i = D_T \; where \; i,s = 1 \ldots k$$

We utilize the term compactness to quality of a given cluster in light of intra and inter cluster minimization, maximization similarities respectively. The clustering problem minimizing in a simple manner is portrayed as takes after.

$$d(D_T, Z) = \sum_{i=1}^{N} . \sum_{j=1}^{k} || D_{Tij} - Z_j ||^2 \quad (4)$$

where k denotes the number of clusters, N the number of dimensions, D_{Tij} is the location of i^{th} dimension of cluster j.

$$\frac{1}{N_j} \sum_{i=1}^{N} D_{Tij} \quad (5)$$

N_i is the number of dimensions in the j^{th} cluster.

Table 1.b. List of fixed-dimension multimodal benchmark functions (F_{14} to F_{23}).

F_Name	Objective Function	Range	V_no	f_{min}
De jong function 5	$F_{14}(x) = \left(\frac{1}{500}\sum_{j=1}^{25}\frac{1}{j+\sum_{i=1}^{2}(x_i-a_{ij})^6}\right)^{-1}$	[-65,65]	2	1
Kowalik function	$F_{15}(x) = \sum_{i=1}^{11}\left[a_i - \frac{x_1(b_i^2+b_ix_2)}{b_i^2+b_ix_3+x_4}\right]^2$	[-5, 5]	4	0.00030
Six-Hump Camel Back function	$F_{16}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$	[-5, 5]	2	1.0316
Branin RCOS function	$F_{17}(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1-\frac{1}{8\pi}\right)\cos x_1 + 10$	[-5, 5]	2	0.398
Goldstein–Price function	$F_{18}(x) = \left[1 + (x_1+x_2+1)^2\left(\begin{matrix}19 - 14x_1 + 3x_1^2 - 14x_2 + \\ 6x_1x_2\end{matrix}\right) + 3x_2^2\right] * [30 + (2x_1 - 3x_2)^2 * (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$	[-2, 2]	2	3
Hartmann function 3	$F_{19}(x) = -\sum_{i=1}^{4} c_i e^{(-\sum_{j=1}^{3} a_{ij}(x_j-p_{ij})^2)}$	[1, 3]	3	-3.86
Hartmann function 6	$F_{20}(x) = -\sum_{i=1}^{4} c_i e^{(-\sum_{j=1}^{6} a_{ij}(x_j-p_{ij})^2)}$	[0, 1]	6	-3.32
Shekel function 5	$F_{21}(x) = -\sum_{i=1}^{5}[(X-a_i)(X-a_i)^T + c_i]^{-1}$	[0, 10]	4	-10.1532
Shekel function 7	$F_{22}(x) = -\sum_{i=1}^{7}[(X-a_i)(X-a_i)^T + c_i]^{-1}$	[0, 10]	4	-10.4028
Shekel function 10	$F_{23}(x) = -\sum_{i=1}^{10}[(X-a_i)(X-a_i)^T + c_i]^{-1}$	[0, 10]	4	-10.5363

II. Related Work

The most recent two decades have seen an exceptional change in the area of computational intelligence for growing popular and efficient optimization techniques which are generally inspired by the nature; these are well known in tackling complex np-hard problems by exploring and mimicking different phenomena of the nature. The absolute most well-known optimization algorithms are comprehensively listed below.

Darwin's theory of evolution has inspired towards the development of GA (Genetic Algorithms) [7]. These algorithms are adaptive heuristic search algorithms intended to imitate the processes in natural system. The basic idea behind the development of GA is the evolution of new chromosomes from the combinations of initial chromosome populations making an expectation that recently evolved ones are better to the old set of chromosomes. The searching capability of GA is exploited, keeping in mind the end goal to search for proper cluster centers in the feature space to such an extent that a similarity metric of the resulting cluster is optimized.

The chromosomes that which are represented as strings of real numbers, encode the centers of a fixed number of clusters [8]. The ACO [9] is probabilistic based method and was planned in view of the natural phenomenon of real ant colonies. It is a populace based MHO calculation motivated by the ant behavior that is used to determine discrete- optimization problems. Many researchers [10], [11] used ACO as tool for finding k optimal clusters of N data objects. This algorithm has been devised and used on several artificial and real datasets. DE [12] is a population of candidate solutions based optimization technique has paved the path by storn and price and it is used to optimize real parameter, real valued functions. It is the one of the quick, robust and efficient global search heuristics of current intrigue. Many authors utilized DE for finding optimized clusters of large unlabeled datasets. As opposed to the vast majority of the current clustering techniques, the DE requires no prior knowledge of the data for finding optimal clusters [13], [14]. The behavioral aspect of many birds flocking and fish schooling patterns has prompted the development of PSO [15] by *Russell Eberhart* and *J. Kennedy*. The key idea has been created from a flock of birds where every individual in the flock determines its closest neighbor and replaces their velocity with that neighbor. Among the numerous nature-inspired algorithms, clustering using PSO technique has identified as robust and efficient in solving clustering problems [16]. PSO can be utilized to discover centroids of a user indicated number of clusters. It is suitable for clustering complex and linearly non-separable datasets [17]. Keeping in mind the end goal to enhance the efficiency of PSO technique, many researchers have proposed distinctive variations of the PSO algorithm furthermore, have grown new thoughts, for example, adaptive heuristics, different fitness functions, kernel-induced similarity measures, and evolution of swarm generations and so on [18-24]. Newton's law of gravity has given rise to an optimization algorithm named as GSA [25]. It is a stochastic population based MHO algorithm was developed based on the *gravity* and *mass* interactions. Variants of GSA also developed based on the concept of antigravity. In GSA randomly created candidate solutions for data clustering problem then the interaction have been made to each and other via Newton's gravitation law to search problem space [26], [27]. The WOA (Whale Optimization Algorithm) [28] has been devised based on the foraging strategy of hump back whales. The strategy of searching for prey, encircling strategy and the mass net searching towards getting the prey was considered in building up this algorithm. The performance comparisons have been made by testing on 23 standard mathematical benchmark functions. Found that WOA perform efficient than many state-of-the-art optimization algorithms including PSO [28]. Authors in [29] improved the optimization performance by introducing new convergence factor in both exploration and exploitation phase. Authors [30] proposed a new parallel meta-heuristic optimization algorithm (NPMOA) which was formulated by hybridizing WOA with CSA. The above variants of WOA (IWAO, CSAWOA) were tested and applied to data clustering problems [31], [32].

In this paper we use the multi-swarm cooperative strategies for enhancing the performance of standard WOA and applying it to data clustering problems. The paper focuses on ring, master-slave and hybrid cooperative strategies for finding the best agent vector $\vec{X}^*$. As earlier said this paper comprises two parts where in the first part we use each multi-swarm cooperative strategy and find the performance. In second part applying the swarm cooperative strategies based WOA for data clustering problems.

The structural organization of this paper is done as follows. The section III presents the Standard Whale Optimization Algorithm. In the Section IV, we have described the swarm cooperative strategies where in A, B and C we presented Ring, Master-Slave and Hybrid cooperative strategies respectively. The performance comparisons of swarm cooperative strategies based WOA with GSA, DE, PSO, WOA and variants of WOA have been made in Section V whereas Section A and B outlines the performance comparisons on 23 standard benchmark functions that are listed in Table 1. a) and b) and data clustering problems that are described in Table 6 respectively. The last section concludes the whole paper and points out the future scope.

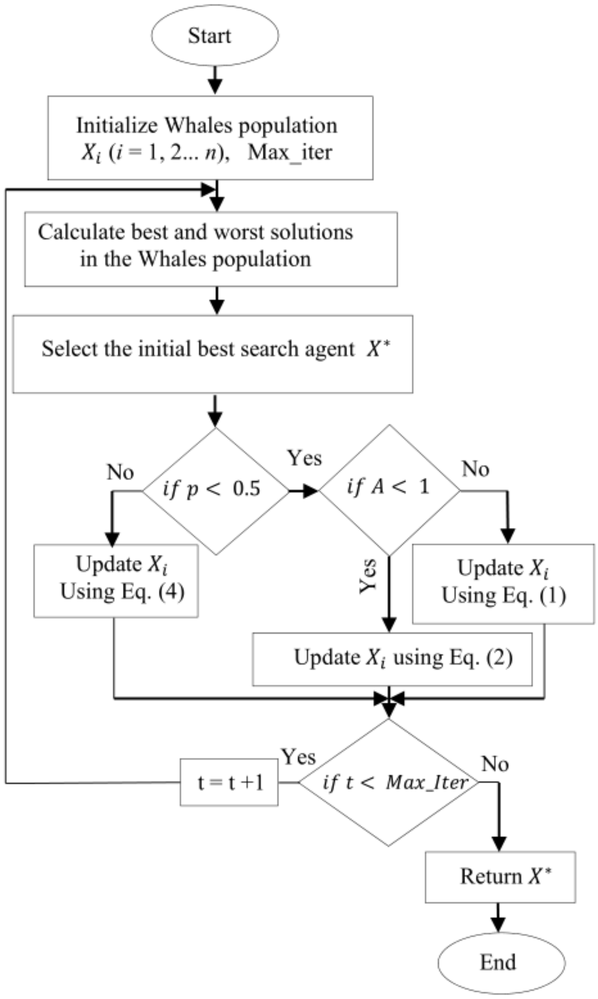


Fig.1. Flowchart of Standard Whale Optimization algorithm

III. Standard WOA

The Whale optimization Algorithm in short WAO was first coined by S. Mirjalili and A. Lewis. The special hunting behavior of humpback whales has paved the path for developing this algorithm [25]. The specialty of whales in hunting prey is they release bubbles in either 'circular' manner or '9' shaped path to move prey to the surface of the ocean. The mathematical model of this behavior is explained in two phases. The first phase is explained as hunting behavior. In this phase observed that whales show two maneuvers. They are shrinking encircling mechanism and Spiral updating position. Eq. (6) and (7) specifies the equations for the whales updating position in circular and spiral path respectively.

$$\vec{X}_{t+1} = \begin{cases} \overrightarrow{X^*}_t - \vec{A}.\vec{D}\,, & p < 0.5 \quad (6) \\ \overrightarrow{D'}.e^{bl}.cos(2\pi l) + \overrightarrow{X^*}_t & p \geq 0.5 \quad (7) \end{cases}$$

where $\overrightarrow{D'} = \left|\overrightarrow{X^*}_t - X_t^*\right|,\ p = (-1).\times rand(0,1)$. In Eq. (6) $\vec{a}$ is decreased to achieve to shrinking the encircling mechanism. Then $\vec{A} = (-2\text{a}).* \text{rand}(1,1) + \text{a}$. where, 'a' is an integer and coefficient $\vec{A}$ is decreased by $\vec{a}$.

In Eq. (2) constant b is used in defining the shape of the logarithmic spiral, $l = (-2).* rand(1,1) + 1$, and $'.'$ is a dot product operator. WOA assumes probability of 50% to choose between either these two approaches. If $p < 0.5$ whales updates their position in *shrinking encircling mechanism* else if $p \geq 0.5$, they move in *Spiral updating position.* The second phase is presented as searching phase. In this phase, WOA assumed that whales search for prey in two different possible ways. If $\overrightarrow{A} \geq 1$ then the updated position of whale using Eq. (8).

$$\vec{X}_{t+1} \leftarrow \vec{X}_{rand} - \vec{A}.\vec{D} \tag{8}$$

where, $\vec{D} = \left|\vec{A}.\vec{X}_{rand} - \vec{X}\right|$,

$\vec{X}_{rand}$ is a random position vector.

Else if $\vec{A} < 1$ new updated position becomes the Eq. (9).

$$\vec{X}_{t+1} \leftarrow \overrightarrow{X^{*}}_{t} - \vec{A}.\vec{D} \tag{9}$$

where, $\vec{D} = \left|\vec{C}.\overrightarrow{X^{*}}_{t} - \overrightarrow{X_t}\right| \vec{A}$ and $\vec{C}$ are internal parameters, the subscript $'t'$ specifies the current iteration, $\vec{X}$ is the position vector, $\overrightarrow{X^{*}}$ is the new position vector. '| |', $'.'$ are absolute value and element-by-element multiplication respectively. Both $\vec{A}$ and $\vec{C}$ are coefficient vectors where, $\vec{A} = 2\vec{a}.\vec{r} - \vec{a}$, $\vec{C} = 2.\vec{r}$. Here, $\vec{a}$ straightaway decreases from 2 to 0 during the course of maneuver (in both phases: exploration and exploitation) and $\vec{r} = (-1).* rand(0,1)$. The flowchart and pseudo code of WOA is given in Fig. 1 and Table 2 respectively.

Table 2. Pseudo code of Standard WOA

1	Initialize population $\mathbf{X_i}$ $(\mathbf{i = 1, 2, \ldots, n})$
2	Find the fitness of each search agent
3	$\mathbf{X^*}$is the best search agent obtained
4	Update a, A, C, l and p every search agent
5	if1 p< 0.5
6	if2 $\lvert\mathbf{A}\rvert$< 1
7	Update position using Eq. (9)
8	else if2 $\mathbf{A} \geq 1$
9	Select a random search agent $\boldsymbol{X_{rand}}$
10	Update the position using Eq. (8)
11	end if2
12	else if1 p ≥ 0.5
13	Update the topography using Eq. (7)
14	end if1
15	Check is any search agent goes beyond the search space and amend it
16	Calculate the fitness of each search agent
17	Update $\mathbf{X^*}$ if there is a better solution
18	Go to Step 4 t times and return $\mathbf{X^*}$

IV. Multi-Swarm Cooperative Strategies

This section presents the multi-swarm cooperative strategies for enhancing the performance of standard WOA by invoking the Ring, Master-Slave and Hybrid strategies individually for finding the best whale population vector. The main objective of invoking these three cooperative strategies is maintaining the balancing between exploration and exploitation phases. According to the standard WOA all the whales are attracted by the best whale position. If the whales in each sub population converge at the best whale position then within a few iteration they will move near to the surface. In invoking Cooperative Strategies into Standard WOA, The whole population is grouped into multiple-cooperative swarms, we named it as sub-swarms. These sub-swarms perform different searching behavior for finding the best whale position vector $Best\_Sub\_Pop_k{}^{*}$. This phenomenon in each whale sub-population leads to the similar updating behavior and the loss of diversity in whales. The fallowing section 4.1, 4.2 and 4.3 describes the Ring, Master-Slave and Hybrid cooperative strategies with neat flowchart and pseudo code.

A. Ring Cooperative Strategy

This Strategy is a heterogeneous search strategy based on *k*-sub-swarms. The basic idea behind this strategy was inspired from ring topology. It is one of the network topology called ring network or ring topology. Many cooperative algorithms used this cooperative strategy to improve the performance of algorithm. In this cooperative strategy each sub-population of *k* sub-population evolves in parallel to find best whale position. Each whale sub-population interchanges its best whale position with the other whale sub-population. Each sub-population replaces its best whale position only if gets the better whale position than the present best whale position. The interchange of best whale position happens continuously. In the optimal model structures of whale population, each sub-swarm can learn from the best whale position given by other sub-swarm and enhance their foraging direction i.e., bringing the prey closer to the surface which is optimal place to catch the prey. The pseudo code and flowchart is given in Table 3 and Fig. 2.

B. Master Slave Cooperative Strategy

This section presents the improved strategy of ring strategy. We mentioned “improved strategy of ring strategy” because the best whale position is interchanged like the ring cooperative strategy. But the difference is instead of interchanging best whale position between each other sub whale population, one sub whale population acts as master and all remaining are as slaves. it also called as multi-population collaborative strategy. In this cooperative strategy each slave sub-population of k-1 sub-population, evolves in parallel to obtain best whale position and share this information to the master. Master sub populations find the best among the received and share the same to the slaves. Master-Slave strategy balances the exploration and exploitation phases. The pseudo code and flowchart is given in Table 4 and Fig. 3.

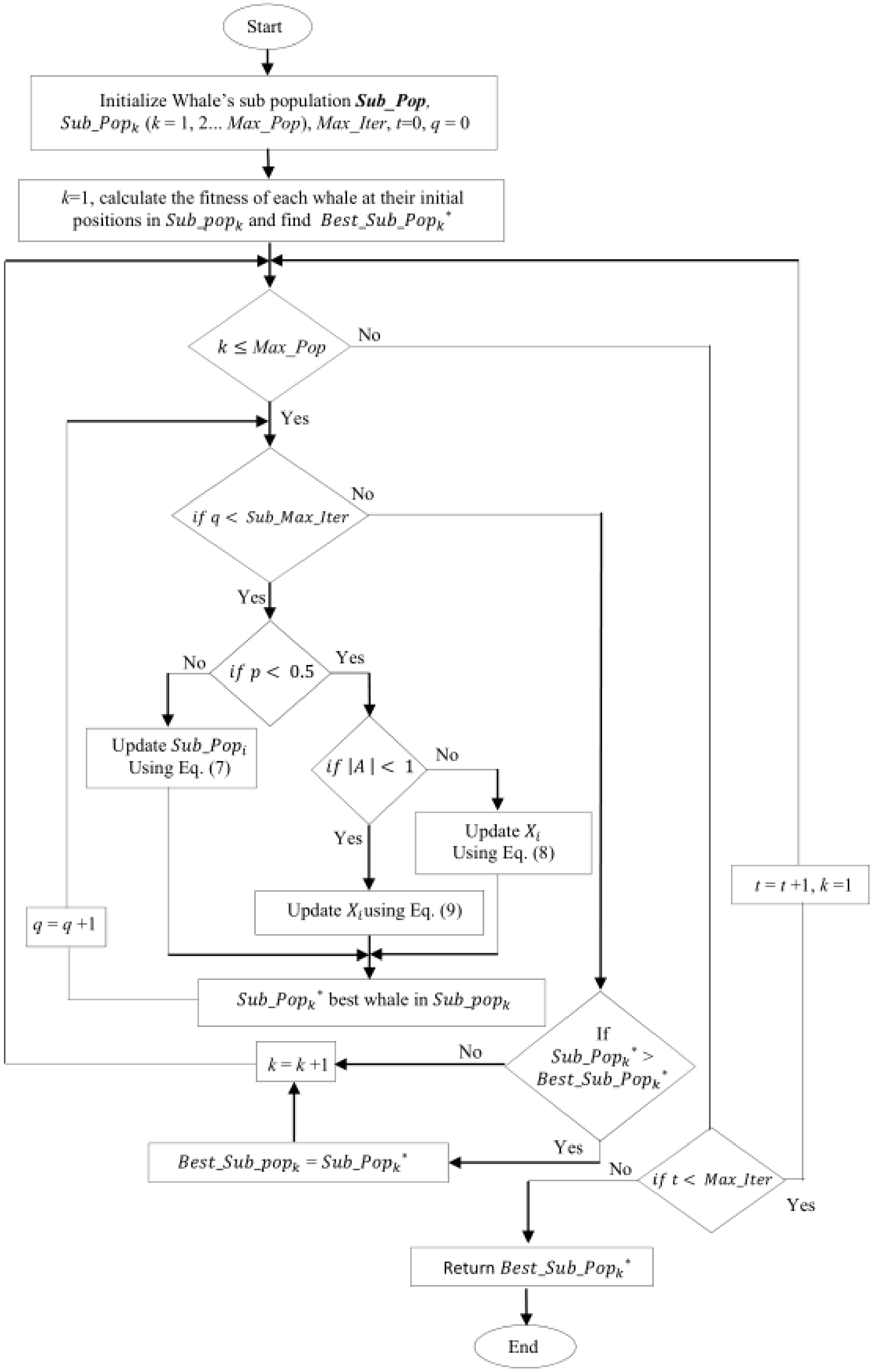


Fig.2. WOA invoking ring cooperative strategy

Table 3. Pseudo code for WOA invoking ring cooperative strategy

Input	Sub_Pop_k (k = 1, 2... *Max_Pop*), *Sub_Max_iter*, *Max_Iter*, t=0, q = 0
Output	Best solution, best optimal values of objective function
1.	Initialize Whale sub population ***Sub_Pop***, Sub_Pop_k (k = 1, 2... *Max_Pop*), *Max_Iter*, k =1, t = 0 and q = 0
2.	calculate the fitness of each whale at their initial positions in Sub_pop_k and find $Best\_Sub\_Pop_k{}^*$ ($Best\_Sub\_Pop_1{}^*$ is the best search agent obtained at k = 1)
3.	Update a, A, C, l and p every search agent
4.	**if** $k \leq Max\_Pop$
5.	**if** p < 0.5
6.	**if** $\lvert A \rvert$< 1
7.	Update the topography using $\vec{X}_{t+1} \leftarrow \overrightarrow{X^*}_t - \vec{A}.\vec{D}$
8.	**else if** $\lvert A \rvert \geq 1$
9.	Select a random search agent X_{rand} update the topography using $\vec{X}_{t+1} \leftarrow \vec{X}_{rand} - \vec{A}.\vec{D}$
10.	**end if**
11.	**else if** p ≥ 0.5
12.	Update the topography using $\vec{X}_{t+1} \leftarrow \overrightarrow{D'}.e^{bl}.\cos(2\pi l) + \overrightarrow{X^*}_t$ Check is any search agent goes beyond the search space and amend it
13.	**end if**
14.	**repeat** *Sub_Max_Iter* times
15.	**if** $Sub\_Pop_k{}^*$ > $Best\_Sub\_Pop_k{}^*$
16.	$Best\_Sub\_pop_k$ = $Sub\_Pop_k{}^*$
17.	$k = k + 1$
18.	**else if** $k = k + 1$
19.	**end if**
20.	**end if**
21.	**repeat** *Max_Iter* times
22.	return $Best\_Sub\_Pop_k{}^*$

C. Hybrid Cooperative Strategy

This section gives the detailed description of a new hybrid cooperative strategy. The discussed cooperative strategies in earlier sections 4.1 and 4.2 are combined and form a new hybrid cooperative strategy. The key objective of this hybridization is infusion of swarm cooperative mechanism of ring cooperative strategy between the slave populations into the Master-Slave cooperative strategy. In this, a duplicate set of the best whale position of each slave sub-swarm population is shared to the master sub-swarm population. Moreover, if after Max_iter (a maximum number of iterations) there is no improvement in the obtained solution, each slave sub-swarm population shares its best whale position obtained with its neighbors. Finally, the general best whale position is detected by the master sub-population.

The cooperation and communication model within the ring cooperative strategy are injected into the master and slave sub-populations and interchanging the individual best whale position of each slave sub-population with master sub-population reduces the possibility of moving outwards from the optimal solution. This hybridization process goes for maximizing the possibility of finding the best position within a low solution cost and the diversification of search can be achieved through the same. On the other hand, as the swarms are searching independently and in real parallelism, comprehensively the algorithm must be more effective than the previous ones. The pseudo code and flowchart of this cooperative strategy are given in Table 5 and Fig. 4 respectively.

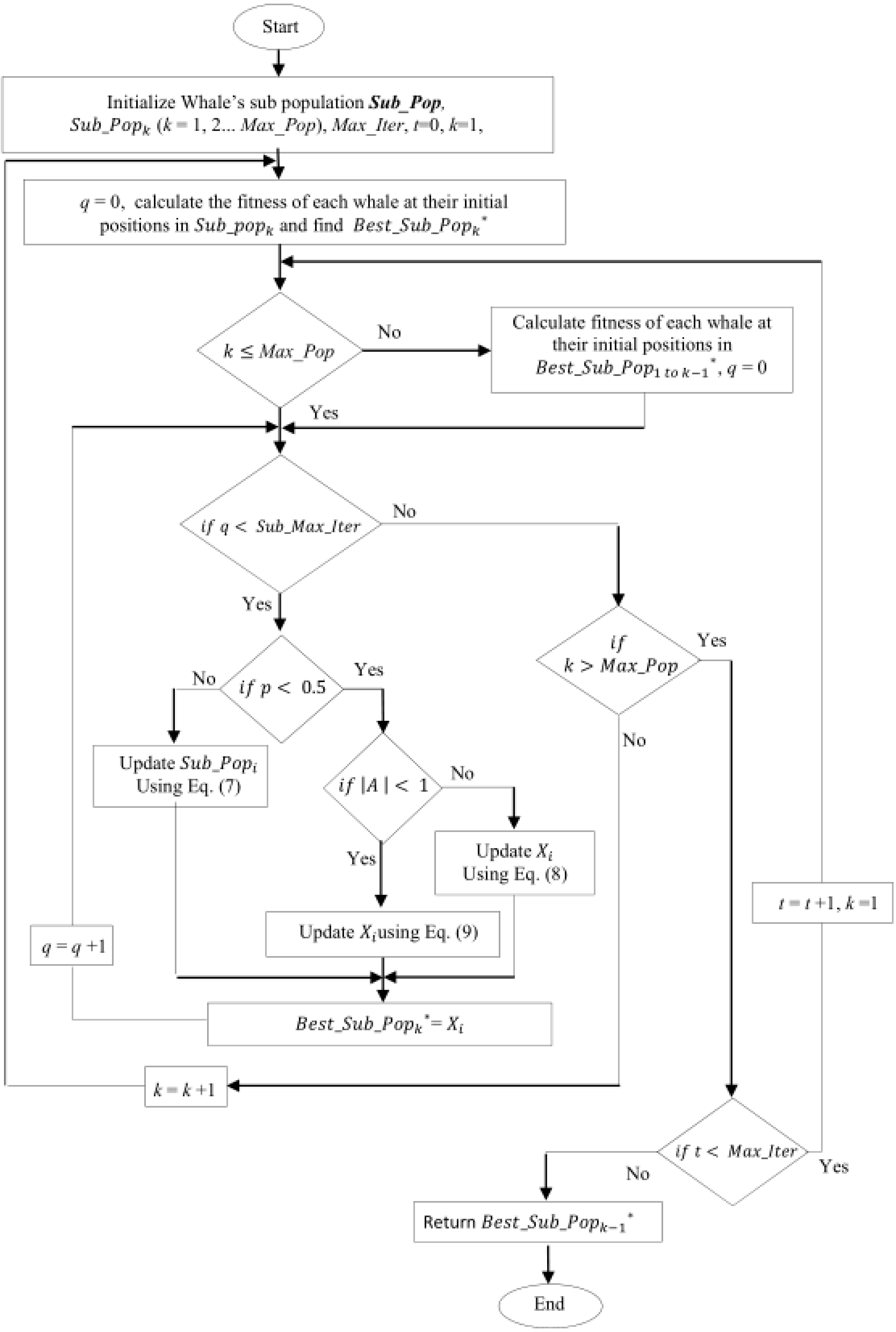


Fig.3. WOA invoking master-slave cooperative strategy

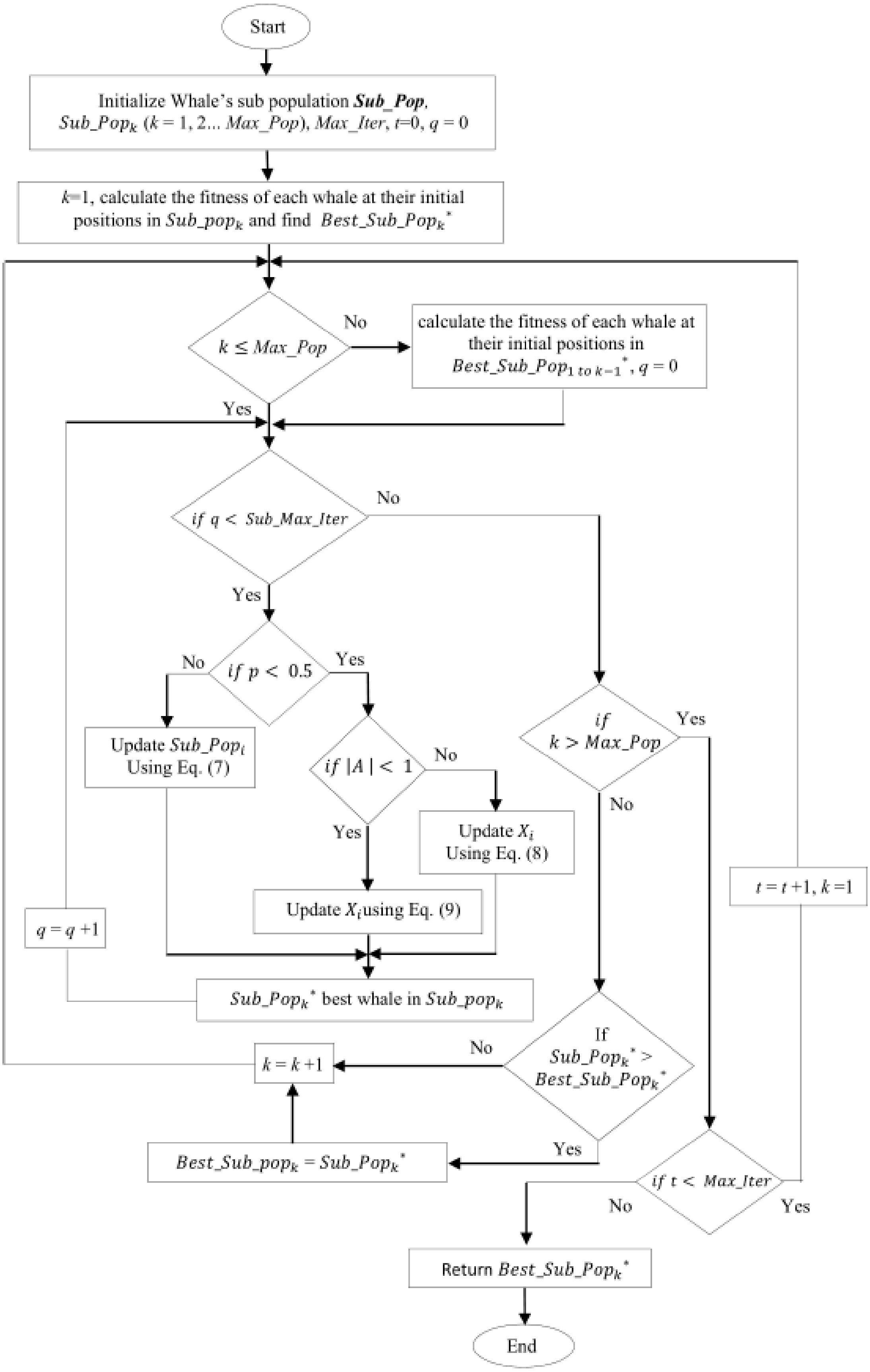


Fig.4. WOA invoking Hybrid cooperative strategy

Table 4. Pseudo code for WOA invoking Master – Slave cooperative strategy

Input	Sub_Pop_k (*k* = 1, 2... *Max_Pop*), *Sub_Max_iter, Max_Iter*, *t*=0, *q* = 0
Output	Best solution, best optimal values of objective function
1.	Initialize Whale sub population ***Sub_Pop***, Sub_Pop_k (*k* = 1, 2... *Max_Pop*), *Max_Iter*, *k* =1, *t* = 0 and *q* = 0
2.	calculate the fitness of each whale at their initial positions in Sub_pop_k and find $Best\_Sub\_Pop_k{}^*$ ($Best\_Sub\_Pop_1{}^*$ is the best search agent obtained at k = 1)
3.	Update *a*, *A*, *C*, *l* and *p* every search agent
4.	**if** $k \leq Max\_Pop$
5.	**if** *p* < 0.5
6.	**if** $\lvert A \rvert < 1$
7.	Update the topography using $\vec{X}_{t+1} \leftarrow \overrightarrow{X^*}_t - \vec{A}.\vec{D}$
8.	**else if** $\lvert A \rvert \geq 1$
9.	Select a random search agent X_{rand} update the topography using $\vec{X}_{t+1} \leftarrow \vec{X}_{rand} - \vec{A}.\vec{D}$
10.	**end if**
11.	**else if** p ≥ 0.5
12.	Update the topography using $\vec{X}_{t+1} \leftarrow \overrightarrow{D'}.e^{bl}.\cos(2\pi l) + \overrightarrow{X^*}_t$ Check is any search agent goes beyond the search space and amend it
13.	**end if**
14.	else if
15.	calculate the fitness of each whale at their initial positions in $Best\_Sub\_Pop_{1\ to\ k-1}{}^*$, *q* = 0
16.	**if** *k* > Max_Pop
17.	goto step 21
18.	else if *k* = *k* +1
19.	goto step 2
20.	**end if**
21.	**repeat step 5 to 20** *Sub_Max_Iter* times
22.	**end if**
23.	**repeat** Sub_*Max_Iter* times
24.	**repeat** step 4 to 23 *Max_Pop* times
25.	**repeat** 4 to 24 *Max_Iter*
26.	return $Best\_Sub\_Pop_k{}^*$

V. Experimental results and discussion

In the current work, the performance of the proposed multi-swarm cooperative strategies (MsWOA) is compared with GSA, DE, PSO, WOA and variants of WOA in solving 23 standard mathematical benchmark functions and data clustering problems. We considered 10 real datasets for addressing the data clustering problems of which are taken from *https://archive.ics.uci.edu*. The detailed description of these datasets is listed in Table 6. In all the experiments, we have taken whales population of 30 and maximum of 500 iterations for both standard mathematical benchmark functions optimization problems and data clustering problems. All the conducted experiments and investigations have been made on an Intel(R) Core(TM)-i5-2400 CPU with a clock rate of 3.10 GHz, 8 GB RAM and proposed multi-swarm cooperative strategies for enhancing standard WOA is described earlier sections has been implemented using MATLAB-R2016a.

A. *Experiment 1: mathematical benchmark functions*

Like the other state-of-the-art and popular population based meta-heuristic optimization techniques such as GSA, DE [12], PSO [15], WOA [25] and advanced variants IWOA [26] and NPMOA [27] of WOA, makes use of a population to investigate the problem space. Population based mechanism bring into play the probability of obtaining optimal solution and abscond from local optima increases.

In this section, the proposed multi-swarm cooperative strategies for enhancing the optimization performance of standard WOA and compared the obtained results with GSA, DE, PSO, WOA, IWOA and NPMOA based on the optimal solutions obtained on 23 standard mathematical benchmark functions. We presented the comprehensive description of these benchmark functions in Table 1 a) and Table 1 b). Note that the Table 1 a) indicates the comprehensive descriptions of 13 standard mathematical benchmark functions which include 7 Uni-modal and 6 Multimodal functions and Table 1 b) presents the description of 10 fixed dimension multimodal standard mathematical benchmark functions. The performance comparison have been made in terms of fitness convergence of GSA, DE, PSO, WOA and advanced variants IWOA and NPMOA of WOA and MsWOA for first four benchmark functions in graphical manner as given in Fig 5. The statistical analysis of obtained solution costs for MsWOA along with GSA, DE, PSO, WOA and variants of WOA is presented in a tabular form as given in Table 7. Hence, it is clear that the MsWOA is good at optimizes the most of the benchmark functions within a low cost solution than the other compared algorithms. Note that we used MsWOA for indicating all Ring (RWOA), Master-Slave (MSWOA) and Hybrid (HWOA) cooperative strategies in a single term.

Table 5. Pseudo code for WOA invoking Hybrid cooperative strategy

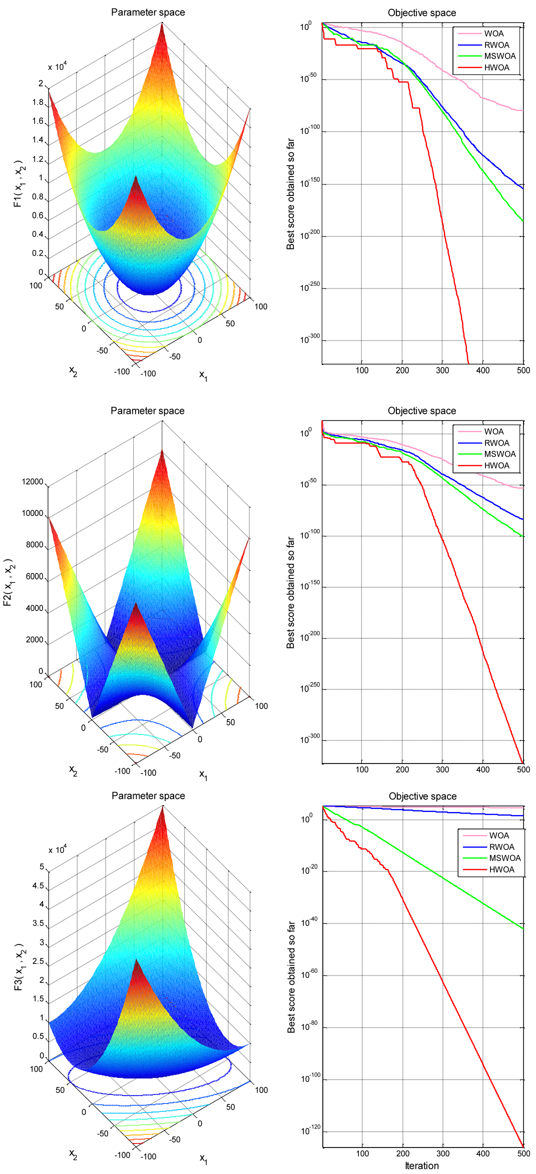
Input	Sub_Pop_k (k = 1, 2... *Max_Pop*), *Sub_Max_iter, Max_Iter*, t=0, q = 0
Output	Best solution, best optimal values of objective function
1.	Initialize Whale sub population ***Sub_Pop***, Sub_Pop_k (k = 1, 2... *Max_Pop*), *Max_Iter*, k =1, t = 0 and q = 0
2.	calculate the fitness of each whale at their initial positions in Sub_pop_k and find $Best\_Sub\_Pop_k{}^*$ ($Best\_Sub\_Pop_1{}^*$ is the best search agent obtained at k = 1)
3.	Update a, A, C, l and p every search agent
4.	**if** $k \leq Max\_Pop$
5.	**if** $p < 0.5$
6.	**if** $\lvert A \rvert < 1$
7.	Update the topography using $\vec{X}_{t+1} \leftarrow \overrightarrow{X^*}_t - \vec{A}.\vec{D}$
8.	**else if** $\lvert A \rvert \geq 1$
9.	Select a random search agent X_{rand} update the topography using $\vec{X}_{t+1} \leftarrow \vec{X}_{rand} - \vec{A}.\vec{D}$
10.	**end if**
11.	**else if** p ≥ 0.5
12.	Update the topography using $\vec{X}_{t+1} \leftarrow \overrightarrow{D'}.e^{bl}.\cos(2\pi l) + \overrightarrow{X^*}_t$ Check is any search agent goes beyond the search space and amend it
13.	**end if**
14.	else if
15.	calculate the fitness of each whale at their initial positions in $Best\_Sub\_Pop_{1\ to\ k-1}{}^*$, q = 0
16.	**if** $k > Max\_Pop$
17.	goto step 21
18.	else if $k = k + 1$
19.	goto step 2
20.	**end if**
21.	**repeat step 5 to 20** *Sub_Max_Iter* times
22.	**end if**
23.	**repeat** Sub_*Max_Iter* times
24.	**repeat** step 4 to 23 *Max_Pop* times
25.	**repeat** 4 to 24 *Max_Iter*
26.	return $Best\_Sub\_Pop_k{}^*$

B. Experiment 2: Data clustering problems

To test the efficiency of MsWOA on data clustering problems, we examined 10 real datasets. The comprehensive descriptions of these datasets are displayed in Table 6. This section for the most part centered around data clustering using the proposed technique. At the point when a partitional clustering algorithm is said to be good, it should return optimal clusters in which the clustering strategy should indulge in minimization and maximization of intra cluster and inter cluster detachment respectively. The experimental and evidentical comparisons have been made with previously established algorithms. We used the fallowing three criteria to measure the clustering performance of each clustering technique. They are,

- Finding quantization error using Eq. (10)
- Intra-cluster minimization, i.e. where the key objective is to minimize the distance between data objects within a cluster.
- Inter-cluster maximization, i.e. where the key objective is to maximize the distance between the centroids of the clusters.

$$\text{Quantization error } J_e = \frac{\sum_{j=1}^{N_c}\left[\sum_{\forall z_p \in C_{i,j}} d(z_p, m_j)/|C_{i,j}|\right]}{N_c} \quad (10)$$



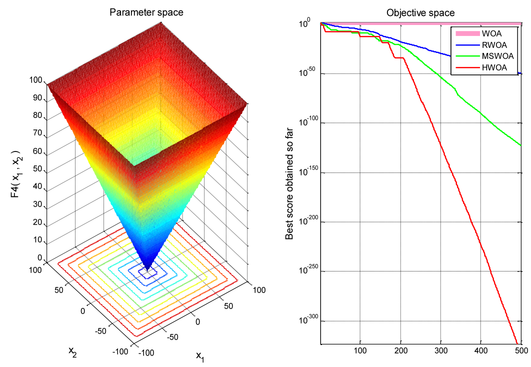


Fig.5. Fitness convergence of first four Uni-modal functions

Table 6. Description of datasets

S. No	***Dataset***	***size***	***dimension***	***k***
1	*Iris*	150	4	3
2	*Glass*	214	9	6
3	*Wine*	178	13	3
4	*Breast*	683	9	2
5	*Pima Indian diabetes*	68	8	2
6	*Haberman's survival*	306	3	2
7	*Hayes-Roth*	160	4	3
8	*E. Coli*	336	7	8
9	*Zoo*	101	16	7
10	*Vowel*	462	10	11

In order to evaluate the MsWOA clustering performance, this algorithm is run on every benchmark dataset. From the obtained results the statistical analysis has been made. The convergence characteristics of data clustering problems are presented in a graphical manner in Fig.6. The mean and standard deviation are calculated for each test. The obtained solution cost, Intra-Cluster and Intra-Cluster distances for MsWOA are summarized in Table 8 along with previously established algorithms GSA, DE, PSO, WOA, IWOA, CSAWOA. In most of the cases the hybrid cooperative strategy improved the performance of WOA. From all the experiment results obtained for data clustering problems, we can testify that MsWOA works well in solving data clustering problems. Note that for all the data cluster problems, *k*-Means partitional clustering algorithm along with the standard Euclidean distance measuring technique is used.

Table 7. Statistical analysis of obtained solution cost for F_1 to F_{23}

Alg.	F_1					F_2				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	121.248e+01	1.1549e-27	8.1245e+03	1.1845e-51	6.7843e+04	5.1532e+07	2.1562e-09	8.4238e+08	3.4525e-47	7.2637e+11
DE	132.328e+01	1.4728e-27	8.0715e+03	1.0548e-55	6.9825e+04	5.4537e+07	2.5416e-09	8.4751e+08	3.9816e-47	7.3519e+09
PSO	115.452e+01	1.4528e-29	8.0282e+03	1.2457e-87	6.8354e+04	4.2193e+07	2.2023e-09	8.7425e+08	3.5812e-47	7.2581e+10
WOA	976.0925	1.4934e-29	6.0966e+03	1.0730e-81	6.2422e+04	2.9207e+07	1.7031e-18	4.6134e+08	3.4225e-52	7.3017e+09
IWOA	141.0601	1.6987e-69	2.8017e+03	8.5331e-315	6.2234e+04	2.6816e+08	1.1694e-32	5.9961e+09	2.1158e-159	1.3408e+11
CSAWOA	143.4836	1.2152e-78	3.2060e+03	0	7.1689e+04	1.0502e+06	3.8957e-42	2.3483e+07	1.1054e-202	5.2510e+08
RWOA	278.4197	3.5175e-53	3.4636e+03	3.0772e-155	6.5930e+04	1.1445e+09	9.0683e-30	2.5592e+10	7.6042e-92	5.7226e+11
MSWOA	148.2499	1.0370e-55	2.1034e+03	3.6666e-186	4.2005e+04	6.2541e+08	2.8133e-31	1.3984e+10	8.5322e-102	3.1270e+11
HWOA	133.7060	3.1538e-94	2.9709e+03	0	6.6431e+04	4.1330e+06	1.0576e-49	9.2418e+07	5.5813e-318	2.0665e+09

Alg.	F_3					F_4				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	5.2318e+05	5.4532e+04	3.8421e+05	2.5136e+03	1.3428e+05	1.7672	2.4365e-36	5.6827	6.4162e-26	112.3518
DE	4.9845e+05	4.1429e+04	3.5432e+05	2.3684e+04	1.3428e+05	1.7516	2.6842e-36	5.6845	6.6273e-26	110.438
PSO	4.7536e+05	4.2137e+04	3.1452e+05	2.2531e+04	1.3549e+05	1.2461	2.6821e-36	5.3245	6.5736e-49	100.2456
WOA	5.7966e+04	5.2965e+04	3.4381e+04	1.6992e+04	1.2689e+05	1.4157	2.2081e-35	5.2690	6.3572e-38	90.7160
IWOA	1.2377e+03	2.4741e-60	1.6451e+04	9.7553e-270	2.5731e+05	0.2815	1.1628e-36	4.2173	4.6036e-149	87.9804
CSAWOA	676.4248	7.3566e-74	9.4417e+03	0.784 e-270	1.6777e+05	0.2573	3.1044e-38	4.2628	1.9567e-208	86.0427
RWOA	2.6761e+04	2.0174e+03	5.0740e+04	30.4146	2.4953e+05	0.8732	6.0729e-24	6.7354	2.3219e-50	92.1388
MSWOA	2.3047e+03	2.8134e-18	1.7782e+04	1.1087e-42	2.1653e+05	0.3378	1.8580e-37	4.3973	1.5083e-123	85.7551
HWOA	635.1985	2.6796e-47	8.1064e+03	9.0514e-127	1.5907e+05	0.2516	1.4527e-76	4.1991	0	87.8658

Alg.	F_5					F_6				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	2.9438e+06	28.5925	1.7486e+08	29.8465	2.7165e+08	648.9124	0.4864	4.8462e+03	0.4934	6.7168e+04
DE	2.7268e+06	28.3516	1.7345e+08	29.5438	2.7465e+08	650.4671	0.45311	4.8513e+03	0.4271	6.4625e+04
PSO	2.8436e+06	28.2816	1.7519e+08	29.1575	2.6427e+08	635.9215	0.3843	4.7364e+03	0.4637	5.4928e+04
WOA	2.6080e+06	27.9527	1.9373e+07	27.9516	2.5396e+08	629.8011	0.3204	4.7976e+03	0.3204	5.5280e+04
IWOA	4.0073e+05	27.8797	8.7116e+06	27.8797	1.9476e+08	168.6582	0.1528	3.2574e+03	0.1528	7.2028e+04
CSAWOA	5.6167e+05	28.6617	1.2501e+07	28.6617	2.7952e+08	135.4935	0.2798	2.8557e+03	0.2798	6.3775e+04
RWOA	6.4358e+05	27.5877	1.0019e+07	27.5877	2.1035e+08	739.3369	0.3262	6.0110e+03	0.3262	7.2705e+04
MSWOA	5.9859e+05	27.9460	1.2597e+07	27.9460	2.8119e+08	607.4678	0.1720	4.6416e+03	0.1720	6.1839e+04
HWOA	5.6736e+05	28.4372	1.2686e+07	28.4371	2.8366e+08	158.9282	0.1634	2.7791e+03	0.1633	6.0336e+04

Alg.	F_7					F_8				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	2.6218	0.9261	12.9861	0.2897	150.6465	-1.2942e+04	-1.5429e+04	1.8435e+03	-1.8415e+04	-3.4791e+03
DE	2.6284	0.8649	12.8534	0.2785	150.6481	-1.2648e+04	-1.6354e+04	1.7926e+03	-1.7816e+04	-3.4627e+03
PSO	1.6845	0.8634	12.6521	0.0654	148.5281	-1.2549e+04	-1.4215e+04	1.7924e+03	-1.8195e+04	-3.4996e+03
WOA	1.3233	0.0057	11.1443	0.0052	139.6725	-1.1491e+04	-1.2337e+04	1.7492e+03	-1.2347e+04	-2.4467e+03
IWOA	0.2960	2.3889e-04	6.2970	4.0297e-05	140.7692	-1.2393e+04	-1.2560e+04	897.0959	-1.2569e+04	-1.7542e+03
CSAWOA	0.2358	8.3541e-04	5.2301	9.0942e-05	116.9492	-1.1683e+04	-1.2494e+04	1.3604e+03	-1.2569e+04	-3.4988e+03
RWOA	0.2866	4.1339e-05	6.3802	4.1339e-05	142.6656	-1.1949e+04	-1.2569e+04	1.4897e+03	-1.2569e+04	-2.2332e+03
MSWOA	0.3021	4.5624e-04	6.7172	7.0342e-05	150.2019	-1.1731e+04	-1.2430e+04	1.5558e+03	-1.2532e+04	-2.2201e+03
HWOA	0.1892	1.3918e-04	4.2067	1.3918e-04	94.0648	-1.1711e+04	-1.2201e+04	1.3690e+03	-1.2201e+04	-1.5414e+03

Alg.	F_9					F_{10}				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	30.6527	0.2241	87.2954	0.0495	490.8134	1.2816	6.8435e-14	2.9043	4.6205e-15	20.6019
DE	30.5468	0.2168	86.6218	0.0413	490.6218	1.1956	6.5269e-15	3.0483	3.9352e-15	20.7294
PSO	30.2458	0.1762	86.1954	0.0361	485.1668	0.6583	6.6024e-15	3.0201	4.6205e-15	20.6903
WOA	27.1213	0	76.7544	0	450.6246	0.4300	6.2172e-15	2.5248	4.4409e-15	20.4593
IWOA	2.8204	0	28.9688	0	467.2299	0.1190	8.8818e-16	1.2921	8.8818e-16	20.8727
CSAWOA	1.6330	0	24.0741	0	447.3239	0.1025	8.8818e-16	1.2812	8.8818e-16	20.6321
RWOA	4.6452	0	35.8709	0	426.5025	0.2884	4.4409e-15	1.9570	4.4409e-15	20.7436
MSWOA	0.9208	0	20.5897	0	460.3999	0.0420	8.8818e-16	0.9380	8.8818e-16	20.9750
HWOA	1.4091	0	22.7304	0	425.1632	0.0448	8.8818e-16	0.9233	8.8818e-16	20.5905

Alg.	F_{11}					F_{12}				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	7.1582	0	49.2604	0	615.1082	7.6207e+06	0.0486	5.7261e+07	0.0486	7.4182e+08
DE	7.3518	0	49.1806	0	615.3409	7.0681e+06	0.0482.	5.8216e+07	0.0482	7.9208e+08
PSO	6.9046	0	40.6608	0	614.7406	6.8106e+06	0.0465	5.6152e+07	0.0465	7.8265e+08
WOA	6.8496	0	51.1150	0	573.1051	6.5367e+06	0.0472	5.6013e+07	0.0472	7.4401e+08
IWOA	2.5364	0	37.8199	0	605.3266	1.7024e+06	0.0101	3.1591e+07	0.0101	6.8750e+08
CSAWOA	1.3268	0	29.4677	0	658.9239	1.4370e+06	0.0073	3.2097e+07	0.0073	7.1770e+08
RWOA	2.5074	0	31.0664	0	610.5245	9.9647e+05	0.0126	1.7356e+07	0.0126	3.7500e+08
MSWOA	1.1732	0	25.8491	0	577.9697	1.3963e+06	0.0149	3.1223e+07	0.0149	6.9817e+08
HWOA	1.1028	0	24.3672	0	501.3912	1.2914e+06	0.0130	2.8878e+07	0.0130	6.4572e+08

Table 7. Continued

Alg.	F_{13}					F_{14}				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	8.8435e+06	0.4261	8.2696e+07	0.4261	3.4791e+08	3.9406	2.9143	2.5492	2.9143	304.6271
DE	9.4926e+06	0.3158	8.8716e+07	0.3158	3.4627e+08	3.9603	2.9084	2.5834	2.9084	304.6419
PSO	9.5109e+06	0.3151	8.3418e+07	0.3151	3.4996e+08	3.5038	2.9864	2.5681	2.9864	201.1934
WOA	9.6103e+06	0.2720	8.2630e+07	0.2720	9.5398e+08	3.4340	2.9821	2.6263	2.9821	43.0259
IWOA	2.5621e+06	0.3602	5.6061e+07	0.3602	1.2533e+09	1.8742	0.9980	18.6189	0.9980	416.9956
CSAWOA	1.2597e+06	0.1022	2.8129e+07	0.1020	6.2899e+08	1.3906	0.9980	5.3740	0.9980	115.8909
RWOA	3.9986e+06	0.2009	6.7180e+07	0.2008	1.4467e+09	1.3106	0.9980	4.8911	0.9980	78.4131
MSWOA	3.4805e+06	0.2839	7.7827e+07	0.2838	1.7403e+09	1.2186	0.9980	2.2341	0.9980	46.0861
HWOA	2.1650e+06	0.2324	4.8411e+07	0.2323	1.0825e+09	1.2154	0.9921	1.0419	0.9921	5.6662

Alg.	F_{15}					F_{16}				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	0.0072	0.0089	0.0348	0.0089	0.6428	-1.0364	-1.0316	0.2487	-1.0316	0.1791
DE	0.0086	0.0075	0.0365	0.0075	0.6612	-1.0363	-1.0316	0.2674	-1.0316	0.1772
PSO	0.0065	0.0081	0.0349	0.0081	0.5241	-1.0316	-1.0316	0.1642	-1.0316	0.1541
WOA	0.0015	3.2965e-04	0.0205	3.2964e-04	0.4502	-1.0305	-1.0316	0.0089	-1.0316	−0.8963
IWOA	9.7449e-04	4.6921e-04	0.0038	4.6921e-04	0.0746	-1.0198	-1.0316	0.1012	-1.0316	0.1630
CSAWOA	7.6984e-04	3.1298e-04	0.0056	3.1298e-04	0.1238	-1.0188	-1.0316	0.1130	-1.0316	1.0587
RWOA	0.0016	4.9407e-04	0.0210	4.9407e-04	0.4707	-0.9916	-0.9999	0.1416	-0.9999	2.1355
MSWOA	6.6054e-04	4.3257e-04	0.0023	4.3257e-04	0.0485	-0.9878	-0.9999	0.1989	-0.9999	3.4022
HWOA	6.0123e-04	3.7381e-04	0.0028	3.7258e-04	0.0599	-0.9931	-1.0000	0.0350	-1.0000	-0.7368

Alg.	F_{17}					F_{18}				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	
GSA	0.5247	0.3971	0.1736	0.3979	1.3601	3.4518	3.0000	3.0814	3.0000	34.4201
DE	0.5716	0.3969	0.1705	0.3979	1.4608	3.5806	3.0000	3.0514	3.0000	34.1089
PSO	0.4975	0.3951	0.1674	0.3979	2.3072	3.4841	3.0000	3.2153	3.0000	34.0045
WOA	0.4256	0.3982	0.1260	0.3979	1.2673	3.5112	3.0000	3.5737	3.0000	34.9664
IWOA	0.4115	0.3979	0.1003	0.3979	1.2713	3.3809	3.0000	3.0850	3.0000	33.3362
CSAWOA	0.4000	0.3979	0.0155	0.3979	0.5795	3.0159	3.0000	0.1765	3.0000	5.7108
RWOA	0.4324	0.3983	0.5675	0.3979	12.8981	3.1832	3.0184	1.3085	3.0183	22.4620
MSWOA	0.4195	0.4024	0.0932	0.3985	1.0922	3.2130	3.0027	0.8532	3.0022	8.7739
HWOA	0.4236	0.4001	0.1144	0.3999	1.6499	3.7103	3.0000	5.2395	3.0000	79.2618

Alg.	F_{19}					F_{20}				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	-3.8514	-3.8567	0.0721	-3.8587	-3.4802	-3.2142	-2.0701	0.0618	-3.1014	-1.7692
DE	-3.8587	-3.8712	0.0506	-3.8654	-2.9634	-2.8401	-2.4602	0.0681	-3.0285	-1.5593
PSO	-3.8297	-3.8409	0.0501	-3.8542	-3.2409	-3.2406	-3.0019	0.0605	-1.4018	-1.4608
WOA	-3.8218	-3.8560	0.0616	-3.8565	-3.4609	-2.9480	-2.9899	0.0695	-2.9900	-2.3866
IWOA	-3.8585	-3.8613	0.0411	-3.8614	-2.9634	-3.2169	-3.3014	0.1696	-3.3015	-1.7692
CSAWOA	-3.8290	-3.8549	0.0429	-3.8549	-3.2802	-2.9522	-3.0017	0.1452	-3.0235	-1.5593
RWOA	-3.7960	-3.8249	0.0552	-3.8260	-3.3592	-3.2815	-3.3143	0.1464	-3.3143	-1.5171
MSWOA	-3.8128	-3.8612	0.0725	-3.8612	-3.5227	-2.6257	-2.6756	0.2733	-2.6756	-0.8829
HWOA	-3.8057	-3.8553	0.1308	-3.8553	-3.0619	-3.0471	-3.1437	0.2353	-3.1437	-1.4009

Alg.	F_{21}					F_{22}				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**	**Average**	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	-3.9021	-4.5407	2.0154	-3.1408	-0.9154	-2.1008	-3.8560	0.3144	-3.0019	-0.4102
DE	-3.7463	-4.3518	2.4106	-3.9215	-0.9425	-3.1473	-3.0417	0.4935	-2.9480	-0.5657
PSO	-4.4026	-4.0407	0.3109	-4.2831	-9.7106	-3.4083	-3.8409	0.2954	-2.9900	-0.9142
WOA	-4.9423	-5.0547	0.4996	-5.0547	-0.5249	-3.6383	-3.7213	0.3054	-3.7213	-0.5457
IWOA	-9.9569	-10.0979	0.8079	-10.1128	-2.0944	-6.3471	-6.8573	0.8925	-6.8695	-1.0269
CSAWOA	-6.9943	-5.0521	2.5053	-9.9571	-0.4070	-8.1818	-9.4109	1.9301	-9.4130	-0.5524
RWOA	-5.0086	-5.0528	0.3067	-5.0529	-0.4786	-9.3501	-10.3475	1.9994	-10.3630	-0.7678
MSWOA	-4.5664	-4.6989	0.2979	-4.6989	-0.4279	-3.3935	-3.8828	0.8519	-3.8828	-0.4594
HWOA	-4.4230	-4.5876	0.4779	-4.5876	-0.3629	-4.9339	-4.9794	0.3872	-4.9794	-0.3981

Alg.	F_{23}				
	Average	**Median**	**Std. Dev**	**Best**	**Worst**
GSA	-1.9215	-4.0019	0.4512	-0.8963	-0.9142
DE	-1.1408	-3.4602	0.4654	-0.1541	-1.5893
PSO	-1.2831	-4.5593	0.4171	-0.1630	-1.5593
WOA	-5.0517	-5.1273	0.3852	-5.1273	-1.3220
IWOA	-2.3074	-2.4192	0.3146	-2.4192	-0.6555
CSAWOA	-5.5784	-5.1016	1.5542	-9.0782	-0.8072
RWOA	-2.3955	-2.4244	0.1182	-2.4251	-0.7247
MSWOA	-3.8022	-4.0244	0.5487	-4.0244	-0.6114
HWOA	-4.7822	-5.0209	0.6715	-5.0209	-1.1131

Table 8. Mean and std. dev. over 30 independent runs for 500 iterations for 10 real datasets

Alg.	Iris			Glass		
	Solution cost	**Intra-cluster**	**Inter-cluster**	**Solution cost**	**Intra-cluster**	**Inter-cluster**
GSA	0.2791±0.0157	2.4925 ±0.3709	1.9247 ±0.2746	0.0565±0.0049	5.2019 ±0.8524	5.2634 ±6.4387
DE	0.2787±0.0151	2.2405 ±0.3918	1.9219 ±0.2778	0.0584±0.0027	5.2113 ±0.8509	4.7738 ±9.2009
PSO	0.2787±0.0152	2.4703 ±0.3725	1.9701 ±0.2784	0.0428±0.0047	5.0216 ±0.7034	6.0021 ±7.3475
WOA	0.2672±0.0141	2.3902 ±0.3562	1.9291 ±0.2789	0.0443±0.0042	5.0199 ±0.7014	4.7812 ±9.2109
IWOA	0.2765±0.0042	2.2258 ±0.3601	1.9121 ±0.2821	0.0436±0.0027	4.9542 ±0.5987	5.2514 ±6.4521
CSAWOA	0.2681±0.0071	2.2258 ±0.3429	1.9623 ±0.2731	0.0435±0.0029	4.8781 ±0.8924	6.0081 ±7.3125
RWOA	0.2689±9.5741e-004	2.1222 ±0.3522	1.9802 ±0.2189	0.0433±0.0017	4.8745 ±0.8957	6.4189 ±6.9702
MSWOA	0.2655±0.0067	2.0641 ±0.3434	1.9291 ±0.2789	0.0429±0.0031	4.0934 ±0.7134	6.7204 ±9.0704
HWOA	0.2645±0.0064	2.0176 ±0.3427	1.9802 ±0.2189	0.0414±0.0014	4.3705 ±0.7447	6.7612 ±6.4308
Alg.	**Wine**			**Breast**		
	Solution cost	**Intra-cluster**	**Inter-cluster**	**Solution cost**	**Intra-cluster**	**Inter-cluster**
GSA	360.8613±1.8143	364.6584 ±8.6347	328.5411 ±3.4305	0.0672±0.0214	17.1289 ±3.7189e-015	13.4628 ±0
DE	359.4155±11.1054	362.0013 ±9.8672	359.8273 ±54.4602	0.0621±0.0275	16.9721 ±7.46e-15	13.4654 ±0
PSO	343.8378±3.5681	350.1967 ±7.5448	349.7465 ±46.7305	0.0563±0.0361	17.0346 ±8.01e-15	13.5438 ±0
WOA	340.1125±10.9298	350.1951 ±7.5439	329.2461 ±71.9199	0.0561±0.0285	17.0246 ±7.95e-15	13.4529 ±0
IWOA	340.8713±1.7802	341.7484 ±8.6213	329.2701 ±72.4532	0.0552±0.0128	17.0238 ±8.01e-15	13.4536 ±0
CSAWOA	342.8378±3.6219	341.0013 ±9.8928	363.7845 ±46.7197	0.0546±0.0341	17.0238 ±7.89e-15	13.4523 ±0
RWOA	341.7219±6.8014e-014	332.7921 ±2.2318	359.8712 ±52.3245	0.0551±6.7872e-004	16.9804 ±7.45e-15	13.5521 ±0
MSWOA	341.6293±1.6702	330.1041 ±7.4414	345.2121 ±47.3042	0.0547±0.1385	16.0426 ±7.75e-15	13.4521 ±0
HWOA	339.6378±1.6716	330.1181 ±7. 4034	329.2711 ±47.9159	0.0534±0.0234	16.4036 ±7.71e-15	13.4520 ±0
Alg.	**Diabetes**			**Haberman**		
	Solution cost	**Intra-cluster**	**Inter-cluster**	**Solution cost**	**Intra-cluster**	**Inter-cluster**
GSA	14.5400±0.1048	730.0298 ±0.1502	30.4512 ±0.4152	5.4800±0.3273	40.3214 ±2.1089e-014	5.2185 ±8.8425e-016
DE	14.5701±0.1814	713.6642 ±74.3624	36.2192 ±12.2457	5.5629±0.2713	35.1085 ±3.5591	8.7361 ±10.8134
PSO	14.4416±0.0548	713.6502 ±74.3512	36.2192 ±12.2401	5.7000±0.3045	35.3451 ±3.6982	7.3547 ±9.8732
WOA	14.2800±0.1654	708.0127 ±74.4521	36.2301 ±14.7728	5.7700±0.2845	35.3026 ±3.6534	7.6424 ±9.8043
IWOA	13.8000±0.0954	707.0087 ±74.6214	36.2354 ±14.7602	5.7800±0.3215	35.2445 ±3.784	7.9401 ±9.9731
CSAWOA	13.8524±0.0459	705.9815 ±74.3714	36.8004 ±14.4543	5.3300±0.3545	35.2064 ±3.0475	7.9821 ±9.9801
RWOA	13.7689±0.0025	705.9802 ±74.3512	36.8425 ±14.4512	5.5689±3.6142e-006	34.9885 ±3.6875	8.1285 ±10.4219
MSWOA	13.7741±0.0454	705.0164 ±74.6151	36.3051 ±14.3508	5.3655±0.3065	34.3721 ±3.5504	7.3047 ±9.8042
HWOA	13.7453±0.0656	705.0045 ±74.3402	36.3801 ±14.3461	5.2814±0.3061	34.3125 ±3.5530	7.3607 ±9.8342
Alg.	**Hayes-Roth**			**E. Coli**		
	Solution cost	**Intra-cluster**	**Inter-cluster**	**Solution cost**	**Intra-cluster**	**Inter-cluster**
GSA	0.3279±0.0324	3.5643 ±0.4428	1.3915 ±0.3541	0.0672±0.0373	0.6648 ±0.0834	0.2445 ±0.0598
DE	0.3928±0.0346	3.518 ±0.2735	1.8519 ±0.7531	0.0663±0.0032	0.6643 ±0.0735	0.3637 ±0.1341
PSO	0.3507±0.0254	3.5081 ±0.4318	1.4189 ±0.6078	0.0671±0.0030	0.6312 ±0.0726	0.3145 ±0.1028
WOA	0.3672±0.0284	3.5492 ±0.4865	1.4256 ±0.6845	0.0641±0.0028	0.6745 ±0.0825	0.3176 ±0.1517
IWOA	0.3765±0.0251	3.5489 ±0.4812	1.4364 ±0.6704	0.0672±0.0028	0.6545 ±0.0754	0.3522 ±0.1102
CSAWOA	0.2681±0.0248	3.5413 ±0.4716	1.7012 ±0.6146	0.0666±0.0030	0.6245 ±0.0689	0.4540 ±0.1056
RWOA	0.2689±0.0129	3.2508 ±0.3512	1.8251 ±0.6702	0.0641±0.0012	0.6012 ±0.0542	0.4643 ±0.1728
MSWOA	0.2655±0.0105	3.3415 ±0.4181	1.4046 ±0.6451	0.0637±0.00124	0.6015 ±0.0621	0.3206 ±0.1410
HWOA	0.2645±0.0108	3.5472 ±0.4160	1.4053 ±0.6415	0.0624±0.00113	0.6018 ±0.0721	0.3603 ±0.1421
Alg.	**Zoo**			**Vowel**		
	Solution cost	**Intra-cluster**	**Inter-cluster**	**Solution cost**	**Intra-cluster**	**Inter-cluster**
GSA	0.0202±0.0026	1.6378 ±1.1528e-15	2.2315 ±5.0214e-016	41.9500±1.3465	901.3251 ±108.8745	348.3268 ±230.3519
DE	0.0220±0.0024	1.6512 ±1.1324e-15	2.6502 ±2.278e-15	39.2800±2.2481	898.617 ±102.9459	353.3712 ±218.2482
PSO	0.0130±0.0018	1.6396 ±1.1244e-15	2.6421 ±2.256e-15	38.4700±2.3618	899.254 ±101.8954	350.3465 ±215.7135
WOA	0.0120±0.0022	1.6401 ±1.1312e-15	2.6471 ±2.274e-15	38.4012±2.2314	899.246 ±101.8705	350.3548 ±215.6548
IWOA	0.0110±0.0021	1.6391 ±1.1268e-15	2.6484 ±2.25e-15	37.6824±1.3545	899.238 ±102.3254	350.7091 ±215.9805
CSAWOA	0.0090±0.0019	1.6381 ±1.1254e-15	2.6487 ±2.258e-15	36.8140±2.3456	899.228 ±102.3815	350.8086 ±215.9501
RWOA	0.0094±8.3665-04	1.5254 ±4.5498e-016	2.6502 ±2.256e-15	35.8941±2.2145	898.145 ±102.9821	353.3282 ±218.2545
MSWOA	0.0100±0.0012	1.5332 ±1.1268e-14	2.6451 ±2.264e-15	34.1553±2.2645	898.224 ±102.3142	350.7301 ±215.9611
HWOA	0.00534±0.0024	1.4301 ±1.1341e-14	2.6432 ±2.261e-15	32.4501±2.2416	898.138 ±102.3107	350.6126 ±215.9411

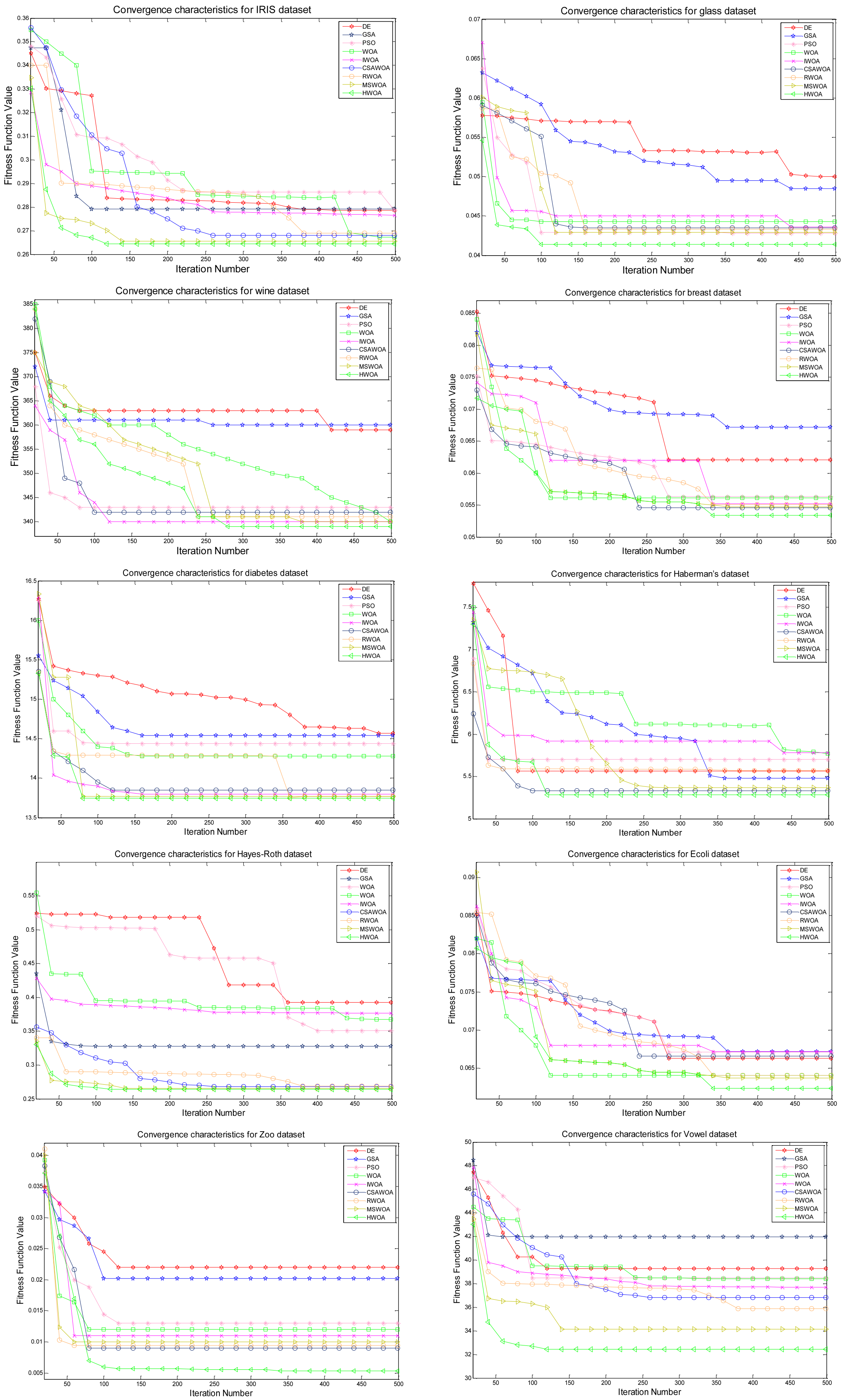


Fig.6. Convergence of data clustering problems

VI. Conclusions and Future Work

In this work we presented multi-swarm cooperative strategies for addressing two problems. They are solving standard mathematical benchmark functions problems and solving the data clustering problems. In this process we have used the standard whale optimization algorithm. We utilized three cooperative strategies namely the Ring, Master-Slave and Hybrid cooperative strategies to enhance the performance of standard WOA. Several experiments have been conducted and the performance of each cooperative strategy is compared with previously established techniques GSA, DE, PSO, WOA and also variants of WOA. The results established the multi-swarm cooperative strategies outperform in solving benchmark functions as well as the data clustering problems compared GSA, DE, PSO, WOA and variants of WOA for 10 different real databases. All the obtained graphical and statistical results are reported respective figures and tables. The reported values were averaged over 40 simulations to specify the algorithms convergence range. In most cases, the proposed multi-swarm cooperative strategies realized lower quantization errors. Generally, the proposed multi-swarm cooperative strategies establish their efficiency in finding optimizing benchmark functions as well as the data clustering problems.

This explore run may be extended for further research in the shadowing dimensions: In this research work, the show on multiple swarm based cooperative strategies proposed for enhancing the numerical optimization performance of WOA and these are compared with other-state-of-the-art algorithms and variants of WOA for its action judgement on 23 standard benchmark functions. As a future line, added new versions of WOA could also be tried to key the unexceeded type in achieving test suite improvement. This research work concentrate on standard WOA optimization. In the future we go through the new population based meta-heuristic optimization algorithms which are also inspired by the nature.

References

[1] Kacprzyk, Janusz, and Witold Pedrycz, eds. Springer handbook of computational intelligence. Springer, 2015.

[2] Amiri, Babak. "Application of teaching-learning-based optimization algorithm on cluster analysis." Journal of Basic and Applied Scientific Research 2.11 (2012): 11795-11802.

[3] Hoang, Duc Chinh, et al. "Real-time implementation of a harmony search algorithm-based clustering protocol for energy-efficient wireless sensor networks." IEEE Transactions on Industrial Informatics 10.1 (2014): 774-783.

[4] Kuo, R. J., et al. "Integration of particle swarm optimization and genetic algorithm for dynamic clustering." Information Sciences 195 (2012): 124-140.

[5] Niknam, Taher, et al. "An efficient hybrid evolutionary optimization algorithm based on PSO and SA for clustering." Journal of Zhejiang University-SCIENCE A 10.4 (2009): 512-519.

[6] Nanda, Satyasai Jagannath, and Ganapati Panda. "A survey on nature inspired metaheuristic algorithms for partitional clustering." Swarm and Evolutionary computation 16 (2014): 1-18.

[7] Holland JH . Genetic algorithms. Sci Am 1992;267:66–72.

[8] Rahman, Md Anisur, and Md Zahidul Islam. "A hybrid clustering technique combining a novel genetic algorithm with K-Means." Knowledge-Based Systems 71 (2014): 345-365.

[9] Dorigo M, Gambardella LM (1997) Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Trans Evol Comput 12:53–66.

[10] Shelokar, P. S., Valadi K. Jayaraman, and Bhaskar D. Kulkarni. "An ant colony approach for clustering." Analytica Chimica Acta 509.2 (2004): 187-195.

[11] Korürek, Mehmet, and Ali Nizam. "A new arrhythmia clustering technique based on Ant Colony Optimization." Journal of Biomedical Informatics 41.6 (2008): 874-881.

[12] Storn R, Price K (2010) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J lobal Optim 23:689–694.

[13] Das, Swagatam, Ajith Abraham, and Amit Konar. "Automatic clustering using an improved differential evolution algorithm." IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans 38.1 (2008): 218-237.

[14] Kwedlo, Wojciech. "A clustering method combining differential evolution with the K-means algorithm." Pattern Recognition Letters 32.12 (2011): 1613-1621.

[15] Kennedy J , Eberhart R . Particle swarm optimization. In: Proceedings of the 1995 IEEE international conference on neural networks; 1995. p. 1942–8.

[16] Ahmed, A., et al. "A review on particle swarm optimization algorithm and its variants to clustering high-dimensional data." The Artificial Intelligence Review 44.1 (2015): 23.

[17] Alam, Shafiq, et al. "Research on particle swarm optimization based clustering: a systematic review of literature and techniques." Swarm and Evolutionary Computation 17 (2014): 1-13.

[18] Clerc M, Kennedy J (2000) The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Trans Evol Comput 6(1):58–73

[19] Mendes R, Kennedy J, Neves J (2004) The fully informed particle swarm: simpler, maybe better. IEEE Trans Evol Comput 8:204–210

[20] Peram T, Veeramachaneni K, Mohan CK (2003) Fitness-distance- ratio based particle swarm optimization. In: Proceedings of the IEEE swarm intelligence symposium, pp 174–181

[21] Naik A, Satapathy SC, Parvathi K (2013) A comparative analysis of results of data clustering with variants of particle swarm optimization. In: International conference on swarm, evolutionary, and Memetic computing, pp 180–192

[22] Yohannes, M. S. "Solving economic load dispatch problem using particle swarm optimization technique." *International Journal of Intelligent Systems and Applications* 4.12 (2012):

[23] Ali Khazaee,"Heart Beat Classification Using Particle Swarm Optimization", *International Journal of Intelligent Systems and Applications*, vol.5, no.6, pp.25-33, 2013.

[24] Poonam Singhal, S. K. Agarwal, Narendra Kumar,"Advanced Adaptive Particle Swarm Optimization based SVC Controller for Power System Stability", *International Journal of Intelligent Systems and Applications*, vol.7, no.1, pp.101-110, 2015.

[25] Rashedi E , Nezamabadi-Pour H , Saryazdi S . GSA: a gravitational search algo- rithm. Inf Sci 2009;179:2232–48.

[26] 20. Hatamlou, A., Abdullah, S., Nezamabadi-pour, H.: A Combined Approach for Clustering Based on K-means and Gravitational Search Algorithms. Swarm and Evolutionary Computation 6, 47–52 (2012)

[27] Hatamlou, A., Abdullah, S., Nezamabadi-pour, H.: Application of Gravitational Search Algorithm on Data Clustering. In: Yao, J., Ramanna, S., Wang, G., Suraj, Z. (eds.) RSKT 2011. LNCS, vol. 6954, pp. 337–346. Springer, Heidelberg (2011)

[28] Mirjalili, Seyedali, and Andrew Lewis. "The Whale Optimization Algorithm."Advances in Engineering Software 95 (2016): 51-67.

[29] Saidala, Ravi Kumar, and Nagaraju Devarakonda. "Improved Whale Optimization Algorithm Case Study: Clinical Data of Anaemic Pregnant Woman." Data Engineering and Intelligent Computing. Springer, Singapore, 2018. 271-281.

[30] Accepted: R. K. Saidala, N. Devarakonda. ―A New Parallel Metaheuristic Optimization Algorithm and It‘s Application in CDM‖, Proceedings of the I2CT IEEE conference 2017, Pune.

[31] Ravi Kumar Saidala, Nagaraju Devarakonda "Improved Whale Optimization Algorithm using Clonal Selection Algorithm for finding optimal structures of data" Under review Information Sciences, Elsevier.

[32] Yevgeniy Bodyanskiy, Olena Vynokurova, Volodymyr Savvo, Tatiana Tverdokhlib, Pavlo Mulesa,"Hybrid Clustering-Classification Neural Network in the Medical Diagnostics of the Reactive Arthritis", International Journal of Intelligent Systems and Applications, Vol.8, No.8, pp.1-9, 2016.

Authors' Profiles

Ravi Kumar Saidala is a research scholar. He is pursuing his Ph.D. in Acharya Nagarjuna University (ANU). He completed his M. Tech in the year 2013 from Computer Science & Engineering with the specialization of Digital Image Processing from ANU. He has worked as a guest faculty in the CSE department, ANU. He has received his B. Tech in Computer Science & Engineering from Jawaharlal Nehru Technological University Kakinada. His research interests are Data Mining, Soft Computing, Digital Image Processing and Pattern Recognition.

Dr. D. Naga Raju is the Professor and HOD of IT Department in Lakireddy Balireddy College of Engineering, Mylavaram. He was awarded his Ph.D in Computer Science & Engineering from the Jawaharlal Nehru Technological University Hyderabad in the year 2014. He completed his master degree M. Tech (CSE) from Jawaharlal Nehru University (JNU), New Delhi in the year 2005. He completed his B. Tech (CSE) from Sri Venkateswara University, Tirupati in the year 2002. He has 16 years of teaching experience and published papers in various International journals, National and International conferences. His research interests are Data Mining, Soft Computing, Machine Learning and Pattern Recognition

How to cite this paper: Ravi Kumar Saidala, Nagaraju Devarakonda, "Multi-Swarm Whale Optimization Algorithm for Data Clustering Problems using Multiple Cooperative Strategies", International Journal of Intelligent Systems and Applications(IJISA), Vol.10, No.8, pp.36-53, 2018. DOI: 10.5815/ijisa.2018.08.04