

Comparative Study of Krill Herd, Firefly and Cuckoo Search Algorithms for Unimodal and Multimodal Optimization

Gobind Preet Singh^{*1}, Abhay Singh²

Received 14th October 2013, Accepted 23th December 2013

Abstract: Today, in computer science, a computational challenge exists in finding a globally optimized solution from an enormously large search space. Various metaheuristic methods can be used for finding the solution in a large search space. These methods can be explained as iterative search processes that efficiently perform the exploration and exploitation in the solution space. In this context, three such nature inspired metaheuristic algorithms namely Krill Herd Algorithm (KH), Firefly Algorithm (FA) and Cuckoo search Algorithm (CS) can be used to find optimal solutions of various mathematical optimization problems. In this paper, the proposed algorithms were used to find the optimal solution of fifteen unimodal and multimodal benchmark test functions commonly used in the field of optimization and then compare their performances on the basis of efficiency, convergence, time and conclude that for both unimodal and multimodal optimization Cuckoo Search Algorithm via Lévy flight has outperformed others and for multimodal optimization Krill Herd algorithm is superior than Firefly algorithm but for unimodal optimization Firefly is superior than Krill Herd algorithm.

Keywords: Metaheuristic Algorithm, Krill Herd Algorithm, Firefly Algorithm, Cuckoo Search Algorithm, Unimodal Optimization, Multimodal Optimization.

1. Introduction

In recent times, nature inspired metaheuristic algorithms are being widely used for solving optimization problems, including NP-hard problems which might need exponential computation time to solve in worst case scenario. In metaheuristics methods [1, 9] we might compromise on finding an optimal solutions just for the sake of getting good solutions in a specific period of time. The main aim of metaheuristic algorithms are to quickly find solution to a problem, this solution may not be the best of all possible solutions to the problem but still they stand valid as they do not require excessively long time to be solved. Two crucial characteristics of metaheuristic algorithms are intensification and diversification. The intensification searches around the current best solution and selects the best candidate or solution. The diversification ensures that the algorithm explores the search space more efficiently. Maintaining balance between diversification and intensification is important because firstly we have to quickly find the regions in the large search space with high quality solutions and secondly not to waste too much time in regions of the search space which are either already explored or which do not provide high quality solutions[3].

In this paper, we have used three metaheuristic algorithms Krill Herd Algorithm (KH) [4], Firefly Algorithm (FA) [1] and Cuckoo Search Algorithm (CS) [5]. First is the Krill Herd Algorithm which was developed by Amir Hossein Gandomi and Amir Hossein Alavi in 2011. The KH algorithm is based on the simulation of the herding behaviour of Krill individuals. Second is the Firefly Algorithm (FA) which was developed by X.-S. Yang in 2007. It was inspired by the flashing pattern of fireflies. Third algorithm is

the Cuckoo search Algorithm which was developed by X.-S. Yang and S. Deb in 2009. It is based on the interesting breeding behaviour of certain species of cuckoos such as brood parasitism. This paper aims to provide the comparison study of the Krill Herd Algorithm (KH) with Firefly Algorithm (FA) and Cuckoo Search (CS) Algorithm via Lévy Flights against unimodal and multimodal test functions. Rest of the paper is organised as follows. First we will briefly explain the Krill Herd Algorithm, Firefly Algorithm, Cuckoo Search Algorithms and several Mathematical benchmark functions in section (2). Then experimental settings and results will be shown in section (3) and then finally we will conclude the paper.

2. Nature Inspired Algorithms and Optimization

2.1. Krill Herd Algorithm

2.1.1. Krill Swarm's Herding Behavior

Many Research have been done in order to find the mechanism that lead to the development non- random formation of groups by various marine animals [11,12]. The significant mechanisms identified are feeding ability, protection from predators, enhanced reproduction and environmental condition [6].

Krills from Antarctic region are one of the best researched marine animals. One of the most significant ability of krills is that they can form large swarms [13, 14]. Yet there are number of uncertainties about the mechanism that lead to distribution of krill herd [15]. There are proposed conceptual models to explain observed formation of krill heard [16] and result obtained from those models states that krill swarms form the basic unit of organization for this species.

Whenever predators (Penguins, Sea Birds) attack krill swarms, they take individual krill which leads in reducing the krill density. After the attack by predators, formation of krill is a multi- objective process mainly including two Goals: (1) Increasing Krill density

¹ Guru Gobind Singh Indraprastha University – India

² Guru Gobind Singh Indraprastha University – India

* Corresponding Author: Email: gobind75@gmail.com

and (2) Reaching food. Attraction of Krill to increase density and finding food are used as objective function which finally lead the krills to herd around global minima. In this mechanism, all individual krill moves towards the best possible solution while searching for highest density and food.

2.1.2. Krill Herd Algorithm

As Predator remove individual from Krill swarm, the average krill density and distance of krill swarm from the food location decreases. We assume this process as the initialization phase in the Krill Herd Algorithm [4]. Value of objective function for each individual is supposed to be combination of distance from food and highest density of krill swarm. Three essential actions [6] considered by Krills to determine the time dependent position of an individual krill are:

- (i) Movement induced by other krill individuals
- (ii) Foraging activity
- (iii) Random Diffusion

We know that all the optimization algorithm should have searching capability in space of arbitrary dimensionality. Hence we generalize Lanrangian model of krill herding to n dimensional decision space.

$$dX_i/dt = N_i + F_i + D_i \quad (1)$$

Here N_i is the motion induced by other krill individuals, F_i is the foraging motion and D_i stands for physical diffusion for i th krill individuals.

2.1.3. Motion Induced by other krill individual

According to research krill individuals move due to the mutual effects by each other so as to maintain high density [6]. Movement for the krill individual is defined as:

$$N_i^{new} = N^{max} \alpha_i + \omega_n N_i^{old} \quad (2)$$

In Eq. (2) N^{max} stands for maximum induced speed which is equal to 0.01 (m/s) [6], α_i is the direction of motion induced which is estimated from target swarm density and local swarm density, ω_n is the inertia weight of the motion induced, N_i^{old} is the last motion induced.

$$\alpha_i = \alpha_i^{local} + \alpha_i^{target} \quad (3)$$

In Eq. (3) α_i^{local} is the local effect due to the neighbors, and α_i^{target} is the target direction effect due to the best krill individual. Attractive or Repulsive effect of the neighbors on an individual krill movement can be formulated as:

$$\alpha_i^{local} = \sum_{j=1}^{NN} \widehat{K}_{i,j} \widehat{X}_{i,j} \quad (4)$$

$$\widehat{X}_{i,j} = X_j - X_i / \|X_j - X_i\| + \varepsilon \quad (5)$$

$$\widehat{K}_{i,j} = K_i - K_j / K^{worst} - K^{best} \quad (6)$$

In Eq. (4) NN is the number of the neighbors. In Eq. (5) and Eq. (6) K^{worst} and K^{best} are, the worst and the best fitness values of the krill individuals till now, K_i represents the fitness value of the i th krill individual, K_j is the fitness of j th neighbor and X represents the related positions.

To choose the neighbor, using actual behavior of Krill individual, a sensing distance (d_s) is calculated using

$$d_{s,i} = 1/5N \sum_{j=1}^N \|X_i - X_j\| \quad (7)$$

In Eq. (7) $d_{s,i}$ is the sensing distance for the i th krill individual and N stands for number of krill individuals. Based on Eq. (7), two krill

individuals are neighbor if the distance between them is less than d_s .

The effect of the individual krill having the best fitness on the i th individual krill is calculated as follow:

$$\alpha_i^{target} = C^{best} \widehat{K}_{i,best} \widehat{X}_{i,best} \quad (8)$$

In Eq. (8) C^{best} is the effective coefficient of the krill with the best fitness and is defined as:

$$C^{best} = 2(rand + I/I_{max}) \quad (9)$$

Where rand is a random value in the range [0, 1], I is the actual iteration number, and I_{max} is the maximum number of iterations.

2.1.4. Foraging motion

Foraging motion is developed in terms of two main effective parameters. One is the food location and the second one is the previous experience about the food location. This motion can be explained for the i th krill individual as follow:

$$F_i = v_f \beta_i + \omega_f F_i^{old} \quad (10)$$

Where

$$\beta_i = \beta_i^{food} + \beta_i^{best} \quad (11)$$

In Eq. (10) v_f is the foraging speed, ω_f is the inertia weight of the foraging motion in the range [0, 1] and F_i^{old} is the last foraging motion. In Eq. (11) β_i^{food} and β_i^{best} are the food attractive effect and best fitness of the i th krill so far respectively. Measured values of the foraging speed [7] is $0.02(\underline{ms}^{-1})$.

The center of food is found at first and then food attraction is formulated. The virtual center of food concentration is estimated according to the fitness distribution of the krill individuals, which is inspired from ‘‘center of mass’’. The center of food for each iteration is formulated as:

$$X^{food} = \sum_{i=1}^N 1/K_i X_i / \sum_{i=1}^N 1/K_i \quad (12)$$

Hence, we can evaluate the food attraction for the i th krill individual by following equation:

$$\beta_i^{food} = C^{food} \widehat{K}_{i,food} \widehat{X}_{i,food} \quad (13)$$

In Eq. (13) C^{food} is the food coefficient. As time passes the effect of food in the krill herding decrease and food coefficient is evaluated as:

$$C^{food} = 2(1 - I/I_{max}) \quad (14)$$

The attraction towards food is defined to attract the krill swarm towards global optima. Based on this definition, the krill individuals normally herd around the global optima after some iteration. This can be considered as an efficient global optimization strategy which helps improving the global optima of the KH algorithm.

The effect of the best fitness of the i th krill individual is also handled using the following equation:

$$\beta_i^{best} = \widehat{K}_{i,ibest} \widehat{X}_{i,ibest} \quad (15)$$

In Eq. (15) K_{ibest} is the best previously visited position of the i th krill individual.

2.1.5. Physical diffusion

The physical diffusion of all the krills is basically a random process. We can express this motion in terms of maximum diffusion speed and a random directional vector. We can formulate

it as follows:

$$D_i = D^{max} \delta \quad (16)$$

In Eq. (16) D^{max} is the maximum diffusion speed [8] and δ is the random directional vector and it includes random values in range [-1, 1]. As krills position gets better, random motion is also reduced. Thus, another term is added to the physical diffusion formula to consider this effect. The effects of the motion induced by other krill individuals and foraging motion gradually decrease increase in iterations. The physical diffusion is a random vector hence it does not steadily reduces with increase in number of iterations due to which another term is added to Eq. (16). This term, linearly decreases the random speed with time and works on the basis of a geometrical annealing schedule:

$$D_i = D^{max}(1 - I/I_{max})\delta \quad (17)$$

2.1.6. Motion Process of the KH Algorithm

The defined motions frequently change the position of a krill individual toward the best fitness. The foraging motion and the motion induced by other krill individuals contain two global and two local strategies. KH a powerful algorithm as all these work in parallel. The formulations of these motions for the i th krill individual show that, if fitness value of each of the above mentioned effective factor like $K_j, K^{best}, K^{food}, K_i^{best}$ is better i.e. less than the fitness of the i th krill, it has an attractive effect else it is a repulsive effect. We can notice from the above formulations that better fitness has more effect on the movement of i th krill individual. The position vector of a krill individual during the interval t to $t + \Delta t$ is given by the following equation:

$$X_i(t + \Delta t) = X_i(t) + \Delta t dX_i/dt \quad (18)$$

Δt should be carefully set according to the optimization problem because this parameter works as a scale factor of the speed vector. Δt completely depends on the search space and it can be obtained simply by the following formula:

$$\Delta t = C_t \sum_{j=1}^{NV} (UB_j - LB_j) \quad (19)$$

In Eq. (19) NV is the total number of variables and LB_j and UB_j are lower and upper bounds of the j th variables respectively. It is empirically found that C_t is a constant Number between [0, 2]. Lower the values of C_t more carefully the krill individuals will search.

2.1.7. Crossover

To improve the performance of the algorithm, genetic reproduction mechanisms are incorporated into the algorithm. One such algorithm is crossover. Crossover is a genetic operator used to vary the programming of chromosomes from one generation to the next. In this Algorithm, an adaptive vectored crossover scheme is employed.

We can control crossover by a crossover probability, Cr , and actual crossover can be performed in two ways: (1) binomial and (2) exponential. The binomial scheme performs crossover on each of the d components or variables/parameters. By generating a uniformly distributed random number between 0 and 1, the m th component of $X_i, x_{i,m}$, is determined as:

$$x_{i,m} = \begin{cases} x_{r,m}, & rand_{i,m} < Cr \\ x_{i,m}, & else \end{cases} \quad (20)$$

$$Cr = 0.2 \bar{R}_{i,best} \quad (21)$$

In Eq. (20) $r \in \{1, 2, \dots, i-1, i+1, \dots, N\}$. With this new crossover

probability, the crossover probability for the global best is equal to zero and it increases with decrease in fitness.

2.2. Firefly Algorithm

2.2.1. Behavior and nature of Fireflies

Fireflies are the creatures that can generate light inside of it. Light production in fireflies is due to a type of chemical reaction. The primary purpose for firefly's flash is to act as a signal system to attract other fireflies. Although they have many mechanisms, the interesting issues are what they do for any communication to find food and to protect themselves from enemy hunters including their successful reproduction. There are around two thousand firefly species, and most of them produce short and rhythmic flashes. The pattern observed for these flashes is unique specific species. The rhythm of the flashes, rate of flashing and the amount of time for which the flashes are observed together forms a pattern that attracts both the males and females to each other. Females of a species respond to individual pattern of the male of the same species.

The light intensity at a particular distance from the light source follows the inverse square law. That is as the distance increases the light intensity decreases. Furthermore, the air absorbs light which becomes weaker and weaker as there is an increase of the distance. There are two combined factors that make most fireflies visible only to a limited distance that is usually good enough for fireflies to communicate each other. The flashing light can be formulated in such a way that it is associated with the objective function to be optimized. This makes it possible to formulate new metaheuristic algorithms.

2.2.2. Firefly algorithm

The firefly (FA) algorithm [1, 9, 10] is a metaheuristic algorithm, inspired by the flashing behavior of fireflies. The primary purpose for a firefly's flash is to act as a signal system to attract other fireflies.

Xin-She Yang formulated this firefly algorithm by taking three assumptions [1]

- i. All fireflies are unisexual, so that one firefly will be attracted to all other fireflies;
- ii. Attractiveness is proportional to their brightness, and for any two fireflies, the less bright one will be attracted by (and thus move to) the brighter one; however, the brightness can decrease as their distance increases;
- iii. If there are no fireflies brighter than a given firefly, it will move randomly.

2.2.3. Light Intensity and Attractiveness

Two core issues in firefly algorithm are (i) The variation of light intensity, (ii) The formulation of the attractiveness.

For simplicity, it is assumed that the attractiveness of a firefly is determined by its brightness which in turn is associated with the encoded objective function of the optimization problems. In the simplest case for maximum optimization problems, the brightness I of a firefly for a particular location x could be chosen as $I(x) \propto f(x)$. Even so, the attractiveness β is relative, it should be judged by the other fireflies. Thus, it will differ with the distance r_{ij} between firefly i and firefly j . In addition, light intensity decreases with the distance from its source, and light is also absorbed by the media, so we should allow the attractiveness to vary with the varying degree of absorption.

Since a firefly's attractiveness is proportional to the light intensity seen by adjacent fireflies, attractiveness β of a firefly can be defined as

$$\beta(r) = \beta_0 e^{-\gamma r^m}, (m \geq 1) \quad (22)$$

In Eq. (22), r or r_{ij} is the distance between the i th and j th fireflies. β_0 is the attractiveness at $r = 0$ and γ is a fixed light absorption coefficient. The distance between any two fireflies i th and j th at x_i and x_j is the Cartesian distance and can be calculated as:

$$r_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2} \quad (23)$$

In Eq. (23), x_{ik} is the k th component of the i th firefly (x_i). The movement of i th firefly, to another more attractive (brighter) j th firefly, is determined by

$$x_i = x_i + \beta_0 e^{-\gamma r_{ij}^2} (x_j - x_i) + \alpha(\text{rand} - 0.5) \quad (24)$$

In Eq. (3) the second term is due to the attraction while the third term is the randomization with α being the randomization parameter. Rand is a random number generator uniformly distributed in the range of $[0, 1]$. For most cases in the implementation, $\beta_0 = 1$ and $\alpha \in [0, 1]$. Furthermore, the randomization term can easily be extended to a normal distribution $N(0, 1)$ or other distributions.

2.3. Cuckoo Search Algorithm via Lévy Flight

2.3.1. Cuckoo's breeding behaviour

Cuckoo is an interesting bird species, known not only for the beautiful sound they make, but also for their aggressive reproductive strategy. Cuckoos are extremely diverse group of birds with regards to breeding system. Many Cuckoo species follow the strategy of brood parasitism by using host individuals either of the same or different species to raise the young of the own. Cuckoo species such as Anis and Guira lay their eggs in communal nest though they may remove others eggs to increase the survival probability of their own eggs. Some host birds can engage in direct conflict with the intruding cuckoos. On recognition of parasitic eggs, the host may kick the parasites eggs out, or build a new nest. Female parasitic cuckoos who are specialized in mimicry lay eggs that closely which resemble the eggs of their host which reduces the probability of their eggs being abandoned.

Parasitic cuckoos often choose a nest where host bird have just laid their eggs. The cuckoo egg hatches earlier as compared to the host's, and the cuckoo chick grows faster than them; In most cases the chick evicts the eggs laid by host species, which increases the cuckoo chick's share of food provided by its host bird. Some cuckoo chick can even replicate the call of host chicks to gain access to more feeding opportunity.

2.3.2. Lévy flight

A Lévy flight is a random walk in which the step-lengths have a probability distribution that is heavy-tailed. Research works have shown that flight behavior of many animals and insects demonstrated the typical characteristics of Lévy flights [17, 18, 19, 20]. Fruit flies or *Drosophila melanogaster*, explore their landscape using a series of straight flight paths punctuated by a sudden 90 degree turn, leading to a Lévy-flight-style intermittent scale free search pattern was shown in a study conducted by Reynolds and Frye. Many researches shows that Lévy flights interspersed with Brownian motion can describe the animals' hunting patterns [24, 25]. Even light can be related to Lévy flights [23]. Latterly, such behavior has been applied to optimization and optimal search, and preliminary results show its promising capability [18, 20, 21, 22].

2.3.3. Cuckoo Search Algorithm

Each egg in the nest represents solution, and Cuckoo egg represents new solution. The aim is to use the new and potentially better solutions (Cuckoos) to replace not-so-good or inferior

solution in the nests. In the simplest form, each nest has one egg. The algorithm [5] can be extended to more complicated cases in which each nest has multiple eggs representing a set of solutions.

Cuckoo search is based on three idealized rules which states that

- i. Each Cuckoo lays one egg, which represents a set of solution coordinates, at a time, and dumps it in a random nest.
- ii. A fraction of the nests containing the best eggs, or solutions, will be carried over to the next generation.
- iii. The number of nests is fixed and there is a probability that a host can discover an alien egg. If this happens, the host can either discard the egg or the nest and this result in building a new nest in a new location.

When generating new solutions x^{t+1} for the i th Cuckoo, Lévy Flight is performed.

$$x_i^t = x_i^{t+1} + \alpha \oplus \text{Levy}(\lambda) \quad (25)$$

In equation (25), $\alpha > 0$ is the step size which should be related to the scales of the problem of interest. In most cases, we can use $\alpha = O(L/10)$ where L is the characteristic scale of the problem of interest. The above equation is essentially the stochastic equation for a random walk. The product \oplus means entry wise multiplications.

The Lévy flight essentially provides a random walk whose random step length is drawn from a Lévy distribution

$$\text{Lévy} \sim u = t^{-\lambda}, (1 < \lambda < 3) \quad (26)$$

This has an infinite variance with an infinite mean. Here the steps essentially form a random walk process with a power-law step-length distribution with a heavy tail. Some of the new solutions should be generated by Lévy walk around the best solution obtained so far, this will speed up the local search. To ensure that the system will not be trapped in a local optimum, a substantial fraction of the new solutions should be generated by far field randomization whose locations should be far enough from the current best solution.

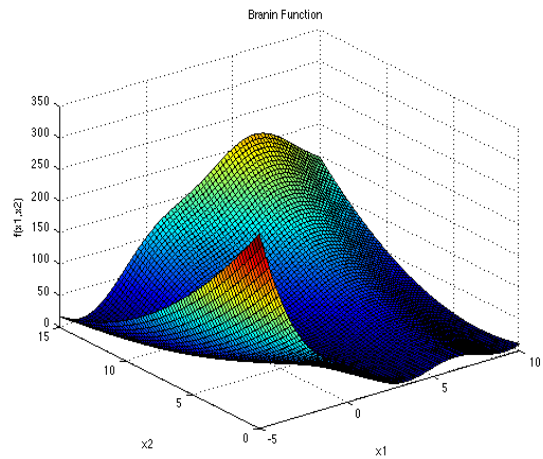
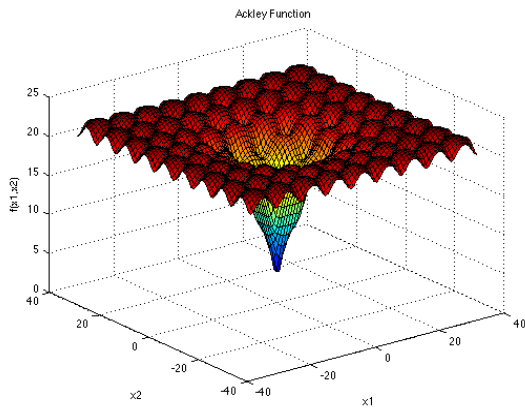
2.4. Testing Optimization Functions

In Literature [26] there are many benchmark test functions which are designed to test the performance of optimization algorithms. In this paper we will compare and validate above mentioned algorithms against these benchmark functions. Seventeen functions [27,28,29] including many multimodal functions are used in this paper in order to compare and verify efficiency and convergence of all three above mentioned Nature Inspired Algorithms. Certain test functions used in our simulations are as follows:

Ackley Function is multimodal function widely used for testing optimization algorithms.

$$f(x) = -20 \exp \left[-0.2 \sqrt{\frac{1}{d} \sum_{i=1}^d x_i^2} \right] - \exp \left[\frac{1}{d} \sum_{i=1}^d \cos(2\pi x_i) \right] + (20 + e)$$

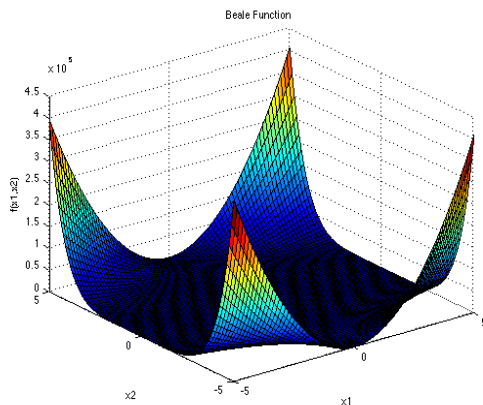
With a global minimum $f(x_*) = 0$ at $x_* = (0,0, \dots, 0)$ in the range of $x_i \in [-32.768, 32.768]$, for all $i = 1, 2, \dots, d$.



Beale function is 2- dimensional multimodal function, with sharp peaks at the corners of the input domain.

$$f(\mathbf{x}) = (1.5 - x_1 + x_1x_2)^2 + (2.25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2$$

Which has minimum $f(x_*) = 0$, at $x_* = (3, 0.5)$ in range $x_i \in [-4.5, 4.5]$, for all $i = 1, 2$.



Branin function is multimodal with three global minima. The recommended values of a, b, c, r, s and t are: a = 1, b = 5.1/(4π²), c = 5/π, r = 6, s = 10 and t = 1/(8π).

$$f(\mathbf{x}) = a(x_2 - bx_1^2 + cx_1 - r)^2 + s(1 - t)\cos(x_1) + s$$

With a global minimum $f(x_*) = 0.397887$ at $x_* = (-\pi, 12.275), (\pi, 2.275)$ and $(9.42478, 2.475)$ in the range $x_1 \in [-5, 10]$ and $x_2 \in [0, 15]$.

Colville is a unimodal test function.

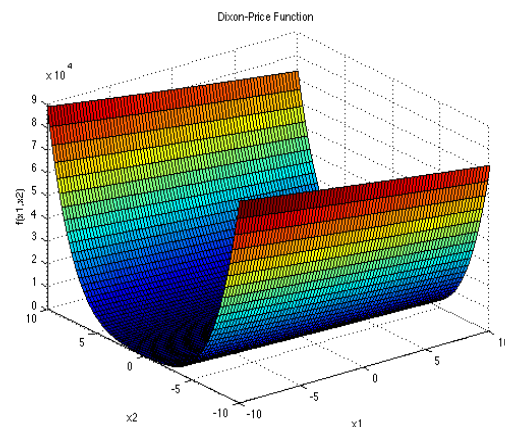
$$f(\mathbf{x}) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_3^2 - x_4)^2 + 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1)$$

Which has minimum $f(x_*) = 0$, at $x_* = (1, 1, 1, 1)$ in range $x_i \in [-10, 10]$, for all $i = 1, 2, 3, 4$.

DIXON-PRICE's unimodal test function

$$f(\mathbf{x}) = (x_1 - 1)^2 + \sum_{i=2}^d i (2x_i^2 - x_{i-1})^2$$

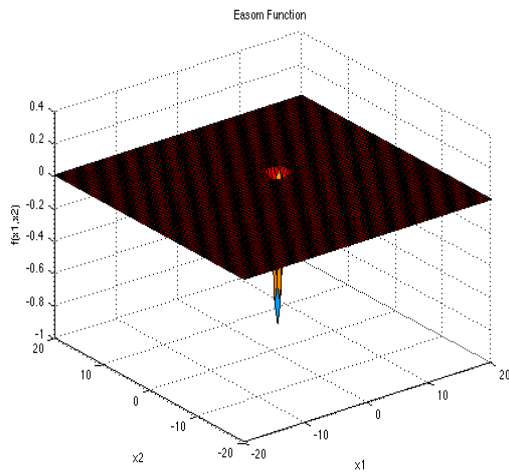
With a global minimum $f(x_*) = 0$ at $x_i = 2 \frac{2^i - 2}{2^i}$ bin the range of $x_i \in [-10, 10]$, for all $i = 1, 2, \dots, d$.



Easom function has several local minima. It is unimodal, and the global minimum has a small area relative to the search space.

$$f(\mathbf{x}) = -\cos(x_1) \cos(x_2) \exp(-(x_1 - \pi)^2 - (x_2 - \pi)^2)$$

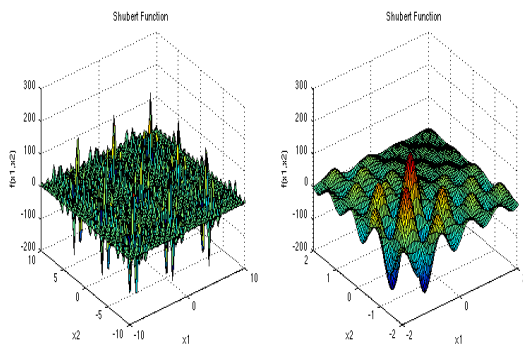
Which has minimum $f(x_*) = -1$, at $x_* = (\pi, \pi)$ in range $x_i \in [-100, 100]$, for all $i = 1, 2$.



Shubert function an multimodal test function has several local minima and many global minima. Its equation is

$$f(\mathbf{x}) = \left(\sum_{i=1}^5 i \cos((i+1)x_1 + i) \right) \left(\sum_{i=1}^5 i \cos((i+1)x_2 + i) \right)$$

Which has minimum $f(x_*) = -186.7309$, in range $x_i \in [-10, 10]$, for all $i = 1, 2$,

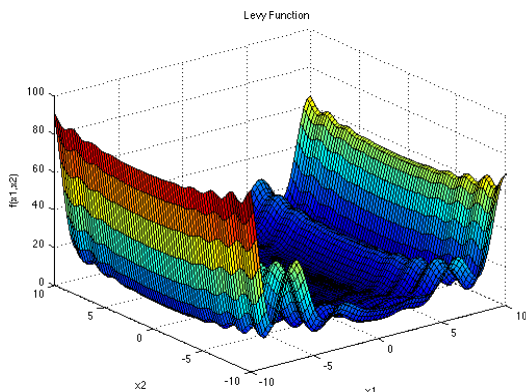


Levy Test function is multimodal function. Its equation is

$$f(\mathbf{x}) = \sin^2(\pi w_1) + \sum_{i=1}^{d-1} (w_i - 1)^2 [1 + 10 \sin^2(\pi w_1 + 1)] + (w_d - 1)^2 [1 + \sin^2(2\pi w_d)], \text{ where}$$

$$w_i = 1 + \frac{x_i - 1}{4}, \text{ for all } i = 1, \dots, d$$

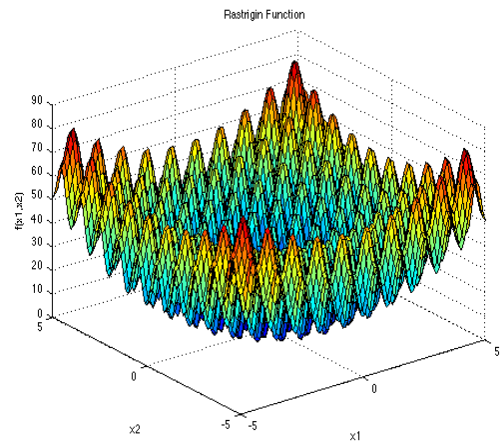
With a global minimum $f(x_*) = 0$ at $x_i = (1, \dots, 1)$ which is evaluated in the range of $x_i \in [-10, 10]$, for all $i = 1, 2, \dots, d$.



Rastrigin function has several local minima. It is highly multimodal, but locations of the minima are regularly distributed.

$$f(\mathbf{x}) = 10d + \sum_{i=1}^d [x_i^2 - 10 \cos(2\pi x_i)]$$

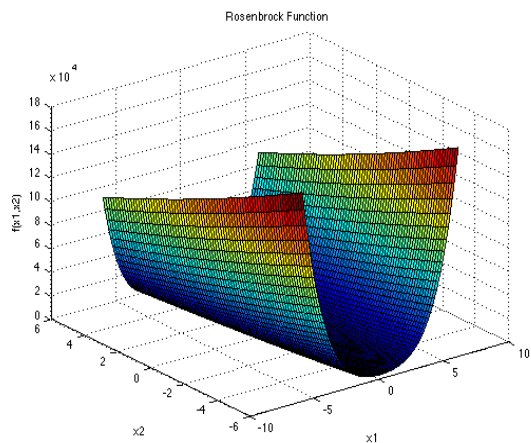
With a global minimum $f(x_*) = 0$ at $x_* = (0, 0, \dots, 0)$ evaluated in the range of $x_i \in [-5.12, 5.12]$, for all $i = 1, 2, \dots, d$.



Rosenbrock function is unimodal, and the global minimum lies in a narrow, parabolic valley.

$$f(\mathbf{x}) = \sum_{i=1}^{d-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$$

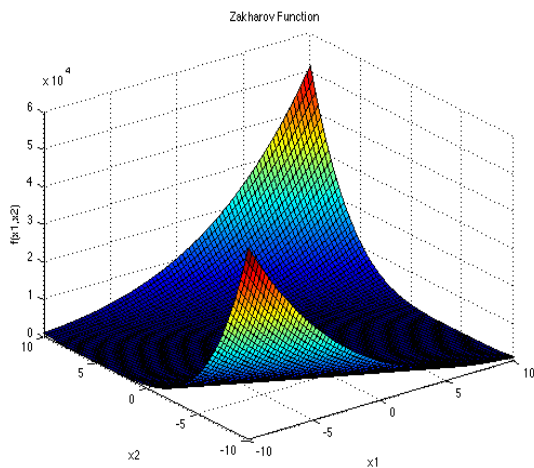
With a global minimum $f(x_*) = 0$ at $x_i = (1, \dots, 1)$ which is evaluated in the range of $x_i \in [-5, 10]$, for all $i = 1, 2, \dots, d$.



Zakharov function has no local minima except the global one. It's a unimodal function and its equation is

$$f(\mathbf{x}) = \sum_{i=1}^d x_i^2 + \left(\sum_{i=1}^d 0.5 i x_i \right)^2 + \left(\sum_{i=1}^d 0.5 i x_i \right)^4$$

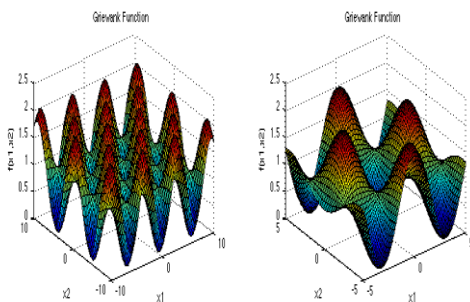
With a global minimum $f(x_*) = 0$ at $x_i = (0, \dots, 0)$ which is evaluated in the range of $x_i \in [-5, 10]$, for all $i = 1, 2, \dots, d$.



Griewank function has many widespread local minima, which are regularly distributed. It is a multimodal test function.

$$f(\mathbf{x}) = \sum_{i=1}^d \frac{x_i^2}{4000} - \prod_{i=1}^d \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

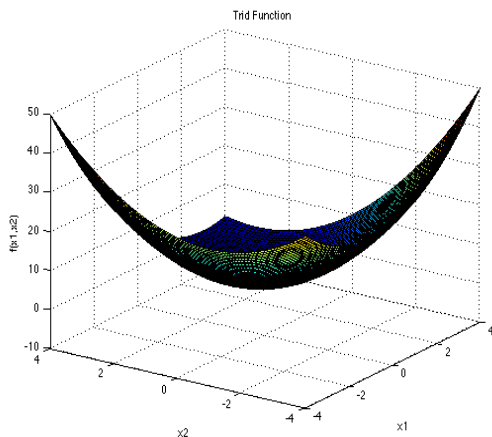
With a global minimum $f(x_*) = 0$ at $x_i = (0, \dots, 0)$ which is evaluated in the range of $x_i \in [-600, 600]$, for all $i = 1, \dots, d$.



Trid function has no local minimum except the global one. It is a unimodal function.

$$f(\mathbf{x}) = \sum_{i=1}^d (x_i - 1)^2 - \sum_{i=2}^d x_i x_{i-1}$$

With a global minimum $f(x_*) = -50$ for $d=6$ and $f(x_*) = -200$ at $d=10$ which is evaluated in the range of $x_i \in [-d^2, d^2]$, for all $i = 1, 2, \dots, d$.



Powell Function is a unimodal function

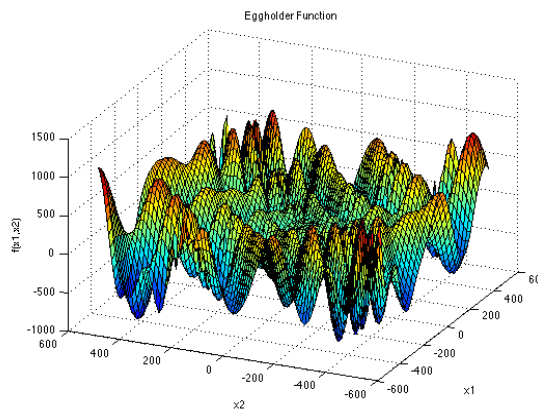
$$f(\mathbf{x}) = \sum_{i=1}^{d/4} [(x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4]$$

This function is usually evaluated on the region $x_i \in [-4, 5]$, for all $i = 1, \dots, d$. having minima $f(x_*) = 0$ at $x_i = (3, -1, 0, 1, \dots, 3, -1, 0)$

Eggholder function is a difficult function to optimize, because of the large number of local minima. It is multimodal test function.

$$f(\mathbf{x}) = -(x_2 + 47) \sin\left(\sqrt{\left|x_2 + \frac{x_1}{2} + 47\right|}\right) - x_1 \sin\left(\sqrt{|x_1 - (x_2 + 47)|}\right)$$

Which has minimum $f(x_*) = -959.6407$, at $x_* = (512, 404.2319)$ in range $x_i \in [-512, 512]$, for all $i = 1, 2$.



3. Implementation and Numerical Experiments

In this section we will compare the performance of Krill Herd algorithm, Firefly algorithm, Cuckoo search algorithm for various benchmark test functions. The benchmarks function include both unimodal and multimodal with both low and high dimensional problems are described in Section 2.4 and for evaluation all computational procedures described above has been implemented in MATLAB™ computer program. In order to compare these algorithms we have carried out extensive simulations and each algorithm has been run 50 times so as to carry out meaningful analysis. The maximum number of function evaluations is set as 10,000 for high dimensional functions and 1000 for low dimensional functions.

Here for Krill Herd Algorithm C_t is set to 0.5 and the inertia weights (ω_n, ω_f) are equal to 0.9 at beginning of search and it linearly decreases to 0.1 at the end. For Firefly algorithm certain constants are fixed as $\alpha = 0.5, \beta = 0.2$ and $\gamma = 1$ for simulation. For cuckoo search algorithm probability for host bird is fixed as $p_\alpha = 0.25$. For simulation we have used various population sizes from $n = 10$ to 150, and found that for most problems, it is sufficient to use $n = 15$ to 50. Therefore, we have used a fixed population size of $n = 50$ in all our simulations for comparison. Now we will divide this section in two parts comprising comparison of algorithms for unimodal test function in first section and comparison of algorithms for multimodal test functions in another. For both the section we will be comparing KH algorithm, FA Algorithm and CS Algorithm via Lévy Flights on the basis of three criteria i.e. Optimization fitness (efficiency), Convergence and processing time.

3.1. Optimization for Unimodal Test Function

In this section we have compared KH algorithm, FA Algorithm and CS Algorithm via Lévy Flights on Eight Unimodal benchmark functions popular for optimization. Unimodal functions are those function which have only single local minima and these function easier to optimize.

3.1.1. Optimization fitness

Here we have calculated the mean fitness value using the above mentioned algorithms for all unimodal test functions mentioned above. Optimized fitness result where global optima is reached are summarized in Table 1.

Table 1. Comparison of Optimization Fitness for Unimodal Test Functions

Function/ Algorithms	KH Algorithm	FFA Algorithm	CS Algorithm
Dixon-Price(d=20)	0.6975	0.6668	0.6667
Rosenbrock (d=20)	31.89	14.49	2.63e-14
Zakharov(d=20)	5.973	3.19e-06	2.77e-28
Powell(d=20)	0.0299	0.002269	5.27e-09
Trid(d=20)	9.34e+2	1.52e+03	1.52e+03
Beale	2.92e-11	5.49e-10	1.34e-30
Easom	-0.96	-0.86	-1
Colville	-1.40e+09	-1.56e+07	-7.57e+17

Here we can see that Cuckoo Search algorithm has outperformed both Krill Herd and Firefly algorithm. For all Unimodal test function, fitness value of CS algorithm is much closer to global optima as compared to other two algorithms. But if we just compare the other two algorithms i.e. Firefly and Krill herd algorithm, their result are very close to each other, but on average, results obtained using Firefly algorithm are slightly better than results obtained using Krill Herd algorithm. As per results from Table 1 it is visible that performance of Krill Herd algorithm is better for low dimensional functions and as we move from low dimensional function to high dimensional functions fitness value for krill herd decreases i.e. distance from global minima increases. According to results in Table 1 performance of Firefly algorithm is better than Krill Herd algorithm for high dimensional functions but for low dimensional function result using Krill Herd algorithm are better than Firefly algorithm. Also in Krill Herd algorithm we have varied dimensions from d= 5 to 20 and observed that as we increases the dimensions, fitness value for function decreases.

3.1.2. Processing time

Here we will compare above mentioned algorithms on basis on processing time. Processing time is basically time consumed by algorithm to process single simulation. It includes time consumed by fixed number of iteration to solve the problem.

Table 2. Comparison of processing time in seconds for Unimodal Test Functions

Function/ Algorithms	KH Algorithm	FFA Algorithm	CS Algorithm
Dixon-Price(d=20)	141.73	125.56	19.69
Rosenbrock (d=20)	137.18	123.44	21.58
Zakharov (d=20)	137.66	125.66	22.92
Powell(d=20)	172.44	138.95	63.35
Trid(d=20)	155.12	146.04	30.72
Beale	13.034	11.18	1.64
Easom	12.69	12.13	2.46
Colville	13.33	12.92	2.59

From Table 2 it is quite easily visible that time taken or consumed by Cuckoo search algorithm is much less than the other two algorithms We can also compute from the Table 2 that time consumed by Firefly algorithm is less than Krill Herd Algorithm although difference is not much, In term of processing time Cuckoo search algorithm again outperform other two algorithms.

3.1.3. Convergence

In this section convergence plots of the benchmark functions for three different algorithm i.e. Krill Herd, Firefly, Cuckoo Search are compared for fixed number of iteration i.e. 10,000 iterations for high dimensional function and 1000 iteration for low dimensional function. Here we will estimate which algorithm gives potentially better and quicker convergence towards optimality. Below In Figure 1-8 Convergence Graph is plotted for all above mentioned Unimodal benchmark functions.

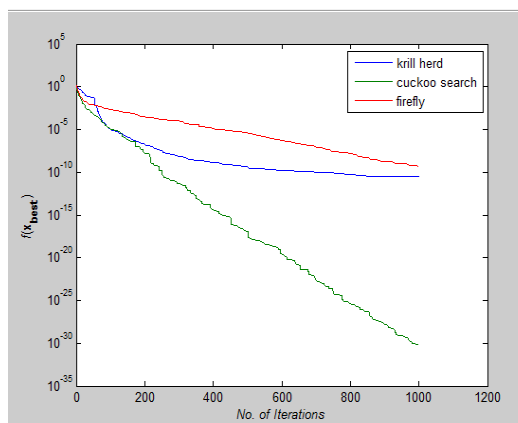


Fig. 1. Convergence Plot for Beale Function

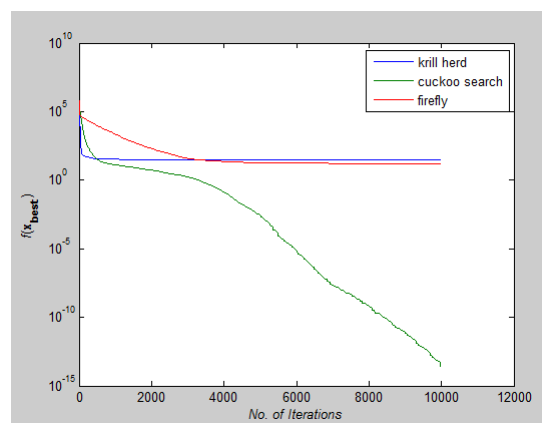


Fig. 2. Convergence Plot for Rosenbrock Function

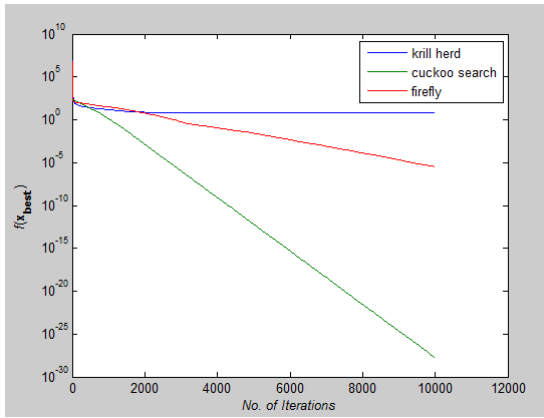


Fig. 3. Convergence Plot for Zakharov Function

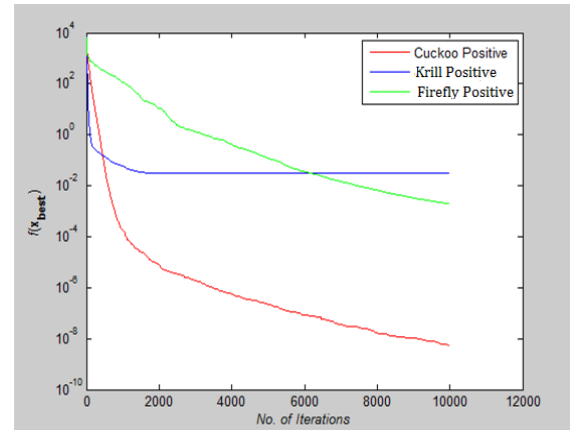


Fig. 7. Convergence Plot for Powell Function

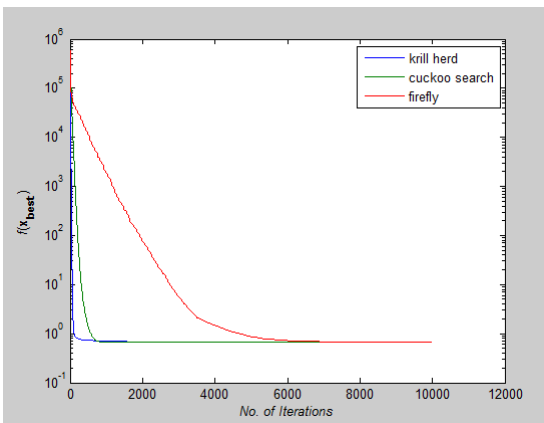


Fig. 4. Convergence Plot for Dixon-Price Function

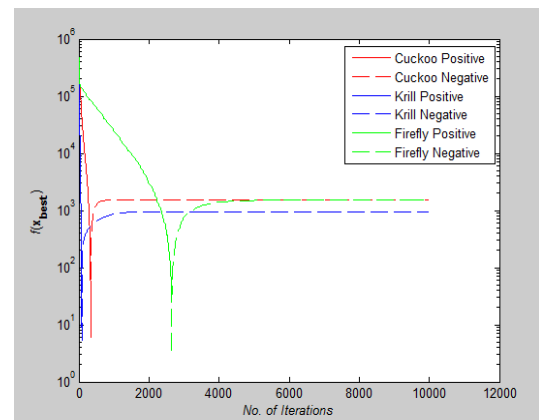


Fig. 8. Convergence Plot for Trid Function

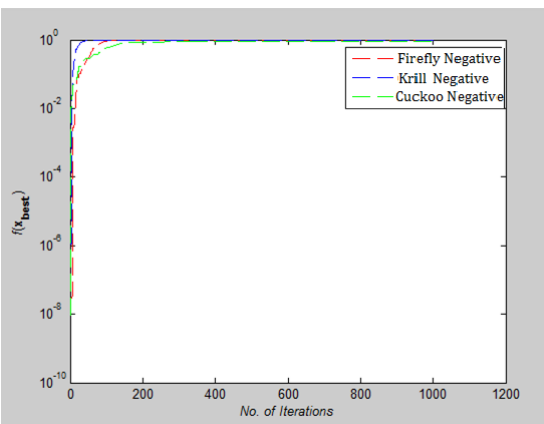


Fig. 5. Convergence Plot for Easom Function

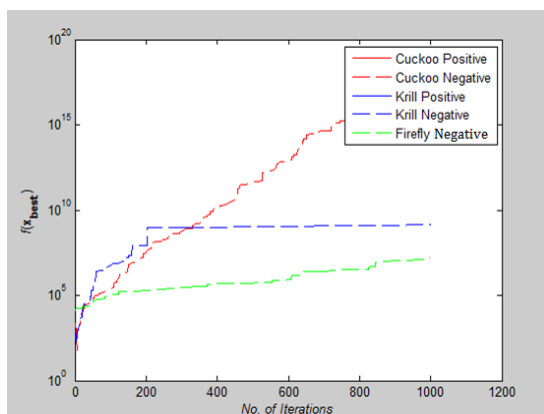


Fig. 6. Convergence Plot for Colville Function

In Fig.5, Fig.6 and Fig.8 we have plotted the absolute value of the fitness function. For Trid, Easom, Colville function, value of global minima is negative and so to plot them on convergence graph we took their absolute vale.

From Fig 1-8 we can interpret that the algorithm that quickly converges to its optimal solution is Krill Herd Algorithm. When we compare them on the number of iteration, Krill Herd Algorithm takes least number of iteration to converge whereas solution of other two algorithm are better in terms of fitness value. It can also be seen from Fig. 1-8 that for most of the test functions, the other two algorithms i.e. Firefly Algorithm and Cuckoo search algorithm do not converge till 10,000 iterations and for functions for which these two algorithm converges before 10,000 iterations, it is the cuckoo search algorithm which converges quickly than firefly for more of function as compare to number of function for which firefly algorithm converges faster than cuckoo search algorithm.

3.2. Optimization for Multimodal Test Function

In this section we have compared KH algorithm, FA Algorithm and CS Algorithm via Lévy Flights on Seven Multimodal benchmark functions popular for optimization. Multimodal functions are those function which have many number local minima and these function comparatively more difficult to optimize.

3.2.1. Optimization fitness

Here we have calculated the mean fitness value using the above mentioned algorithms for all multimodal test functions mentioned in Section 2.4. Optimized fitness result where global optima is reached are summarized in Table 3.

Table 3. Comparison of Optimization Fitness for Multimodal Test Functions

Function/ Algorithms	KH Algorithm	FFA Algorithm	CS Algorithm
Ackley($d=20$)	1.19e-05	7.36e-03	4.44e-15
Levy($d=20$)	0.066	2.28e-06	1.14e-06
Rastrigin ($d=20$)	8.69	20.36	2.84
Griewank ($d=20$)	1.37e-09	5.89e-04	0
Branin	0.3979	0.3981	0.3981
Shubert	-186.7309	-186.7309	-186.7309
Eggholder	-893.0205	-930.2513	-959.6407

From Table 3 we can see that Cuckoo Search algorithm has outperformed both Krill Herd and Firefly algorithm in multimodal optimization as well. For all Multimodal test function, fitness value of CS algorithm is much closer to global optima as compared to other two algorithms. But if we together compare the other two algorithms i.e. Firefly and Krill herd algorithm for multimodal test functions, result for them are in contradiction with results in previous section for unimodal test functions. In multimodal optimization for most of the low dimensional function ,result for both the algorithms are almost equivalent, for most of the time both are able to find the global optima but for high dimensional multimodal functions fitness value obtained using Krill Herd is better than fitness value obtained using Firefly Algorithm. Results in Table 3 are in contradiction with results in Table 1 as in unimodal test function optimization fitness for high dimensional function using firefly was better than krill herd but in multimodal function optimization fitness using krill herd algorithm is better than firefly algorithm for high dimensional function.

3.2.2. Processing time

In this section we will compare Krill Herd, Firefly, Cuckoo Search algorithms on basis on processing time for multimodal optimization functions. Processing time is basically time consumed by algorithm to process single simulation. It includes time consumed by fixed number of iteration to solve the problem.

Table 4. Comparison of processing time in seconds for Multimodal Test Functions

Function/ Algorithms	KH Algorithm	FFA Algorithm	CS Algorithm
Ackley($d=20$)	138.06	123.49	24.78
Levy($d=20$)	169.37	129.03	38.86
Rastrigin($d=20$)	139.55	126.31	24.58
Griewank($d=20$)	137.18	123.44	21.58
Branin	13.36	11.28	1.91
Shubert	13.83	11.33	2.56
Eggholder	13.04	11.23	2.83

Results in Table 4 are quite similar with the results in Table 2. In multimodal optimization function as well, time taken or consumed by Cuckoo search algorithm is much less than the other two algorithms It can also be interpreted from the Table 4 that time consumed by Firefly algorithm is less than Krill Herd Algorithm although difference is not much, In term of processing time Cuckoo search algorithm again outperform other two algorithms.

3.2.3. Convergence

Here convergence plots of the benchmark functions for three

different algorithm i.e. Krill Herd, Firefly, Cuckoo Search are compared for fixed number of iteration i.e. 10,000 iterations for high dimensional function and 1000 iteration for low dimensional functions. Here we will estimate which algorithm gives potentially better and quicker convergence towards optimality. Convergence Graph for all above mentioned Multimodal benchmark functions are plotted in Figure 9-15.

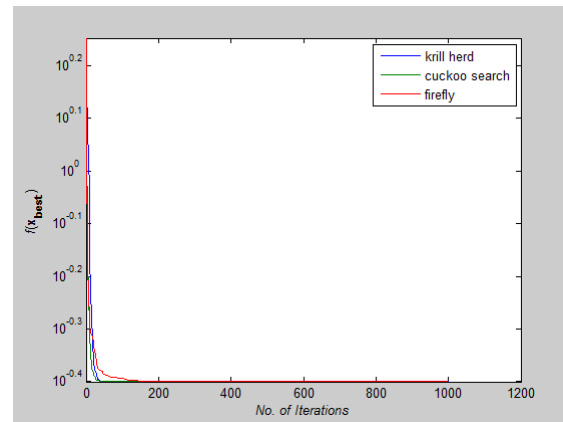


Fig. 9. Convergence Plot for Branin Function

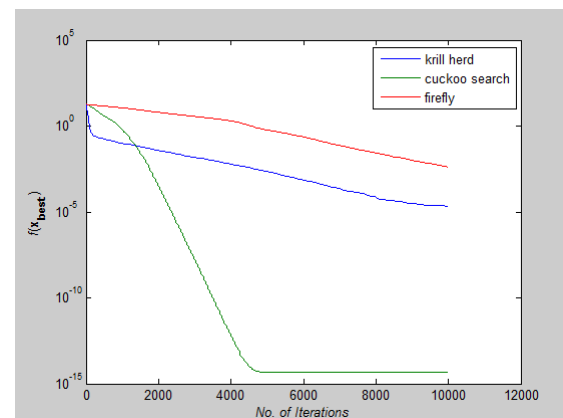


Fig. 10. Convergence Plot for Ackley Function

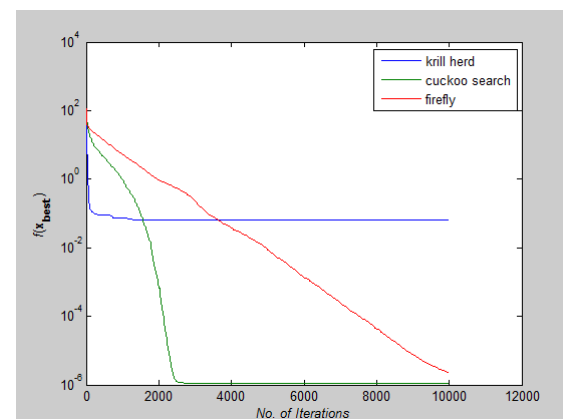


Fig. 11. Convergence Plot for Levy function

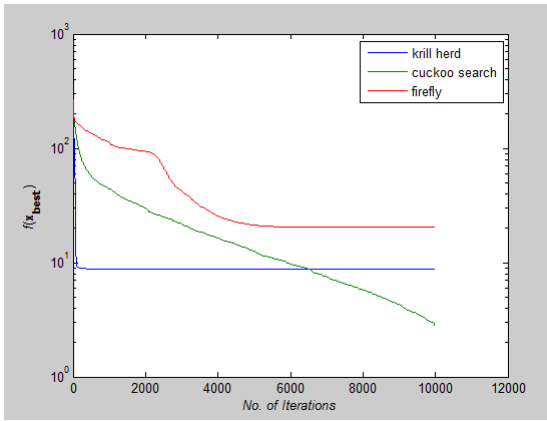


Fig.12 . Convergence Plot for Rastrigin Function

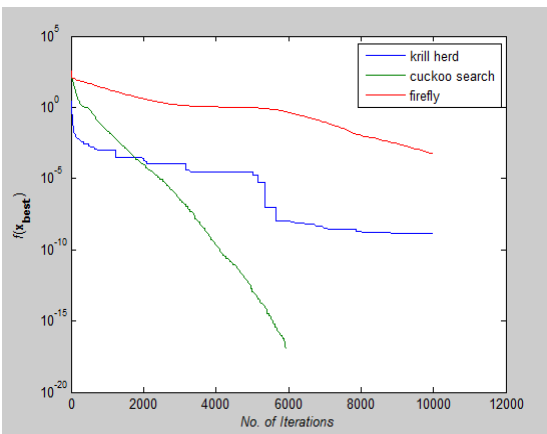


Fig. 13. Convergence Plot for Griewank Function

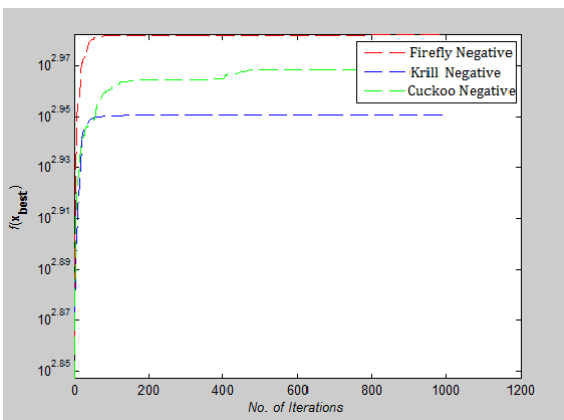


Fig. 14. Convergence Plot for Eggholder Function

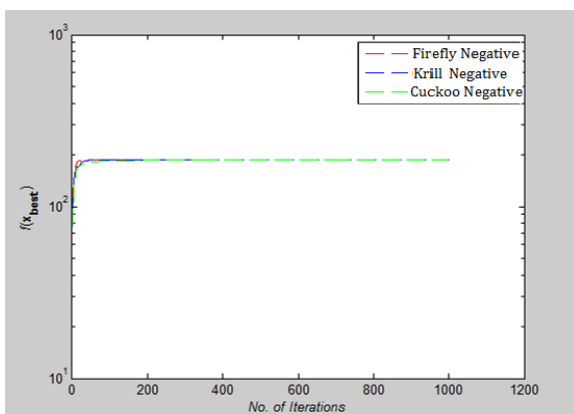


Fig. 15. Convergence Plot for Shubert Function

In Fig.14, and Fig.15 we have plotted the absolute value of the

fitness function. In multimodal optimization Shubert and Eggholder function value of global minima is negative and so to plot them on convergence graph we took their absolute value.

From Fig 9-15 we can interpret that although for many function all the algorithms are not able to converge before 10,000th iteration but for test functions like Rastrigin, Branin, Griewank, Levy and Eggholder, Krill herd algorithm is the fastest to converge to its optimal solution. When we compare them on the number of iteration, Krill Herd Algorithm takes least number of iteration to converge. If we see Fig.10 it is visible that for Ackley function Cuckoo Search Algorithm is the fastest to converge to its optimal solution others are not able to converge before 10,000th iteration. Also in Fig. 11 Cuckoo Search Algorithm is second fastest after Krill Herd algorithm to converge. From Fig 9-15 we can interpret that Firefly algorithm don not converge to its optimal solution for any of the high dimensional functions till 10,000th iteration. For multidimensional functions it is the Krill herd algorithm which is fastest to converge to its optimal solution, then after Krill Herd algorithm it is cuckoo search algorithm to converge to its optimal solution and at last is Firefly algorithm.

4. Conclusion

In this paper we have compared latest metaheuristic algorithms such as Krill Herd Algorithm, Firefly Algorithm and Cuckoo Search algorithm via Lévy Flights on basis of three criteria i.e. optimization fitness (efficiency), time processing and convergence. Results obtained by simulation of these algorithms on unimodal and multimodal test functions shows that Cuckoo search algorithm is superior for both unimodal and multimodal test function in terms of optimization fitness and time processing whereas when comparison comes down to line between Krill Herd Algorithm and Firefly Algorithm, KH Algorithm is superior than FFA algorithm for multimodal optimization of both high and low dimensional functions whereas for unimodal optimization FFA algorithm is superior than KH Algorithm for High dimensional function but KH algorithm is superior for low dimensional functions but in terms of time processing FFA Algorithm is surpasses KH Algorithm for both unimodal and multi modal optimization. When we compare these algorithms on basis of convergence Krill Herd is the fastest of all to converge to its optimal solution after which comes the Cuckoo search algorithm and at last comes the Firefly algorithm which do not converge for most of the function to it optimal solution.

During simulation we also noticed that as we increase the dimension, fitness value or optimization fitness for KH algorithm decreases although it outperformed FFA algorithm for high dimensional multimodal function but on average as dimension increases optimization fitness for KH algorithm decreases.

As all these powerful optimization strategy are able to optimize both unimodal and multimodal test function effectively hence we can easily extend them to study multi objective optimization applications with various constraints and even to NP-hard problems.

References

- [1] X. S. Yang, "Nature-Inspired Metaheuristic Algorithms", Luniver Press, 2008.
- [2] Christian Blum, Maria Jos'e Blesa Aguilera, Andrea Roli, Michael Sampels, Hybrid Metaheuristics, An Emerging Approach to Optimization, Springer, 2008 .
- [3] Christian Blum, and Maria Jos'e Blesa Aguilera. Metaheuristics in Combinatorial Optimization: Overview

and Conceptual Comparison, Springer, 2008.

- [4] Amir Hossein Gandomi, Amir Hossein Alavi. Krill herd: A new bio-inspired optimization algorithm, Elsevier, 2012.
- [5] X.-S. Yang, S. Deb, "Cuckoo search via Lévy flights", in: Proc. Of World Congress on Nature & Biologically Inspired Computing (NaBIC 2009), December 2009, India. IEEE Publications, USA, pp. 210-214 (2009).
- [6] Hofmann EE, Haskell AGE, Klinck JM, Lascara CM. Lagrangian modelling studies of Antarctic krill (*Euphausia superba*) swarm formation. *ICES J Mar Sci* 2004;61:617–31.
- [7] Price HJ. Swimming behavior of krill in response to algal patches: a mesocosm study. *Limnol Oceanogr* 1989;34:649–59.
- [8] Morin A, Okubo A, Kawasaki K. Acoustic data analysis and models of krill spatial distribution. Scientific Committee for the Conservation of Antarctic Marine Living Resources, Selected Scientific Papers, Part I; 1988. p.311–29.
- [9] Sh. M. Farahani, A. A. Abshouri, B. Nasiri, and M. R. Meybodi, "A Gaussian Firefly Algorithm", *International Journal of Machine Learning and Computing*, Vol. 1, No. December 2011.
- [10] Xin-She Yang, Chaos-Enhanced Firefly Algorithm with Automatic Parameter Tuning, *International Journal of Swarm Intelligence Research*, December 2011.
- [11] Flierl G, Grunbaum D, Levin S, Olson D. From individuals to aggregations: the interplay between behavior and physics. *J Theor Biol* 1999;196:397–454.
- [12] Okubo A. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. *Adv Biophys* 1986;22:1–94.
- [13] Hardy AC, Gunther ER. The plankton of the South Georgia whaling grounds and adjacent waters, 1926–1927. *Disc Rep* 1935;11:1–456.
- [14] Marr JWS. The natural history and geography of the Antarctic krill (*Euphausia superba* Dana). *Disc Rep* 1962;32:33–464.
- [15] Nicol S. Living krill, zooplankton and experimental investigations. Proceedings of the international workshop on understanding living krill for improved management and stock assessment marine and freshwater behaviour and physiology 2003;36(4):191–205.
- [16] Murphy EJ, Morris DJ, Watkins JL, Priddle J. Scales of interaction between Antarctic krill and the environment. In: Sahrhage D, editor. Antarctic Ocean and resources variability. Berlin: Springer-Verlag; 1988. p. 120–30.
- [17] Brown C., Liebovitch L. S., Glendon R., Lévy flights in Dobe Ju/'hoansi foraging patterns, *Human Ecol.*, 35, 129-138 (2007).
- [18] Pavlyukevich I., Lévy flights, non-local search and simulated annealing, *J. Computational Physics*, 226, 1830-1844 (2007).
- [19] Pavlyukevich I., Cooling down Lévy flights, *J. Phys. A:Math. Theor.*, 40, 12299-12313 (2007).
- [20] Reynolds A. M. and Frye M. A., Free-flight odor tracking in *Drosophila* is consistent with an optimal intermittent scale-free search, *PLoS One*, 2, e354 (2007).
- [21] Shlesinger M. F., Zaslavsky G. M. and Frisch U. (Eds), *Lévy Flights and Related Topics in Physics*, Springer, (1995).
- [22] Shlesinger M. F., Search research, *Nature*, 443, 281- 282 (2006).
- [23] Barthelemy P., Bertolotti J., Wiersma D. S., A Lévy flight for light, *Nature*, 453, 495-498 (2008).
- [24] Viswanathan, G. M. et al. Optimizing the success of random searches. *Nature* 401, 911–914(1999)
- [25] Bartumeus, F. et al. Optimizing the encounter rate in biological interactions: Lévy versus Brownian strategies. *Phys. Rev. Lett.* 88, 097901 (2002)
- [26] Yao X, Liu Y, Lin G. Evolutionary programming made faster. *IEEE Trans Evolut Comput* 1999;3:82–102.
- [27] www.sfu.ca/~ssurjano/optimization.html
- [28] www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar_files/TestGO_files/Pa ge364.htm
- [29] MominJamil,Xin-SheYang, Hans-Ju`rgenZepernick. " Test Functions for Global Optimization: A Comprehensive Survey" in *Swarm Intelligence and BioInspired Optimization*, Elsevier, Part I; 2013.p. 193-222