

Article

A New Swarm Intelligence Approach for Clustering Based on Krill Herd with Elitism Strategy

Zhi-Yong Li ¹, Jiao-Hong Yi ² and Gai-Ge Wang ^{3,*}

¹ School of Digital Creation and Animation, Shenzhen Polytechnic, Shenzhen 518055, China; E-Mail: lizhiyong@szpt.edu.cn

² School of Environmental Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China; E-Mail: yijiaohong@163.com

³ School of Computer Science and Technology, Jiangsu Normal University, Xuzhou 221116, China

* Author to whom correspondence should be addressed; E-Mail: gaigewang@163.com; Tel.: +86-138-5200-6092.

Academic Editor: Javier Del Ser Lorente

Received: 4 August 2015 / Accepted: 15 October 2015 / Published: 22 October 2015

Abstract: As one of the most popular and well-recognized clustering methods, fuzzy C-means (FCM) clustering algorithm is the basis of other fuzzy clustering analysis methods in theory and application respects. However, FCM algorithm is essentially a local search optimization algorithm. Therefore, sometimes, it may fail to find the global optimum. For the purpose of getting over the disadvantages of FCM algorithm, a new version of the krill herd (KH) algorithm with elitism strategy, called KHE, is proposed to solve the clustering problem. Elitism strategy has a strong ability of preventing the krill population from degrading. In addition, the well-selected parameters are used in the KHE method instead of originating from nature. Through an array of simulation experiments, the results show that the KHE is indeed a good choice for solving general benchmark problems and fuzzy clustering analyses.

Keywords: clustering; krill herd; elitism strategy; swarm intelligence

1. Introduction

Currently, fuzzy clustering is one of the important research branches in many fields, such as knowledge discovery, image processing, machine learning, and pattern recognition. With the expansion of scope of the study, more accurate clustering results are required from various aspects in scientific and practical application. Fuzzy C-Means (FCM) clustering is one of the most popular and well-recognized clustering methods. This method uses the concept of the geometric closeness of data points in Euclidean space. It allocates these data to different clustering, and the distance between these clusters is then determined. The FCM clustering algorithm is the basis of other fuzzy clustering analysis methods in theory and application respects, and it is therefore most widely-used among various clustering algorithms. However, the FCM algorithm is essentially a local search optimization algorithm. Herein, if its initial value is selected improperly, it will converge to a local minimum. Therefore, this drawback limits the FCM algorithm to be used in many applications.

Aiming at the disadvantages of the FCM algorithm, researchers have proposed several methods to improve its performance. Apart from the traditional methods, recently, various metaheuristic algorithms have been proposed by the inspiration of nature and successfully addressed all kinds of application problems, such as grey wolf optimizer (GWO) [1,2], genetic algorithm (GA) [3], biogeography-based optimization (BBO) [4–6], animal migration optimization (AMO) [7], gravitational search algorithm (GSA) [8–10], cuckoo search (CS) [11–15], stud genetic algorithm (SGA) [16], wolf search algorithm (WSA) [17], multi-verse optimizer (MVO) [18], dragonfly algorithm (DA) [19], moth-flame optimization (MFO) [20], earthworm optimization algorithm (EWA) [21], harmony search (HS) [22,23], firefly algorithm (FA) [24–26], particle swarm optimization (PSO) [27–29], monarch butterfly optimization (MBO) [30], ant colony optimization (ACO) [31], bat algorithm (BA) [32–36], differential evolution (DE) [37–40], and interior search algorithm (ISA) [41]. Among them, swarm-based metaheuristic search, so called swarm intelligence methods, are one of the most well-known paradigms in nature-inspired algorithms. Due to its remarkable performance, they have dealt with a variety of applications, such as reliability [42,43], knapsack problems [44], quantitative interpretation [45], scheduling [46], path planning [47], parameter estimation [48], global numerical optimization [49–51], neural network training [52,53] and feature selection [54]. The KH method that is inspired by the krill herding behavior of krill in sea was first proposed by Gandomi and Alavi in 2012 [55,56]. Since it has high stability and strong robustness when solving optimization problems, many researchers have made in-depth studies about it and various improved version of KH methods have been proposed [57–60]. The main differences between the KH algorithm and other swarm intelligence algorithms is that the parameters used in the KH algorithm is fully originated from the real krill herd in nature. Here, a new version of KH algorithm with elitism strategy, called KHE, is proposed. Elitism tragedy can prevent the krill population from degrading. In KHE, the well-selected parameters are used instead of originating from nature [61,62]. Furthermore, the KHE method is applied to solve clustering problem for the purpose of escaping a local minimum. Moreover, with the aim of showing the performance of KHE method, it is compared with six other metaheuristic algorithms through seven complicated benchmark problems. The results show that the KHE method is able to find the better function values on given benchmark problems than six others; KHE is also a good choice of implementing fuzzy clustering analyses.

Section 2 provides a basic knowledge of FCM clustering algorithm. Section 3 reviews the optimization process of KH, and then a framework of KHE method is given. This is followed by the usage of KHE method to solve the clustering problem. With the aim of the showing the performance of the KHE method, several simulation results comparing KHE with other methods for general benchmark functions and clustering are presented in Section 4. The discussion and future work orientation can be provided in Section 5.

2. Fuzzy C-Means (FCM) Clustering Algorithm

Let $X = \{x_1, x_2, \dots, x_n\}$ be n data samples; c ($2 \leq c \leq n$) is the number of the divided categories for these data samples; $\{A_1, A_2, \dots, A_c\}$ indicates that the corresponding c categories, and \mathbf{U} is their similarity classification matrix, whose cluster centers are $\{v_1, v_2, \dots, v_c\}$; $\mu_k(x_i)$ is the membership degree of x_i in the category A_k (abbreviated as μ_{ik}). The objective function J_b can be expressed as follows:

$$J_b(\mathbf{U}, \mathbf{v}) = \sum_{i=1}^n \sum_{k=1}^c (\mu_{ik})^b (d_{ik})^2 \tag{1}$$

where d_{ik} is the Euclidean distance that is used to measure distance between the i -th sample x_i and the center point of the k -th category. It can be calculated as follows:

$$d_{ik} = d(x_i - v_k) = \sqrt{\sum_{j=1}^m (x_{ij} - v_{kj})^2} \tag{2}$$

where m is the number of characteristics of the data sample; b is the weighting parameter and its range is $1 \leq b \leq \infty$. The FCM clustering algorithm is to find an optimal classification, so that the classification is able to produce the smallest function value J_b . It is required that the sum of the values of membership degree for a sample in terms of each cluster is 1. That is to say, it can be described as

$$\sum_{j=1}^c \mu_j(x_i) = 1, i = 1, 2, \dots, n \tag{3}$$

As stated before, μ_{ik} is the membership degree of x_i in the category A_k , and it can be updated as

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{b-1}}} \tag{4}$$

Subsequently, all the cluster centers $\{v_i\}$ are calculated as

$$v_{ij} = \frac{\sum_{k=1}^n (\mu_{ik})^b x_{kj}}{\sum_{k=1}^n (\mu_{ik})^b} \tag{5}$$

here, we suppose $I_k = \{i \mid 2 \leq c < n; d_{ik} = 0\}$. For all of category i , $i \in I_k$, $\mu_{ik} = 0$.

The updating process mentioned above is repeated by Equations (4) and (5) until the method converges. When the algorithm converge, in theory, various cluster centers for each sample and the membership degree in terms of each category are obtained at this time, thus fuzzy clustering partition

has been done by now. Although FCM has a high search speed, it is essentially a local search algorithm, and is therefore very sensitive to initial cluster centers. If the cluster centers have the initial poor choice, it will converge to a local minimum.

3. KHE Method for Clustering Problem

3.1. KH Method

Krill herd (KH) [55] is a novel swarm intelligence method for solving optimization problems. It is the simplification and idealization of the herding of the krill swarms in sea. The position of an individual krill is determined by three motions as:

- (i) movement induced by other krill individuals;
- (ii) foraging action; and
- (iii) random diffusion

In KH, the Lagrangian model is used in a d -dimensional decision space as shown in Equation (6).

$$\frac{dX_i}{dt} \approx N_i + F_i + D_i \tag{6}$$

where N_i is the motion induced by other krill individuals; F_i is the foraging motion, and D_i is the physical diffusion of the i -th krill individuals.

3.1.1. Motion Induced by Other Krill Individuals

The direction of motion induced, α_i , is approximately evaluated by the target effect, a local effect, and a repulsive effect. For krill i , it can be defined as:

$$N_i^{new} = N^{max} \alpha_i + \omega_n N_i^{old} \tag{7}$$

where

$$\alpha_i = \alpha_i^{local} + \alpha_i^{target} \tag{8}$$

and N^{max} is the maximum induced speed, ω_n is the inertia weight of the motion induced, N_i^{old} is the last motion induced, α_i^{local} is the local effect provided by the neighbors and α_i^{target} is the target direction effect provided by the best krill individual.

3.1.2. Foraging Motion

The foraging motion is influenced by the two main factors: The previous and current food location. For the i -th krill individual, this motion can be expressed as follows:

$$F_i = V_f \beta_i + \omega_f F_i^{old} \tag{9}$$

where

$$\beta_i = \beta_i^{food} + \beta_i^{best} \tag{10}$$

and V_f is the foraging speed, ω_f is the inertia weight of the foraging, F_i^{old} is the last foraging motion, β_i^{food} is the food attractiveness and β_i^{best} is the effect of the best fitness of the i -th krill so far.

3.1.3. Random Diffusion

This motion can be expressed in terms of a maximum diffusion speed and a random directional vector. It can be formulated as follows:

$$D_i = D^{\max} \delta \quad (11)$$

where D_{\max} is the maximum diffusion speed, and δ is the random directional vector.

Based on the above motions, the position of a krill individual from t to $t + \Delta t$ is given by the following equation:

$$X_i(t + \Delta t) = X_i(t) + \Delta t \frac{dX_i}{dt} \quad (12)$$

It should be noted that Δt is a constant that can be determined by problem of interest. More details about the KH algorithm can be found in [55].

3.2. KH Method with Elitism Strategy (KHE)

As stated before, the KH method can always include the best krill individual in the population. However, the positions of all the krill individuals in the population will be updated during the optimization process regardless of its good and bad. When the best one is being updated, there is a probability of worsening the best one. If this happens, the whole population will deteriorate so that it may lead to slow convergence.

With the aim of preventing the krill population degrade, an elitism strategy is incorporated into the basic KH method. That is, in our current work, a new version of the KH method with elitism strategy (abbreviated as KHE) is proposed. In KHE method, certain best krill individuals are memorized, and then all the krill are updated by three motions. Finally, certain worst krill individuals in the new population will be replaced by the memorized best ones in the last generation. Elitism strategy can forbid the best ones being destroyed by three krill motions, and can guarantee the population can always proceed to the better status. Limited by the length of the paper, the more detailed process of elitism strategy can be referred to in [4,63].

3.3. KHE Method for Clustering Problem

The clustering problem is essentially an optimization problem. Therefore, clustering problem can be solved by the KHE method. As per Sections 2, 3.1, and 3.2, the optimization process of KHE method for clustering problem can be simply represented as follows:

- (1) Initialize the control parameters. All the parameters used in KHE are firstly initialized.
- (2) Randomly initialize c cluster centers, and generate the initial population, calculate membership degree of each cluster center for all samples by Equation (4), and the fitness of each krill individual value f_i , where $i = 1, 2, \dots, NP$. Here, NP is the number of population size.
- (3) Set $t = 0$.

- (4) Save the *KEEP* best krill individuals as *BEST*.
- (5) Implement three motions and update the positions of krill individuals in population.
- (6) Replace the *KEEP* worst krill individuals with the *KEEP* best krill individuals saved in *BEST*.
- (7) Calculate c clustering centers, membership degree and fitness for each individual.
- (8) If the $t < Maxgen$, $t = t + 1$, go to Equation (4); Otherwise, the algorithm is over and finds the final global optimal solution.

Based on the above steps, a brief presentation of KHE for clustering problem is shown in Figure 1.

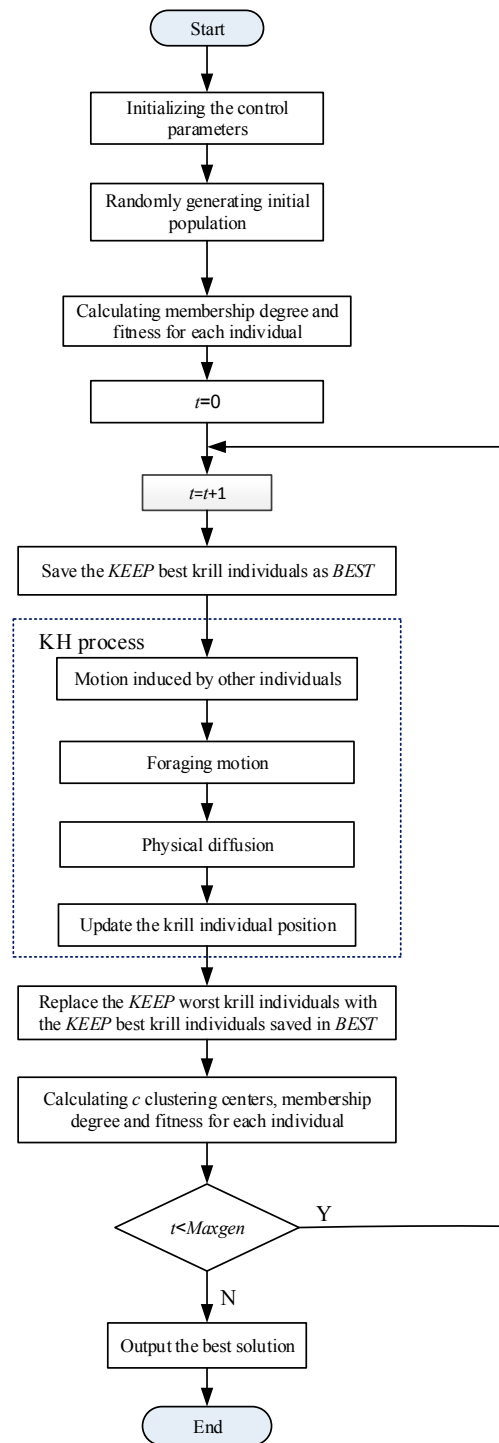


Figure 1. Flowchart of FCM algorithm by using the KHE method.

4. Simulation Results

In this section, after function evaluation through an array of experiments conducted in benchmark functions (see Table 1), the clustering problem is dealt with by the KHE method. More detailed descriptions of all the benchmarks can be referred to in [4,64,65]. Note that the dimensions of functions are thirty. In order to obtain fair results, all the implementations are conducted under the same conditions as shown in [59].

Table 1. Benchmark functions.

No.	Name	Definition
F01	Dixon & Price	$f(\vec{x}) = (x_1 - 1)^2 + \sum_{i=2}^n i(2x_i^2 - x_{i-1})^2$
F02	Griewank	$f(\vec{x}) = \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$
F03	Holzman 2 function	$f(\vec{x}) = \sum_{i=1}^n ix_i^4$
F04	Powell	$f(\vec{x}) = \sum_{i=1}^{n/4} (x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4$
F05	Quartic with noise	$f(\vec{x}) = \sum_{i=1}^n (i \cdot x_i^4 + U(0,1))$
F06	Rosenbrock	$f(\vec{x}) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$
F07	Sphere	$f(\vec{x}) = \sum_{i=1}^n x_i^2$

The parametric study about KH has been done in [61]. The parameters for KHE method are the same as [61], which are set as follows: $V_f = 0.02$, $D^{\max} = 0.005$ and $N^{\max} = 0.01$. For the parameters used in the other methods, their settings can be referred to in [4,63].

In order to remove the influence of the randomness and get the relatively representative statistical results, 200 implementations have been done independently on each benchmark. The population size and maximum generation number are set to 50 in the experiments conducted in Section 4.1. In the following experiments, the optimal solution for each test problem is bolded.

4.1. Convergent Performance Compared KHE with Six Other Methods

The performance of KHE was compared with basic KH and other five optimization methods (ACO [31], GA [3], HS [22,23], PSO [27,66] and SGA [16]) on seven optimization problems (see Table 2).

From Table 2, it can be seen that, for the best, mean and worst function values, KHE has the best performance on all the seven benchmarks on average. For other methods, their obtained function values are similar. Carefully looking at Table 2, generally speaking, SGA has the relatively better final optimization values than the other five methods. The results in Table 2 indicate that the KHE method is the proper strategy for most optimization problems.

Table 2. Mean, best and worst function values obtained by different methods.

	Function	ACO	GA	HS	KH	KHE	PSO	SGA
MEAN	F01	3.36E5	1.25E5	8.26E5	1.55E5	18.90	2.77E5	8.64E3
	F02	32.55	106.40	403.60	67.46	1.13	172.70	29.50
	F03	8.72E4	3.51E4	2.07E5	3.74E4	1.86	8.36E4	2.29E3
	F04	5.92E3	1.88E3	6.06E3	3.70E3	36.01	2.47E3	182.80
	F05	17.75	7.92	54.68	10.14	4.44E−4	13.77	0.66
	F06	5.47E3	1.86E3	3.99E3	1.22E3	31.60	1.38E3	313.10
	F07	85.19	21.57	119.50	20.03	0.04	50.05	11.02
BEST	F01	1.14E5	1.71E4	3.13E5	5.31E4	3.37	2.44E4	1.41E3
	F02	14.90	33.08	266.10	35.84	1.02	91.07	9.64
	F03	2.36E4	6.30E3	9.39E4	1.76E4	0.03	1.16E4	300.60
	F04	2.43E3	388.10	2.26E3	1.00E3	2.41	1.01E3	52.56
	F05	6.12	1.28	25.52	5.10	5.13E−6	4.22	0.08
	F06	3.73E3	513.10	2.20E3	697.30	28.19	508.00	137.50
	F07	55.53	5.80	70.20	10.44	3.84E−3	29.45	4.16
WORST	F01	8.46E5	3.63E5	1.26E6	2.85E5	167.60	2.39E6	4.90E4
	F02	69.45	235.90	498.70	101.70	1.63	568.30	68.65
	F03	1.76E5	1.35E5	3.29E5	6.26E4	20.86	6.00E5	8.98E3
	F04	8.89E3	4.42E3	1.07E4	6.35E3	218.40	4.72E3	558.50
	F05	37.48	28.87	81.76	17.74	7.82E−3	32.10	5.08
	F06	8.08E3	4.14E3	5.70E3	1.90E3	46.15	2.88E3	688.00
	F07	126.40	42.72	143.40	33.13	0.31	65.62	21.11

4.2. Clustering Problem Compared KHE with Seven Other Methods

As stated before, a clustering problem is essentially an optimization problem, so it can be solved by the KHE method. Here, KHE is compared with pure FCM and the other five metaheuristic methods including the basic KH method. The dataset used in this paper is the same with the data in [67]. The data have four-hundred data samples, and its characteristic dimension is two. Now, we will divide these data samples into four categories. Therefore, each krill contains eight elements. Population size and maximum generation number are set to 16 and 25, respectively. For other algorithms, their parameter settings are the same as Section 4.1. Figure 2 is the clustering results by pure FCM clustering algorithm when its final objective function value is 3.620176. Figure 3 shows the optimization process of KHE method for clustering problem. From Figure 3, we can see, KHE has a fast convergent speed for clustering problem. Figure 4 is the clustering results by KHE algorithm when its final objective function value is 3.303485. From Figures 2 and 4, we can see that, KHE method can obtain more accurate clustering results than pure FCM. More results can be recorded in Table 3. From Table 3, on average, the KHE method has the most accurate clustering results, and both SGA and KHE have the optimal performance for the best clustering results. For the worst performance, all the methods except FCM have the similar clustering results that are significantly better than pure FCM. For standard deviation (STD), KHE has the second performance that is only inferior to HS. From Table 3 and Figures 2–4, it can be seen that, the KHE method can solve the clustering problem better than other comparative methods in most cases.

It should be pointed out that, each run may generate completely different results. This is because the clustering results are dependent on the initial clustering centers.

Table 3. Optimization results for the fuzzy C-means (FCM) problem.

	ACO	FCM	GA	HS	KH	KHE	PSO	SGA
MEAN	3.303556	3.368558	3.303527	3.303536	3.303624	3.303510	3.303542	3.303523
BEST	3.303474	3.303478	3.303466	3.303468	3.303471	3.303462	3.303463	3.303462
WORST	3.303766	3.728121	3.303766	3.303766	3.303766	3.303766	3.303766	3.303766
STD	5.6032E-5	0.09555	5.9076E-5	4.0144E-5	1.1470E-4	4.6495E-5	5.0819E-5	5.1780E-5

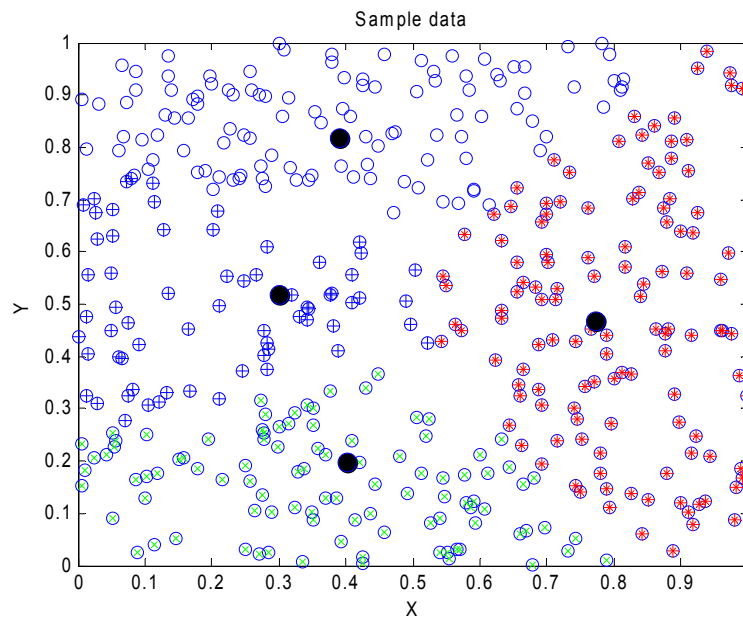


Figure 2. Clustering results of FCM algorithm.

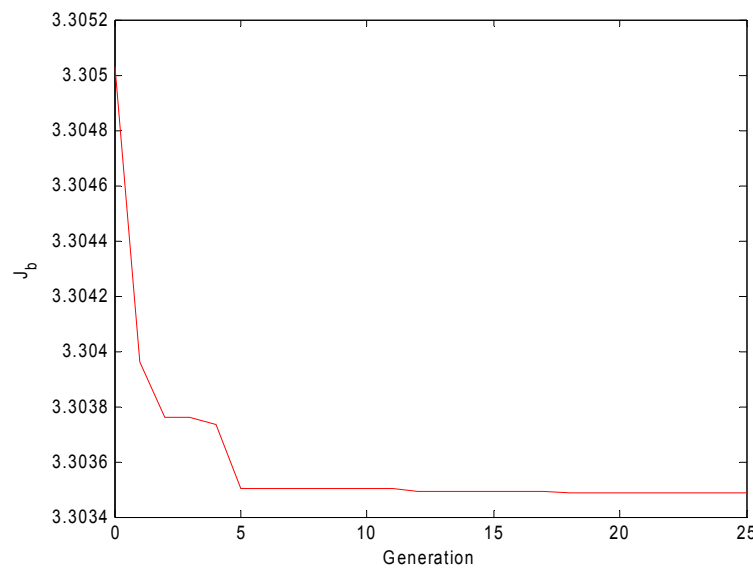


Figure 3. Optimization process for clustering problem of the KHE method.

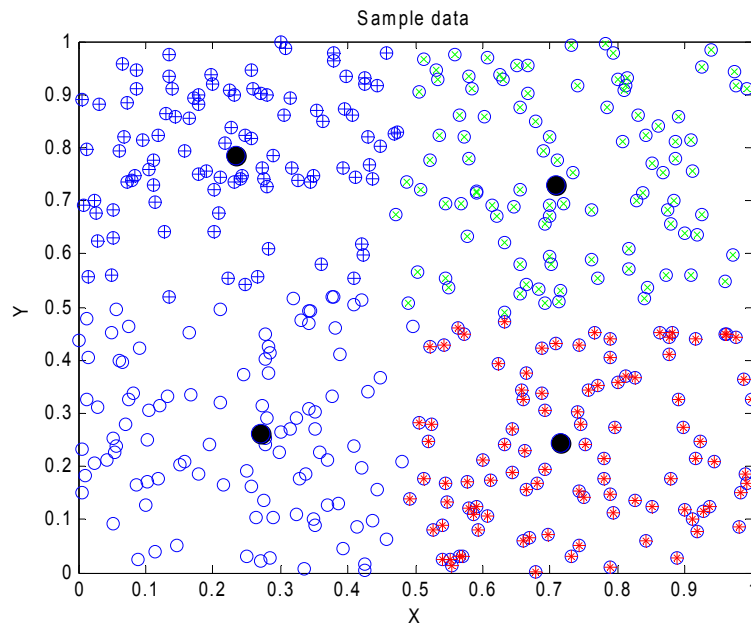


Figure 4. Clustering results of the KHE method.

5. Discussion and Conclusions

In many application fields, fuzzy clustering, especially fuzzy C-means (FCM) clustering, is one of the important hot research branches. FCM clustering algorithm is the most widely-used one among various clustering algorithms and has been used to successfully solve several application problems. However, the FCM algorithm is essentially a local search optimization algorithm. Herein, if its initial value is selected improperly, it will converge to a local minimum. Aiming at the disadvantages of the FCM algorithm, a new kind of swarm-based metaheuristic search, called KHE, is proposed to solve the clustering problem. Elitism strategy used in the KHE method can prevent the krill population from degrading. In KHE, the well-selected parameters are used instead of originating from nature. Furthermore, the KHE method is applied to solve the clustering problem for the purpose of escaping a local minimum. Moreover, with the aim of showing the performance of KHE method, it is compared with six other metaheuristic algorithms through seven complicated benchmark problems. The results show that the KHE method performs well on given benchmark problems and fuzzy clustering analyses.

Moreover, there are no additional operators added to the basic KH method. Therefore, the KHE method is simple and easy to implement.

Despite the above advantages of the KHE method, two prospective research points should be oriented as follows. On the one hand, in the current work, there is no study of computational requirements. The research of computational requirements should be made in future. On the other hand, only a few test problems and the clustering problem is solved by the KHE method in the present work. More problems should be used to test the KHE method from various aspects, and then it is used to solve more application problems, such as image segmentation, constrained optimization, knapsack problems, scheduling, dynamic optimization, antenna and microwave design problems, and water, geotechnical and transport engineering.

Acknowledgments

This work was supported by Jiangsu Province Science Foundation for Youths (No. BK20150239) and the National Natural Science Foundation of China (No. 61503165).

Author Contributions

All of the authors contributed to the content of this paper. Zhi-Yong Li and Jiao-Hong Yi participated in the algorithm analyses, design, algorithm implementation and draft preparation. Gai-Ge Wang analyzed the experimental data and revised this paper. All authors read and approved the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61.
2. Saremi, S.; Mirjalili, S.Z.; Mirjalili, S.M. Evolutionary population dynamics and grey wolf optimizer. *Neural. Comput. Appl.* **2014**, *26*, 1257–1263.
3. Goldberg, D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*; Addison-Wesley: New York, NY, USA, 1998.
4. Simon, D. Biogeography-based optimization. *IEEE Trans. Evol. Comput.* **2008**, *12*, 702–713.
5. Saremi, S.; Mirjalili, S.; Lewis, A. Biogeography-based optimisation with chaos. *Neural. Comput. Appl.* **2014**, *25*, 1077–1097.
6. Wang, G.G.; Gandomi, A.H.; Alavi, A.H. An effective krill herd algorithm with migration operator in biogeography-based optimization. *Appl. Math. Model.* **2014**, *38*, 2454–2462.
7. Li, X.; Zhang, J.; Yin, M. Animal migration optimization: An optimization algorithm inspired by animal migration behavior. *Neural. Comput. Appl.* **2014**, *24*, 1867–1877.
8. Rashedi, E.; Nezamabadi-Pour, H.; Saryazdi, S. Gsa: A gravitational search algorithm. *Inf. Sci.* **2009**, *179*, 2232–2248.
9. Mirjalili, S.; Wang, G.G.; Coelho, L.D.S. Binary optimization using hybrid particle swarm optimization and gravitational search algorithm. *Neural. Comput. Appl.* **2014**, *25*, 1423–1435.
10. Mirjalili, S.; Lewis, A. Adaptive gbest-guided gravitational search algorithm. *Neural. Comput. Appl.* **2014**, *25*, 1569–1584.
11. Yang, X.S.; Deb, S. Cuckoo search via lévy flights. In *Proceeding of the World Congress on Nature & Biologically Inspired Computing (NaBIC 2009)*, Coimbatore, India, 9–11 December 2009; Abraham, A., Carvalho, A., Herrera, F., Pai, V., Eds.; IEEE Publications: Coimbatore, India, 2009; pp. 210–214.
12. Li, X.; Wang, J.; Yin, M. Enhancing the performance of cuckoo search algorithm using orthogonal learning method. *Neural. Comput. Appl.* **2013**, *24*, 1233–1247.
13. Wang, G.G.; Deb, S.; Gandomi, A.H.; Zhang, Z.; Alavi, A.H. Chaotic cuckoo search. *Soft Comput.* **2015**, doi: 10.1007/s00500-015-1726-1.

14. Wang, G.G.; Gandomi, A.H.; Zhao, X.; Chu, H.E. Hybridizing harmony search algorithm with cuckoo search for global numerical optimization. *Soft Comput.* **2014**, *25*, 1423–1435.
15. Wang, G.G.; Gandomi, A.H.; Yang, X.S.; Alavi, A.H. A new hybrid method based on krill herd and cuckoo search for global optimization tasks. *Int. J. Bio-Inspir. Comput.* **2012**. Available online: <http://www.inderscience.com/info/ingeneral/forthcoming.php?jcode=ijbic> (accessed on 21 October 2015).
16. Khatib, W.; Fleming, P. The stud ga: A mini revolution? In Proceedings of the 5th International Conference on Parallel Problem Solving from Nature, New York, NY, USA, 4–9 May 1998; Eiben, A., Back, T., Schoenauer, M., Schwefel, H., Eds.; Springer-Verlag: New York, NY, USA, 1998; pp. 683–691.
17. Fong, S.; Deb, S.; Yang, X.S. A heuristic optimization method inspired by wolf preying behavior. *Neural. Comput. Appl.* **2015**, *26*, 1725–1738.
18. Mirjalili, S.; Mirjalili, S.M.; Hatamlou, A. Multi-verse optimizer: A nature-inspired algorithm for global optimization. *Neural. Comput. Appl.* **2015**, doi:10.1007/s00521-015-1870-7.
19. Mirjalili, S. Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural. Comput. Appl.* **2015**, doi:10.1007/s00521-015-1920-1
20. Mirjalili, S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowl-Based Syst.* **2015**, doi:10.1016/j.knosys.2015.07.006.
21. Wang, G.G.; Deb, S.; Coelho, L.D.S. Earthworm optimization algorithm: A bio-inspired metaheuristic algorithm for global optimization problems. *Int. J. Bio-Inspir. Comput.* **2015**. Available online: <http://www.inderscience.com/info/ingeneral/forthcoming.php?jcode=ijbic> (accessed on 21 October 2015).
22. Geem, Z.W.; Kim, J.H.; Loganathan, G.V. A new heuristic optimization algorithm: Harmony search. *Simulation* **2001**, *76*, 60–68.
23. Wang, G.; Guo, L.; Duan, H.; Wang, H.; Liu, L.; Shao, M. Hybridizing harmony search with biogeography based optimization for global numerical optimization. *J. Comput. Theor. Nanosci.* **2013**, *10*, 2318–2328.
24. Yang, X.S. Firefly algorithm, stochastic test functions and design optimisation. *Int. J. Bio-Inspir. Comput.* **2010**, *2*, 78–84.
25. Wang, G.G.; Guo, L.; Duan, H.; Wang, H. A new improved firefly algorithm for global numerical optimization. *J. Comput. Theor. Nanosci.* **2014**, *11*, 477–485.
26. Wang, G.; Guo, L.; Duan, H.; Liu, L.; Wang, H. A modified firefly algorithm for ucav path planning. *Int. J. Hybrid Inf. Technol.* **2012**, *5*, 123–144.
27. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceeding of the IEEE International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.
28. Mirjalili, S.; Lewis, A. S-shaped versus v-shaped transfer functions for binary particle swarm optimization. *Swarm. Evol. Comput.* **2013**, *9*, 1–14.
29. Mirjalili, S.; Lewis, A.; Sadiq, A.S. Autonomous particles groups for particle swarm optimization. *Arab. J. Sci. Eng.* **2014**, *39*, 4683–4697.

30. Wang, G.G.; Deb, S.; Cui, Z. Monarch butterfly optimization. *Neural. Comput. Appl.* **2015**, doi:10.1007/s00521-015-1923-y.
31. Dorigo, M.; Maniezzo, V.; Colorni, A. Ant system: Optimization by a colony of cooperating agents. *IEEE Trans. Syst. Man Cybern. B Cybern.* **1996**, *26*, 29–41.
32. Yang, X.S. *Nature-Inspired Metaheuristic Algorithms*; 2nd ed.; Luniver Press: Frome, UK, 2010.
33. Yang, X.S.; Gandomi, A.H. Bat algorithm: A novel approach for global engineering optimization. *Eng. Comput.* **2012**, *29*, 464–483.
34. Mirjalili, S.; Mirjalili, S.M.; Yang, X.S. Binary bat algorithm. *Neural. Comput. Appl.* **2013**, *25*, 663–681.
35. Zhang, J.W.; Wang, G.G. Image matching using a bat algorithm with mutation. *Appl. Mech. Mater.* **2012**, *203*, 88–93.
36. Yang, X.S.; He, X. Bat algorithm: Literature review and applications. *Int. J. Bio-Inspir. Comput.* **2013**, *5*, 141–149.
37. Storn, R.; Price, K. Differential evolution-A simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **1997**, *11*, 341–359.
38. Zou, D.; Wu, J.; Gao, L.; Li, S. A modified differential evolution algorithm for unconstrained optimization problems. *Neurocomputing* **2013**, *120*, 469–481.
39. Zou, D.; Liu, H.; Gao, L.; Li, S. An improved differential evolution algorithm for the task assignment problem. *Eng. Appl. Artif. Intell.* **2011**, *24*, 616–624.
40. Wang, G.G.; Gandomi, A.H.; Yang, X.S.; Alavi, A.H. A novel improved accelerated particle swarm optimization algorithm for global numerical optimization. *Eng. Comput.* **2014**, *31*, 1198–1220.
41. Gandomi, A.H. Interior search algorithm (isa): A novel approach for global optimization. *ISA Trans.* **2014**, *53*, 1168–1183.
42. Zou, D.; Gao, L.; Wu, J.; Li, S.; Li, Y. A novel global harmony search algorithm for reliability problems. *Comput. Ind. Eng.* **2010**, *58*, 307–316.
43. Zou, D.; Gao, L.; Li, S.; Wu, J. An effective global harmony search algorithm for reliability problems. *Expert Syst. Appl.* **2011**, *38*, 4642–4648.
44. Zou, D.; Gao, L.; Li, S.; Wu, J. Solving 0–1 knapsack problem by a novel global harmony search algorithm. *Appl. Soft. Comput.* **2011**, *11*, 1556–1564.
45. Li, X.; Yin, M. Application of differential evolution algorithm on self-potential data. *PLoS ONE* **2012**, *7*, e51199.
46. Li, X.; Yin, M. An opposition-based differential evolution algorithm for permutation flow shop scheduling based on diversity measure. *Adv. Eng. Softw.* **2013**, *55*, 10–31.
47. Wang, G.; Guo, L.; Duan, H.; Liu, L.; Wang, H.; Shao, M. Path planning for uninhabited combat aerial vehicle using hybrid meta-heuristic de/bbo algorithm. *Adv. Sci. Eng. Med.* **2012**, *4*, 550–564.
48. Li, X.; Yin, M. Parameter estimation for chaotic systems by hybrid differential evolution algorithm and artificial bee colony algorithm. *Nonlinear Dyn.* **2014**, *77*, 61–71.
49. Li, X.; Wang, J.; Zhou, J.; Yin, M. A perturb biogeography based optimization with mutation for global numerical optimization. *Appl. Math. Comput.* **2011**, *218*, 598–609.
50. Li, X.; Yin, M. Multi-operator based biogeography based optimization with mutation for global numerical optimization. *Comput. Math. Appl.* **2012**, *64*, 2833–2844.

51. Li, X.; Yin, M. Self-adaptive constrained artificial bee colony for constrained numerical optimization. *Neural. Comput. Appl.* **2012**, *24*, 723–734.
52. Mirjalili, S.; Mohd Hashim, S.Z.; Moradian Sardroudi, H. Training feedforward neural networks using hybrid particle swarm optimization and gravitational search algorithm. *Appl. Math. Comput.* **2012**, *218*, 11125–11137.
53. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Let a biogeography-based optimizer train your multi-layer perceptron. *Inf. Sci.* **2014**, *269*, 188–209.
54. Li, X.; Yin, M. Multiobjective binary biogeography based optimization for feature selection using gene expression data. *IEEE Trans. Nanobiosci.* **2013**, *12*, 343–353.
55. Gandomi, A.H.; Alavi, A.H. Krill herd: A new bio-inspired optimization algorithm. *Commun. Nonlinear Sci. Numer. Simul.* **2012**, *17*, 4831–4845.
56. Gandomi, A.H.; Talatahari, S.; Tadbiri, F.; Alavi, A.H. Krill herd algorithm for optimum design of truss structures. *Int. J. Bio-Inspir. Comput.* **2013**, *5*, 281–288.
57. Wang, G.G.; Gandomi, A.H.; Alavi, A.H.; Hao, G.S. Hybrid krill herd algorithm with differential evolution for global numerical optimization. *Neural. Comput. Appl.* **2014**, *25*, 297–308.
58. Wang, G.G.; Gandomi, A.H.; Alavi, A.H. Stud krill herd algorithm. *Neurocomputing* **2014**, *128*, 363–370.
59. Wang, G.; Guo, L.; Wang, H.; Duan, H.; Liu, L.; Li, J. Incorporating mutation scheme into krill herd algorithm for global numerical optimization. *Neural. Comput. Appl.* **2014**, *24*, 853–871.
60. Guo, L.; Wang, G.G.; Gandomi, A.H.; Alavi, A.H.; Duan, H. A new improved krill herd algorithm for global numerical optimization. *Neurocomputing* **2014**, *138*, 392–402.
61. Wang, G.G.; Gandomi, A.H.; Alavi, A.H. Study of lagrangian and evolutionary parameters in krill herd algorithm. In *Adaptation and Hybridization in Computational Intelligence*; Fister, I., Fister, I., Jr., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 18, pp. 111–128.
62. Wang, G.G.; Gandomi, A.H.; Alavi, A.H.; Deb, S. A hybrid method based on krill herd and quantum-behaved particle swarm optimization. *Neural. Comput. Appl.* **2015**, doi:10.1007/s00521-015-1914-z.
63. Wang, G.G.; Guo, L.; Gandomi, A.H.; Hao, G.S.; Wang, H. Chaotic krill herd algorithm. *Inf. Sci.* **2014**, *274*, 17–34.
64. Yang, X.S.; Cui, Z.; Xiao, R.; Gandomi, A.H.; Karamanoglu, M. *Swarm Intelligence and Bio-Inspired Computation*; Elsevier: Waltham, MA, USA, 2013.
65. Jamil, M.; Yang, X.S. A literature survey of benchmark functions for global optimisation problems. *Int. J. Math. Model. Numer. Optim.* **2013**, *4*, 150–194.
66. Wang, G.G.; Gandomi, A.H.; Alavi, A.H. A chaotic particle-swarm krill herd algorithm for global numerical optimization. *Kybernetes* **2013**, *42*, 962–978.
67. Liu, Q.; Wang, Z.; Liu, S. A optimization clustering algorithm based on simulated annealing and genetic algorithm. *Microcomput. Inf.* **2006**, *22*, 270–272.