A whale optimization algorithm (WOA) approach for clustering

Jhila Nasiri and Farzin Modarres Khiyabani

Abstract: Clustering is a powerful technique in data-mining, which involves identifying homogeneous groups of objects based on the values of attributes. Meta-heuristic algorithms such as particle swarm optimization, artificial bee colony, genetic algorithm and differential evolution are now becoming powerful methods for clustering. In this paper, we propose a new meta-heuristic clustering method, the Whale Clustering Optimization Algorithm, based on the swarm foraging behavior of humpback whales. After a detailed formulation and explanation of its implementation, we will then compare the proposed algorithm with other existing well-known algorithms in clustering, including PSO, ABC, GA, DE and k-means. Proposed algorithm was tested using one artificial and seven real benchmark data sets from the UCI machine learning repository. Simulations show that the proposed algorithm can successfully be used for data clustering.

1. Introduction

Data mining is the procedure of identifying correlation and patterns among attributes in databases by using appropriate techniques.

ABOUT THE AUTHORS
J. Nasiri Roveshti is currently Ph.D. student in IAU, Tabriz Branch, Iran. Her current interests include data mining and optimization.

Farzin Modarres Khiyabani is an assistant professor of mathematics at Tabriz Islamic Azad University. His research interests include operation research, meta-heuristic algorithms, numerical optimization and image processing. His articles have been published in various international journals indexed in WOS.

PUBLIC INTEREST STATEMENT
Clustering is an important and useful operation in data-mining, which involves classifying a particular set of unlabeled data into two or more groups, so that there is maximum similarity among the data of each cluster based on the selection criteria. Clustering algorithms cover a broad spectrum of utilization, having applications in such diverse fields as medical science, decision making, manufacturing, image processing, etc. Nowadays, considering the existence of such a massive amount of non-labeled data, the deployment of intelligent methods for data clustering has become a necessity. Therefore, this study deploys whale optimization algorithm (WOA) to solve clustering problems. The obtained solutions of the proposed algorithm are more accurate than those achieved using existing methods; moreover, due to its search methodology, the possibility of local optima entrapment is very low.
Data mining levels are preprocessing of data, selection of efficient and appropriate algorithm and to analyze the data. Some of the methods in data mining are the summation, association, and clustering. Data clustering is the most popular and practical method of them, that have been used for a variety of applications in data mining as pattern recognition (Andrews, 1972; Chen, Kim, & Mahnassani, 2014; Chen, Qi, Fan, Limeng, & Yong, 2015), image analysis (Aziz, Ewees, & Hassani, 2018; Chien-Chang, Hung-Hun, Meng-Yuan, & Horng-Shing, 2018; Lillesand & Keifer, 1994), security, vector quantization and other fields (Chiranjeevi & Jena, 2016; Gribkova, 2015; He, Pan, & Lin, 2006).

Clustering is gathering unlabelled objects into groups with respect to similarities between these objects. Such that the objects in the same cluster are more similar to each other than objects in different clusters according to some predefined criteria (Elhag & Ozcan, 2018; Zhang, Ouyang, & Ning, 2010). A number of algorithms have been proposed that take into account the nature of the data, the quantity of the data and other input parameters in order to cluster the data. The similarity criteria in clustering are various in different researches. Most of the clustering problems have exponential complexity in terms of the number of clusters. Because most of the similarity criterion functions are non-complex and nonlinear, clustering problems have several local solutions (Welch, 1982).

Clustering algorithms can be simply classified as hierarchical clustering and partitional clustering (Frigui & Krishnapuram, 1999; Han & Kamber, 2001; Leung, Zhang, & Xu, 2000; Sander, 2003). Hierarchical clustering groups data with a sequence of partitions either from singleton clusters to a cluster, including all objects or vice versa. This study is centralized on partitional clustering, that divide the data set into a set of disjoint clusters. The most popular partitional clustering algorithms are the prototype-based clustering algorithms where each cluster is represented by the center of the cluster, and the used objective function is the sum of the distance from the object to the center.

\(k\)-means is a popular, center-based clustering approach due to its simplicity and efficiency with linear complexity. However, the solution of the \(k\)-means algorithm depends on the initial random state and always converges to the local optimum (Jain & Dubes, 1998; MacQueen, 1967). Recently, researchers to overcoming this problem have presented heuristic clustering algorithms. Due to the large amount of information and the complexity of the problems, classical optimization methods are incapable of solving most of the optimization problems; therefore, researchers have started to use meta-framework algorithms. Today, nature-inspired algorithms are widely used to solve these problems in various fields (Faieghi & Baeanu, 2012; Farnad & Baleana, 2018; Sharma & Buddhirju, 2018). Meanwhile, clustering techniques, as well as other data mining and data analysis steps, have made significant progress using the collective intelligence algorithms.

Clustering with heuristic algorithms is emerging as an alternative to more conventional clustering techniques (Cui, 2017; Zhang et al., 2010).

Selim and Al-Sultan (Selim & Al-Sultan, 1991) used a simulated annealing approach for the clustering problem. Predetermined parameters of the algorithm are discussed and its convergence to a global solution of the clustering problem is demonstrated.

Mualik and Mukhopadhyay (Mualik & Mukhopadhyay, 2010) presented a combined clustering algorithm. They combined SA with artificial neural networks to improve solution quality. The proposed hybrid algorithm was used to cluster three real-life microarray data sets and the results of the proposed approach were compared with some commonly used clustering algorithms. The results indicated the superiority of the new algorithm. Mualik and Bandyopadhyay (Mualik & Bandyopadhyay, 2000) presented an approach based on genetic algorithm to solve the clustering problem. They examined the algorithm on synthetic and real-life data sets to evaluate its performance.
Shelokar et al. (Shelokar, Jayaraman, & Kulkarni, 2004) proposed a clustering algorithm based on ant colony optimization (ACO). The proposed algorithm was tested on some artificial and real-life data sets. The performance of this technique in comparison with popular algorithms such as GA, SA, and TS appeared to be very promising. Merve et al. (Van Der Merve & Engelhrecht, 2003) presented an approach to solving clustering problem used the particle swarm optimization (PSO) algorithm. A PSO clustering and a hybrid method were used, where the particles of the swarm are selected by the answers of the k-means algorithm. Both methods were compared with the k-means algorithm and the results indicated that the proposed algorithms and better answers.

Tunchan (Tunchan, 2012) presented a new PSO approach to the clustering problem that is efficient, easy-to-tune and applicable when the number of clusters is known or unknown. Karaboga et al. (Karaboga & Ozturk, 2011) used the artificial bee colony algorithm to solve the clustering problem. The results of simulations on 13 test problems from UCI indicated the superior performance of the proposed algorithm in comparison to PSO algorithm and some other approaches. Furthermore, the authors were found that the ABC algorithm can be appropriate to solve multivariate clustering problems.

Zhang et al. (Zhang et al., 2010) proposed an artificial bee colony (ABC) clustering algorithm to clustering that Deb’s rule is used to selection process instead of greedy selection. They test the algorithm on several well-known real data sets and compared with other popular heuristics algorithms in clustering. Results were very encouraging in terms of the quality of clusters. Armando and Farmani (Armando & Farmani, 2014) proposed a method that is the combined of k-means and ABC algorithms to improve the efficiency of k-means in finding a global optimum solution.

Karthikeyan and Christopher (Karthikeyan & Christopher, 2014) propose an algorithm by a combination of PSO algorithm and ABC algorithm used for data. This algorithm is compared with other existing clustering algorithms to evaluate the performance of the proposed approach. Sandeep and Pankaj (Sandeep & Pankaj, 2014) proposed a new hybrid sequential clustering approach, that uses PSO algorithm in sequence with the Fuzzy k-means algorithm in data clustering. Experimental results show that the new approach improves quality of formed clusters and avoids being trapped in local optimta.

Recently, Mirjalili and Lewis (Mirjalili & Lewi, 2016) described a new swarm based meta-heuristic optimization algorithm that mimicking the social behavior of humpback whales in hunting that called whale optimization algorithm (WOA)(Mirjalili, Mirjalili, & Lewis, 2014, Mirjalili & Lewi, 2016, Horng, Dao, Shieh, & Nguyen, 2017). The algorithm is inspired by the bubble_net hunting strategy. They have tested the WOA algorithm with 29 mathematical benchmark optimization problems and compared the performance of WOA algorithm with other conventional modern heuristic algorithms such as PSO (Kenedy & Eberhart, 1995), Differential Evolutional (Storn & Price, 1997), Gravitational Search Algorithm (Rashedi, Nezamabadi- Pour, & Saryazdi, 2009) and Fast Evolutionary Programming (Yao & Liu, 1999). WOA was recognized to be enough competitive with other well-known and popular meta-heuristic methods.

In this paper, the WOA algorithm is extended for solving the clustering problem as an optimization problem. We intend to use the advantages of the whale optimization algorithm, such as the low number of parameters and lack of local optima entrapment, in solving clustering problems. Our main goal is to cluster unlabeled data using the whale optimization algorithm so that we can get better results with simple solutions and do a complete search compared to the existing methods. The performance of the proposed algorithm has been tested on various data sets and compared with several clustering algorithms. The remaining of this paper is organized as follows: Section 2 discussed the clustering problem. Section 3 describes the WOA algorithm. The new WOA clustering algorithm was proposed in section 4. Section 5 provided the experimental results. Finally, conclusions and some future research direction are in section 6.
2. The clustering problem
Clustering is applied to grouping data objects of a given data set based on some similarity measures. Similar objects are ideally put in the same cluster while dissimilar objects are placed in different clusters. Most of the researchers used instance measurement for evaluating similarities between objects, which is obtained from the Minkouski metric.

In general, the problem can be expressed as follows: Suppose that 
\[ s = \{x_1, x_2, \ldots, x_n\} \]
be a set of \( n \) objects of a data set with \( m \) dimensions. Each \( x_i \) is described by a real-valued \( m \) dimensional vector as 
\[ x_i = (x_{i1}, x_{i2}, \ldots, x_{im}) \]
where each \( x_{ij} \) denotes the value of \( j \)th attribute of the \( i \)th object. The goal of clustering is assigned each object \( x_i \) to one of the \( k \) cluster in the set of partition 
\[ z = \{c_1, c_2, \ldots, c_k\} \]
such that distance between \( x_i \) and \( z_k \) center of \( k \)th cluster center be the minimum and

1. \( \forall i \neq j \) \( c_i \cap c_j = \emptyset \),
2. \( \bigcup_{i=1}^{k} c_i = s \),
3. \( \forall c_i \neq \emptyset \).

So, clustering problem is minimizing the following Euclidean distance:

\[
d(s, z) = \left( \sum_{i=1}^{n} \sum_{k=1}^{K} w_{ik} \|x_i - z_k\|^2 \right)^{\frac{1}{2}},
\]  \quad (2.1)

where \( x_i \) denotes \( i \)th data object and \( z_k \) represents the \( k \)th cluster center. \( w_{ik} \) is the association weight of pattern \( x_i \) with cluster \( k \) and defined as

\[
w_{ik} = \begin{cases} 
1 & \text{if } \|x_i - z_k\|^2 = \min_{1 \leq j \leq n} \|x_i - z_j\|^2 \\
0 & \text{otherwise} 
\end{cases}.
\]  \quad (2.2)

According to Equation (2.2), we assign each object to the nearest cluster center out of the all cluster centers.

3. Whale optimization algorithm
Whale optimization algorithm was proposed by Jalili and Lewis for optimizing numerical problems (Mirjalili & Lewis, 2016). The algorithm simulates the intelligence hunting behavior of humpback whales. This foraging behavior is called bubble-net feeding method that is only be observed in humpback whales. The whales create the typical bubbles along a circle path while encircling prey during hunting. Simply, bubble-net hunting behavior could describe such that humpback whales dive down approximation 12 m and then create the bubble in a spiral shape around the prey and then swim upward the surface following the bubbles. In order to perform optimization, the mathematical model for spiral bubble-net feeding behavior is given as follows:

3.1. Encircling prey
Humpback whales can find the place of prey and encircle them. The WOA algorithm considers; current best search agent position be the target prey or close to the optimum point, and other search agents will try to update their position towards the best search agent. This behavior is formulated as the following equations:

\[
D = |C \cdot \overline{x'}(t) - \overline{x}(t)|,
\]  \quad (3.1)

\[
\overline{x}(t + 1) = \overline{x'}(t) - A \cdot D,
\]  \quad (3.2)
where $t$ indicates the current iteration, $\mathbf{X}^*$ is the position vector of the best solution have been obtained so far iteration $t$, $\mathbf{x}$ is the position vector of each agent, $|$ | is the absolute value, and $.$ is an element-by-element multiplication. The coefficient vectors $\mathbf{A}$ and $\mathbf{C}$ are calculated as follows:

$$\tilde{\mathbf{A}} = 2a \, r - \bar{a}, \quad (3.3)$$

$$\tilde{\mathbf{C}} = 2r. \quad (3.4)$$

where $\bar{a}$ is linearly decreased from 2 to 0 over the course of the iteration and $r$ is a random number $[0,1]$. 

### 3.2. Bubble-net attacking method

The Bubble-net strategy is hybrid of combined two approaches that can be mathematically model as follows:

#### a. Shrinking Encircling Mechanism

This behavior of Whales simulated by decreasing the value of $\bar{a}$ in the equation (3.3). Note that the fluctuation range of $\mathbf{A}$ is also decreased by $\bar{a}$. In other words, $\mathbf{A}$ is a random value in the interval $[-\bar{a}, \bar{a}]$ where $\bar{a}$ is decreased from 2 to 0 over the course of iterations. Setting random values for $\tilde{\mathbf{A}}$ in $[-1,1]$, the new position of a search agent can be defined anywhere in between the original position of the agent and the position of the current best agent.

#### b. Spiral Updating Position

In this approach, a spiral equation is created between the position of whale and prey to simulate the helix-shaped movement of humpback whales as follows:

$$\mathbf{D}^0 = \left| \mathbf{X}^*(t) - \mathbf{x}(t) \right|, \quad (3.5)$$

$$\mathbf{x}(t + 1) = \mathbf{D}^0 \, e^{bl} \cos(2\pi l) + \mathbf{X}^*(t), \quad (3.6)$$

where $\mathbf{D}^0$ is the distance between the whale and prey, $b$ is constant defines the logarithmic shape, $l$ is random in $[-1,1]$ and is an element-by-element multiplication.

Indeed, humpback whales swim along a spiral-shaped path and at the same time within shrinking circle. Assuming a probability of 50%, choosing either the shrinking encircling movement or the spiral model movement is simulated during iterations of the algorithm. It means that:

$$\mathbf{x}(t + 1) = \begin{cases} \mathbf{X}^*(t) - \mathbf{A} \cdot \mathbf{D} & \text{if } p < 0.5; \\ \mathbf{D}^0 \, e^{bl} \cos(2\pi l) + \mathbf{X}^*(t) & \text{if } p \geq 0.5, \end{cases} \quad (3.7)$$

where $p$ is a random number in $[0,1]$. 

### 3.3. Search for prey

Almost all meta-heuristic algorithms explore the optimum using random selection. In the bubble-net method, the position of the optimal design is not known, so humpback whales search for prey randomly. In contrast to the exploitation phase with $\mathbf{A}$ in interval $[-1,1]$ in this phase consider, $\mathbf{A}$ be a vector of the random values greater than 1 or less than $-1$. With this assumption, search agent able to move far away from a reference whale. In return, the position of search agent will be updated according to randomly chosen from search agent, instead of the best search agent found so far. These two actions formulated as follows:
\[ \bar{D} = |C \cdot \overrightarrow{X_{\text{rand}}} - \bar{X}|, \]  

(3.8)

\[ \bar{X}(t + 1) = \overrightarrow{X_{\text{rand}}} - \bar{A} \cdot \bar{D}. \]  

(3.9)

where \( \overrightarrow{X_{\text{rand}}} \) is a random position vector.

The WOA algorithm starts from a set of random solutions. At each iteration, search agents update their position according to the above explanations. WOA is a global optimizer. Adaptive variation of the search vector \( \bar{A} \) allows the WOA algorithm easily transit between exploration and exploitation. Furthermore, WOA includes only two main internal parameters to be adjusted. High exploration ability of WOA is due to the position updating mechanism of Whales using (3.9). High exploitation and convergence are emphasized, which originate from (3.6) and (3.2). These equations show that the WOA algorithm is able to provide high local optima avoidance and convergence speed during the course of the iteration.

4. Whale optimization-based clustering algorithm

Whale optimization algorithm is a new meta-heuristic optimization algorithm that simulates the intelligence bubble-net hunting behavior of humpback whales. WOA is a simple, robust and swarm based stochastic optimization algorithm. Population-based WOA has an ability to avoid local optima and get a global optimal solution. These advantages cause WOA to be an appropriate algorithm for solving different constrained or unconstrained optimization problems for practical applications without structural reformation in the algorithm.

In the context of WOA, a swarm refers to a number of potential solutions to the optimization problem, where each potential solution is referred to as a search agent. The aim of the WOA is to find the search agent position that results in the best evaluation of a given objective function.

In this section, we are going to solve the clustering problem using WOA. Inspired by the context of clustering, assume that search agent represents \( k \) cluster centers (\( k \) is predefined and shows the number of clusters). Each search agent \( X_i \) is constructed as follows:

\[ X_i = (z_{i1}, z_{i2}, \ldots, z_{ik}), \]  

(4.1)

where \( z_{ij} \) refers to the \( j \)th cluster center vector of the \( i \)th search agent in cluster \( c_{ij} \). Therefore, a swarm represents a number of candidates clustering for the vectors of the data set. We prefer intra-distance of clusters be the fitness function that measures the distance between cluster center and data vectors of the same cluster according to Equations (2.1) and (2.2).

According to above assumptions the pseudo code of whale optimization clustering algorithm proposes as follows: pseudocode

Algorithm 1 The Whale Optimization-based Clustering Algorithm

1: procedure
2:  Load data samples
3:  Initialize each search agent to contain \( k \) randomly cluster centers
4:  while \( t < \) Iteration do
5:      for each search agent \( i \) do
6:          for each data vector \( x_p \) do
7:              Calculate the Euclidean distance of \( x_p \) to all cluster centers.
8:              Assign \( x_p \) to the cluster \( c_{ij} \) such that

\[ |x_p - z_j| = \min_{c=1,2,\ldots,k} |x_p - z_c|. \]
9: Calculate the fitness using (4.2)

\[
\text{fitness} = \sum_{j=1}^{\mathbf{k}} \sum_{i=1}^{n} w_{ij} |x_j - z_j|.
\]

(4.2)

where \( w_{ij} = \begin{cases} 
1 & \text{if } |x_i - z_j| = \min_{1 \leq m \leq k} |x_i - z_m| \\
0 & \text{else}
\end{cases} \).

10: end for
11: end for
12: \( \mathbf{X}^* \) is the best search agent
13: for each search agent do
14: update \( \mathbf{a}, \mathbf{A}, \mathbf{C}, \mathbf{I} \) and \( \mathbf{p} \)
15: if \( p < 0.5 \) then
16: if \( |\mathbf{A}| < 1 \) then
17: update search agent by (3.2)
18: else if \( |\mathbf{A}| \geq 1 \) then
19: select random search agent
20: update current search agent by (3.9)
21: end if
22: else if \( p \geq 0.5 \) then
23: update the position of current search agent by (3.6)
24: end if
25: end if
26: end for
27: \( t = t + 1 \)
28: end while
29: return \( \mathbf{X}^* \)

From theoretical standpoint, WOA clustering algorithm can be considered as a global optimizer because it includes exploration and exploitation ability simultaneously. Furthermore, the proposed hyper-cube mechanism defines a search space in the neighborhood of the best solution and allows other search agents to exploit the current best record inside that domain. Adaptive variation of the search vector \( \mathbf{A} \) allows the WOA clustering algorithm to smoothly transit between exploration and exploitation. It means that by decreasing \( \mathbf{A} \), some iterations are concentrated to exploration, and the rest is dedicated to exploitation.

5. Experimental study

In this work, the performance of WOA clustering approach was compared to well-known algorithms namely the ABC clustering proposed by Karaboga and Ozturk (Karaboga & Ozturk, 2011), the PSO clustering proposed by Merve and Engelbrecht (Van Der Merve & Engelbrecht, 2003), the differential evolution-based (DE) clustering algorithm proposed by Sakar et al. (Sarkar, Yegnanafayana, & Khemani, 1997), the genetic algorithm-based clustering technique, called GA-clustering (GA), proposed by Mualik and Bandyopadhyay (Mualik & Bandyopadhyay, 2000) and \( k \)-means a vector quantization algorithm proposed by MacQueen (MacQueen, 1967). All algorithms were programmed in Matlab 2013b and executed on an Intel core, i7 CPU, 4 Gb and 1.73 GHz computer running Microsoft Windows XP. The parameter settings are the same in the original corresponding papers. Here, 8 data sets are used to evaluate the performance of proposed algorithm compared with above heuristics. One artificial data set that is generated in Matlab environment using mean vector \( \mu \) and variance matrix sigma. Other data sets are from classification problems from the UCI databases (Blake & Merz, 1998) that are Iris, wine, contraceptive method choice (CMC), Balance, Breast Cancer, Glass and Thyroid data sets. The data sets are used in this study can be described as follows:
Artificial data set ART: This data set contains 300 objects and three clusters. Samples are drawn from three independent bivariate normal distribution, where classes were distributed according to
\[
\begin{align*}
\mu_1 &= [0, 0], \\
\mu_2 &= [3.4, 4], \\
\mu_3 &= [6.1, 1]
\end{align*}
\]
and
\[
\sigma = \begin{pmatrix} 0.5 & 0.05 \\ 0.05 & 0.5 \end{pmatrix}
\]
is the covariance matrix.

The artificial produced data set are described in Figure 1.

Figure 1 visualizes the clusters performed by WOA clustering approach on art data set. Attention at position of the cluster centers specified by WOA shows that proposed algorithm has a high ability to find the global optimal answers.

Iris data set: This data set contains 150 random samples of flowers from the Iris species setosa, versicolor and virginica used by Fisher (Fisher, 1936). From each species, there are 50 observations with four attributes, which are sepal length, sepal width, petal length and petal width in cm.

Wine data set: This data set are the results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars (Forina, Leardi, Armanino, & Lanteri, 1998). The analysis determined the quantities of 13 constituents found in each of the three types of wines. So, there are 178 instances with 13 numeric attributes in Wine data set.
CMC data set are a subset of the 1987 Indonesia Contraceptive Prevalence survey. The samples are married women who were either not pregnant or do not know if they were at the time of the interview. The problem involves predicting the choice of the current contraceptive method of a woman based on her demographic and socio-economic characteristics. This data set contains 1473 objects with 9 attributes and 3 clusters.

Balance data set: This data set was generated to model psychological experimental results. Each example is classified as having the balance scale tip to the right, tip to the left or be balanced. The data set include 4 inputs, 3 classes and there are 625 examples.

Cancer data set: This data set is based on Breast Cancer Wisconsin-Original set. It contains 569 patterns with 11 attributes and 2 clusters.

Glass data set: This data set contains one of the biggest numbers of classes. It is used to classify glass types as float processed building windows, vehicle windows, containers, tablewares or headlamps. Nine inputs are based on 9 chemical measurements with one of six types of glass, which are continuous with 70, 76, 17, 13, 9 and 29 instances of each class respectively. The total number of instances is 214.

Thyroid data set: This data set is the diagnosis of thyroid, whether it is hyper or hypofunction. Five inputs are used to classify three classes of Thyroid function as being over function, normal function or under function. The data set is based on new-thyroid data and contains 215 patterns. There are two control parameters in WOA algorithm. The swarm size of whales set in 50 and the maximum number of iteration supposed to 300. The parameter settings of ABC, PSO, GA, DE and ACO are set the same as their original papers. The effectiveness of stochastic algorithms is greatly dependent on 20 times individually for their own effectiveness test, each time with randomly generated initial solutions.

Table 1 summarizes results, obtained from the clustering algorithms for the data sets described above. The values reported are averages of intra-cluster distances over 20 simulations, standard deviations, and ranking of the techniques based on mean values. At a glance, it is obvious that the WOA algorithm gets the best performance in six of the problems and the second rank in two problems. In artificial problem and Iris, Balance, Cancer and Thyroid data sets proposed algorithm is in rank 1 in comparison with other algorithms. In Glass problem, that contains one of the biggest numbers of clusters, also the WOA clustering algorithm is in rank 1 and mean value of intra-cluster distance function is 231.29. This value is very smaller in comparison with other algorithms. In Glass problem PSO algorithm is in rank 2 and mean value, in this case, is 240.89. This result shows the excellence of WOA clustering algorithm between the other swarm intelligence algorithms. The high performance of WOA clustering in average values will be highlighted when reviewing the values of standard deviations in different clustering algorithms. For example, in Iris and Cancer problems, proposed algorithm has the smallest SD values. In Wine, CMC and Thyroid SD values are in second place.

It is worthy to note that, biggest value of SD is in WOA clustering algorithm is 4.51 that be seen in Glass problem. According to mean value of WOA in CMC problem, the high value of SD is negligible. As a result of above discussion, solutions of using WOA to solve clustering problem is very significant and successful. By foreseeing the benefits of WOA algorithm namely, the lowest number of predetermined parameters, simplicity in implementation, ability to avoid local optima and get a globally optimal solution, It can be said that WOA clustering is an excellent offer to solve clustering problems.

The convergence of curves of the WOA and PSO clustering for ART, Cancer, Glass and Thyroid data sets are provided in Figure 2. As it is obvious in this figure, usually the WOA clustering (with intra-cluster distance fitness function) behavior in the convergence towards the optimum only in final iterations. This is probably due to avoiding local optimum. This means that, the algorithm fails in local search and as a result does exploration in the initial steps of iteration avoiding local search,
but in the last iterations, algorithm cleverly tends towards the global optimum response. In 100 last iterations, the performance of algorithm in local search improves significantly and finally PSO fails compared with WOA.

6. Conclusion

Nowadays, simulation the intelligence behavior of animals and insects for solving the search and optimization problems is very common. Whale optimization algorithm which is inspired by bubble-net the haunting strategy of humpback whales is most closely studied a meta-heuristic algorithm in the area of swarm intelligence, which is a new, simple and robust optimization approach. In this paper, the whale optimization algorithm developed to solve popular clustering problem. Clustering is gathering data into clusters such that the data in the same cluster have a high degree of similarity and data from different clusters being as possible as dissimilar. The results of this algorithm is compared with well-known k-means clustering approach and other popular stochastic algorithms such as PSO, artificial bee colony, differential evolution, and genetic algorithm clustering. The Preliminary computational experience in terms of the intra-cluster distance function and standard deviation shows that the whale optimization algorithm can successfully be applied to solve clustering problems. Moreover, these results proposed algorithm was effective, easy to implementation and robust as compared with other approaches. There are some directions that can improve the performance of the proposed algorithm in the future. The combination of WOA clustering algorithm with other clustering approaches and using other fitness functions in clustering approach can be further researches.
Figure 2. Convergence of curves of the WOA and PSO for some data sets.

References


Funding
The authors received no direct funding for this research.

Author details
Jhila Nasiri1
E-mail: jhila_nasiri@yahoo.com
ORCID ID: http://orcid.org/0000-0003-3140-4211
Farzin Modarres Khiyabani1
E-mail: farzin.modarres@gmail.com
ORCID ID: http://orcid.org/0000-0002-7688-2535
1 Department of Mathematics, Tabriz Branch, Islamic Azad University, Tabriz, Iran.

Citation information
Cite this article as: A whale optimization algorithm (WOA) approach for clustering, Jhila Nasiri & Farzin Modarres Khiyabani, Cogent Mathematics & Statistics (2018), 5: 1483565.


