Applied Soft Computing 46 (2016) 230–245

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

An improved krill herd algorithm with global exploration capability for solving numerical function optimization problems and its application to data clustering

R. Jensi*, G. Wiselin Jiji

Dr. Sivanthi Aditanar College of Engineering, Tiruchendur, Tamilnadu, India

ARTICLE INFO

Article history:
Received 28 December 2015
Received in revised form 7 April 2016
Accepted 20 April 2016
Available online 10 May 2016

Keywords:
Global optimization
High convergence
Global exploration
Metaheuristic
Data clustering

ABSTRACT

Krill herd algorithm is a stochastic nature-inspired algorithm for solving optimization problems. The performance of krill herd algorithm is degraded by poor exploitation capability. In this study, we propose an improved krill herd algorithm (IKH) by making the krill the global search capability. The enhancement comprises of adding global search operator for exploration around the defined search region and thus the krill individuals move towards the global best solution. The elitism strategy is also applied to maintain the best krill during the krill update steps. The proposed method is tested on a set of twenty six well-known benchmark functions and is compared with thirteen popular optimization algorithms, including original KH algorithm. The experimental results show that the proposed method produced very accurate results than KH and other compared algorithms and is more robust. In addition, the proposed method has high convergence rate. The high performance of the proposed algorithm is then employed for data clustering problems and is tested using six real datasets available from UCI machine learning laboratory. The experimental results thus show that the proposed algorithm is well suited for solving even data clustering problems.

1. Introduction

Over the last few decades, many nature-inspired algorithms have been proposed for solving numerical optimization problems. Nature-inspired algorithms [3–5] play a vital role in solving many engineering optimization problems [6–10,23–25,32] owing to the global exploration and exploitation ability. These algorithms imitate the behavior of living things in nature such as animals, birds, fishes, etc. Several heuristic algorithms have been developed in the literature. Genetic algorithm (GA) [17] was proposed by Goldberg in 1998 simulating the survival of fittest among individuals in the population over many generations. Evolutionary strategy (ES) [16] and differential evolution (DE) [13,14] algorithms belong to the sub class of evolutionary algorithms which use selection, mutation and recombination operators. Population-based Incremental Learning (PBIL) [18] developed by Shumeet Baluja in 1994, is an optimization algorithm which combines the genetic algorithm with simple competitive learning. Particle Swarm Optimization (PSO) [19] was proposed by Eberhart and Kennedy in 1995, imitates the food foraging behavior of bird flocks or fish school. Ant Colony Optimization (ACO) [12] was initially developed by Marco Dorigo, simulates the behavior of ants searching for a path between food source and colony. Artificial bee colony (ABC) [11] is another optimization algorithm developed by Karaboga in 2005 based on the foraging behavior of honey bee swarm. Biogeography-based optimization (BBO) [15] is developed by Simon in 2008, is imitating the migrating behavior of species between islands. StudGA (SGA) [20], a variant of GA, is an optimization algorithm in which fittest individual is selected rather than stochastic selection and shares this information among others using genetic algorithm operators. Cuckoo search (CS) [21] which is a recently proposed optimization algorithm by Yang et al. in 2009, inspired by the obligate brood parasitic behavior of cuckoo species. Firefly algorithm (FFA) [22] is developed by Yang in 2010 based on the flashing behaviour of fireflies. In Refs. [2,35] authors proposed a new PSO algorithm combined with levy flight for solving optimization problems. Krill herd (KH) [1] is introduced by Gandomi and Alavi in recent times based on the imitation of krill individuals behavior. Even though there exists several optimization algorithms, research is still going on in the development of optimization algorithm which will provide high convergence rate and global optimum solution.

* Corresponding author.
E-mail address: r_jensi@yahoo.co.in (R. Jensi).

http://dx.doi.org/10.1016/j.asoc.2016.04.026
1568-4946/

The remaining section of the paper is organized as follows. Section 2 provides the variants of krill herd algorithm in solving function optimization problems. Section 3 briefly explains the original krill herd algorithm, and Section 4 presents the proposed IKH approach. Section 5 provides the experimental results, Section 6 provides the experimental results of data clustering and Section 7 is concluded with future discussion.

2. Variants of krill herd algorithm

In Ref. [1], the authors tested four different krill herd algorithms such as KH without any genetic operators, KH with crossover operator, KH with mutation operator and KH with crossover and mutation operators. They concluded that KH with crossover operator performed well for solving the optimization problems. Hence, hereafter KH refers to KH with crossover operator. Krill herd would fail to reach the global optima due to the poor exploitation capability. In order to overcome the problem of KH, the researchers developed novel variants of KH method. Krill herd (KH) is hybridized with other methods by utilizing the advantages of all the methods employed.

Wang et al. [28] employed an updated genetic reproduction mechanism such as stud selection and crossover operator into KH instead of stochastic selection while updating krill individual position. The algorithm is named as Stud krill herd because the operators are derived from Stud GA. G.G. Wang et al. [29] introduced an effective biogeography-based krill herd (BBKH) in which migration operator from BBO is combined with KH. In addition, Wang et al. [30] proposed a hybrid differential evolution KH method in which differential evolution operator is hybridized with KH during krill individual position updating process. Wang et al. [31] improved KH performance by including levy flight method into KH resulting Levy-flight krill herd algorithm. Wang et al. [37] developed a hybrid metaheuritic algorithm called HS/KH in which Harmony search is combined with KH in order to perform exploration and exploitation effectively. Wang et al. [38] introduced a Levy flight distribution and elitism strategy for updating KH motion calculation. Wang et al. [39] proposed a hybridized algorithm FKH where operators used in firefly algorithm is combined into KH. Li et al. [43] proposed an improved KH algorithm with linear decreasing step. In Ref. [44] authors introduced opposition based learning strategy and free search operator into KH algorithm in order to avoid the diversity problem. Wang et al. [45] developed a hybrid metaheuristic cuckoo search and krill herd (CSKH) algorithm where operators from cuckoo search are combined with KH to enhance its effectiveness and reliability. Wang et al. [46] proposed a hybrid simulated annealing-based krill herd algorithm for solving global optimization problems.

3. Krill herd (KH) algorithm

Krill herd (KH) [1] is a new metaheuristic population based global optimization algorithm. The inspiration of KH algorithm is the herding behaviour of krill swarm when looking for food and communication with each other. The implementation of KH method is based on three movements such as:

(i) Movement influenced by other krill individual.
(ii) Foraging action.
(iii) Physical diffusion.

KH approach follows Lagrangian model for effective search and it is described as:

$$\frac{dX_i}{dt} = N_i + F_i + D_i \tag{1}$$

where N_i is the movement induced by other krill individuals, F_i is the foraging action and D_i is the random physical diffusion of the ith krill individuals.

The direction of motion induced, $\propto_i$, depends on the three components, namely local swarm density, a target swarm density and a repulsive swarm density. The movement of a krill individual N_i is defined as:

$$N_i^{\text{new}} = N^{\max}\propto_i + \omega_n N_i^{\text{old}} \tag{2}$$

where

$$\propto_i = \propto_i^{\text{local}} + \propto_i^{\text{target}} \tag{3}$$

and $N^{\max}$ is the maximum induced speed, ω_n is the inertia weight, N_i^{old} is the motion induced previously, $\propto_i^{\text{local}}$ is the local effect offered by neighbours and $\propto_i^{\text{target}}$ is the best krill individual's target effect.

The second movement of KH approach foraging action F_i depends on two parameters, namely current food location and information about previous food location. The i[th] krill individual's motion is described as:

$$F_i = V_f\beta_i + \omega_f F_i^{\text{old}} \tag{4}$$

where

$$\beta_i = \beta_i^{\text{food}} + \beta_i^{\text{best}} \tag{5}$$

and V_f is the foraging speed, ω_f is the inertia weight of the foraging action, F_i^{old} is the previous foraging motion, β_i^{food} is the food attractive and β_i^{best} is the best fitness found by the ith krill so far. The value for ω_n, ω_f is equal to 0.9 at the first iteration and decreases gradually to 0.1 at the end of the iteration.

The third movement of KH approach is random physical diffusion. The physical diffusion of the ith krill individual depends on two components, namely maximum diffusion speed and a random directional vector and it is defined as:

$$D_i = D^{\max}\left(1 - \frac{I}{I_{\max}}\right)\delta \tag{6}$$

where $D^{\max}$ is the maximum diffusion speed, δ is the random vector in the range [−1, 1], I is the current generation and $I_{\max}$ is the maximum generation.

Based on the three movements defined above, the position of ith krill individual during the time interval is:

$$X_i(t+\Delta t) = X_i(t) + \Delta t\frac{dX_i}{dt} \tag{7}$$

It is clearly seen that Δt is an important parameter and its value determines the convergence speed. For more details refer [1].

4. The proposed IKH algorithm

Sometimes KH gets stuck in local optima due to the poor exploration and exploitation. To improve the performance of KH, in this study we incorporate global search towards the best solution into KH so that the proposed algorithm reaches promising region within the defined boundary. In the proposed method, original KH algorithm is performed to move the krill into new position. After exploration stage, global search is carried out with a small step to achieve global best solution. Here the step size is determined by using the dimension of the problem. Step size is calculated using Eq. (8):

$$\text{stepsize} = \begin{cases} \dfrac{\exp(-\text{nd})}{\exp(-2)}, \text{if nd} > 10 \\[2ex] \dfrac{1}{\exp(-0.5)}, \text{otherwise} \end{cases} \tag{8}$$