Clustering Approach Based On Von Neumann Topology Artificial Bee Colony Algorithm

Wenping Zou[1,2], Yunlong Zhu[1], Hanning Chen[1], Tao Ku[1]
[1]Key Laboratory of Industrial Informatics, Shenyang Institute of Automation,
Chinese Academy of Sciences,
110016, Shenyang, China
[2]Graduate School of the Chinese Academy of Sciences,
100039, Beijing, China
{zouwp, ylzhu, chenhanning, kutao}@sia.cn

Abstract - *Article Bee Colony (ABC) is one of the most recently introduced algorithms based on the intelligent foraging behavior of a honey bee swarm. This paper proposes a new variant of the ABC algorithm based on Von Neumann topology structure, namely Von Neumann Neighborhood Article Bee Colony (VABC). VABC significantly improves the original ABC in solving complex optimization problems. Clustering is a popular data analysis and data mining technique. The most popular technique for clustering is k-means algorithm. However, the k-means algorithm highly depends on the initial state and converges to local optimum solution. In this work, VABC algorithm is tested on a set of widely-used benchmark functions and is used for solve data clustering on several benchmark data sets. The performance of VABC algorithm is compared with ABC and Particle Swarm Optimization (PSO) algorithms. The simulation results show that the proposed VABC outperforms the other two algorithms in terms of accuracy, robustness, and convergence speed.*

Keywords: Article Bee Colony; Von Neumann topology; Swarm Intelligence; Data cluster; k-means;

1 Introduction

Artificial Bee Colony (ABC) algorithm is a new swarm intelligent algorithm, which was first introduced by Karabog in Erciyes University of Turkey in 2005 [1], and the performance of ABC is analyzed in 2007 [2]. The ABC algorithm imitates the behaviors of the real bees on searching food source and sharing the information of food sources to the other bees. Since the ABC algorithm is simple in concept, easy to implement, and has fewer control parameters, it has been widely used in many fields, such as constrained optimization problems [3], neural networks [4] and clustering [5].

In [6][7], James Kennedy et al. had introduced the effects of various population topologies on the PSO algorithm in detail and they considered that the PSO with Von Neumann configuration performed more better than the other topologies structure. Hence, this paper applies Von Neumann topology structure to the ABC. In order to evaluate the performance of the VABC, we compared the performance of the VABC algorithm with that of ABC and PSO on a set of well-known benchmark functions. From the simulation results, the VABC algorithm shows remarked performance improvement over the ABC and PSO algorithms in all benchmark functions.

Data clustering is the process of grouping data into a number of clusters. The goal of data clustering is to make the data in the same cluster share a high degree of similarity while being very dissimilar to data from other clusters. Clustering algorithms can be simply classified as hierarchical clustering and partitional clustering [8]. This paper mainly focuses on partitional clustering. The most popular partitional clustering algorithm is k-means. In the past three decades, k-means clustering algorithm has been used in various domains. However, k-means algorithm is sensitive to the initial states and always converges to the local optimum solution. In order to overcome this problem, many methods have been proposed. Over the last decade, more and more stochastic, population-based optimization algorithms have been applied to clustering problems. For instance, Shelokar et al. have introduced an evolutionary algorithm based on ACO algorithm for clustering problem [9][10], Merwe et al. have presented PSO to solve the clustering problem [11][12] and Karaboga et al. have used the ABC algorithm [13]. In this paper, a VABC algorithm is applied to solve the clustering problem, which has been tested on a variety of data sets. The performance of the VABC on clustering is compared with results of the ABC and PSO algorithms on the same data sets. The above data sets are provided from the UCI database [14].

The rest of the paper is organized as follows. In Section 2，we will introduce the original ABC algorithm. Section 3 will discuss the Von Neumann topology structure, and our Von Neumann topology implementations of the ABC algorithm will be presented. Section 4 tests the algorithms on the benchmarks, and the results obtained are presented and discussed. The application of VABC algorithm on clustering is shown in Section 5, and the performance of VABC algorithm is compared with ABC and PSO algorithms on clustering problem in this section. Finally, conclusions are given in Section 6.

2 The article bee colony algorithm

The artificial bee colony algorithm is a new population-based metaheuristic approach, initially proposed by Karaboga [1][2] and further developed by Karaboga and Basturk [3][4]. It has been used in various complex problems. The algorithm simulates the intelligent foraging behavior of honey bee swarms. It is a very simple and robust optimization algorithm. In the ABC algorithm, the colony of artificial bees is classified into three categories: employed bees, onlookers and scouts. Employed bees are associated with a particular food source which they are currently exploiting or are “employed” at. They carry with them information about this particular source and share the information to onlookers. Onlooker bees are those bees that are waiting on the dance area in the hive for the information to be shared by the employed bees about their food sources, and then make decision to choose a food source. A bee carrying out random search is called a scout. In the ABC algorithm, first half of the colony consists of the employed artificial bees and the second half includes the onlookers. For every food source, there is only one employed bee. In other words, the number of employed bees is equal to the number of food sources around the hive. The employed bee whose food source has been exhausted by the bee becomes a scout. The position of a food source represents a possible solution to the optimization problem and the nectar amount of a food source corresponds to the quality (fitness) of the associated solution represented by that food source. Onlookers are placed on the food sources by using a probability based selection process. As the nectar amount of a food source increases, the probability value with which the food source is preferred by onlookers increases [1][2]. The main steps of the algorithm are given in Table I:

In the initialization phase, the ABC algorithm generates a randomly distributed initial food source positions of *SN* solutions, where *SN* denotes the size of employed bees or onlooker bees. Each solution x_i (i =1, 2,..., *SN*) is a *D*-dimensional vector. Here, *D* is the number of optimization parameters. And then evaluate each nectar amount fit_i. In the ABC algorithm, nectar amount is the value of benchmark function.

In the employed bees’ phase, each employed bee finds a new food source v_i in the neighbourhood of its current source x_i. The new food source is calculated using the following equation (1):

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}) \tag{1}$$

where $k \in$ (1, 2,..., *SN*) and $j \in$ (1, 2,..., *D*) are randomly chosen indexes, and *k* has to be different from *i*. ϕ_{ij} is a random number between [-1,1]. And then employed bee compares the new one against the current solution and memorizes the better one by means of a greedy selection mechanism.

In the onlooker bees’ phase, each onlooker chooses a food source with a probability which related to the nectar amount (fitness) of a food source shared by employed bees. The probability is calculated using the following equation (2):

TABLE I
MAIN STEPS OF THE ABC ALGORITHM

1:cycle=1
2:Initialize the food source positions x_i, i=1 ...*SN*
3:Evaluate the nectar amount (fitness fit_i) of food sources
4:repeat
5: Employed Bees’ Phase
  For each employed bee
    Produce new food source positions v_i
    Calculate the value fit_i
    Apply greedy selection mechanism
  EndFor.
6: Calulate the probability values p_i for the solution.
7: Onlooker Bees’ Phase
  For each onlooker bee
    Chooses a food source depending on p_i
    Produce new food source positions v_i
    Calculate the value fit_i
    Apply greedy selection mechanism
  EndFor
8: Scout Bee Phase
  If there is an employed bee becomes scout
  Then replace it with a new random source positions
9: Memorize the best solution achieved so far
10 cycle=cycle+1.
11:until cycle=Maximum Cycle Number

$$p_i = fit_i / \sum_{n=1}^{SN} fit_i$$

(2)

In the scout bee phase, if a food source can not be improved through a predetermined cycles, called “limit”, it is removed from the population and the employed bee of that food source becomes scout. The scout bee finds a new random food source position using the equation (3) below:

$$x_i^j = x_{\min}^j + rand[0,1](x_{\max}^j - x_{\min}^j) \quad (3)$$

where $x_{\min}^j$ and $x_{\max}^j$ are lower and upper bounds of parameter j, respectively.

These steps are repeated through a predetermined number of cycles, called Maximum Cycle Number (MCN), or until a termination criterion is satisfied [1][2][15].

3 Article bee bolony algorithm based on von neumann topology structure

The essence of driving swarm algorithm activity is social communication. The individual of the swarm will communicate their knowledge with their neighborhoods. Hence, the different concepts for neighborhood lead to different neighborhood topologies. Different neighborhood topologies primarily affect the communication abilities. Some kinds of population structures work well on some functions, while other kinds work well better on other functions. In [7], Kennedy theorized that populations with fewer connections might perform better on highly multimodal problems, while highly interconnected populations would be better for unimodal problems. After studying the various population topologies on the PSO performance, Kennedy considered that Von Neumann topology structure worked well on a wide range of problems [7]. Actually, the original ABC algorithm is a star topolgy structure, which is a fully connected neighhor relation. From equation (1), we noticed that k could be every particle except i. That means every particle is neighbor of ith particle. Because of Kennedy’s studying, in this paper, we will we apply Von Neumann topology structure to the ABC, namely VABC.

Von Neumann topology was proposed by Kennedy and Mendes [7]. In the Von Neumann topolgy structure, an individual can communicate with four of its neighbors using a rectangular lattice topology. A graphical representation of the Von Neumann model is shown in Figure 1. In order to form the Von Neumann topology structure for M particles, we adopt below approch:

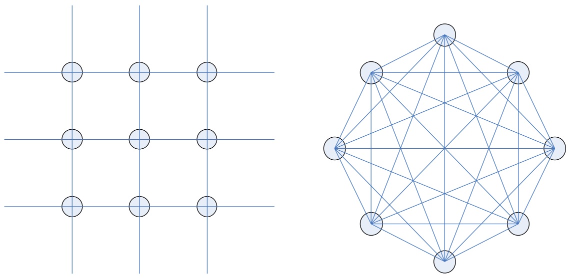

Fig. 1. The Von Neumann neighborhood is shown on the left and the star neighborhood is shown on the right.

(A) Arrange the M particles in rows rows and cols columns, that is M=*rows*cols*

(B) For the ith particle, $i \in \{1, 2, \ldots, M\}$:

a) Up neighbor: $N_i(1) = (i\text{-}cols) \bmod M$, if $N_i(1) = 0$, $N_i(1)=M$

b) Left neighbor: $N_i(2) = i\text{-}1$, if $(i\text{-}1) \bmod cols = 0$, $N_i(2) = i-1+cols$.

c) Right neighbor: $N_i(3) = i+1$, if $i \bmod cols = 0$, $N_i(3) = i+1\text{-}cols$.

d) Down neighbor: $N_i(4) = (i+cols) \bmod M$, if $N_i(4) = 0$, $N_i(4)=M$

Notice the negative number in mod calculation: (-5) mod 20 = 15 [16]

In the VABC algorithm, in the employed bees’ phase, we use Von Neumann topology. In the equation (1), k could only be the up, left, down right neighbor of particle i. However，in the onlooker bees’ phase, we don’t change the original ABC algorithm. This will improve convergence speed.

4 Experiments

4.1 Benchmark functions

In order to compare the performance of the proposed VABC algorithm with ABC and PSO, we used a set of well-known benchmark functions. The formulas and the properties of these functions are listed as follows:

Sphere function:

$$f_1(x) = \sum_{i=1}^{D} x_i^2$$

$$x \in [-5.12, 5.12]$$

Rosenbrock function:

$$f_2(x) = \sum_{i=1}^{D} 100(x_i^2 - x_{i+1})^2 + (1 - x_i)^2$$

$$x \in [-15, 15]$$

Griewank function:

$$f_3(x)=\frac{1}{4000}(\sum_{i=1}^{D}x_i^2)-(\prod_{i=1}^{D}\cos(\frac{x_i}{\sqrt{i}}))+1$$

$$x\in[-600,600]$$

Rastrigin function:

$$f_4(x)=\sum_{i=1}^{D}(x_i^2-10\cos(2\pi x_i)+10)$$

$$x\in[-15,15]$$

Ackley function:

$$f_5(x)=20+e-20\exp(-0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D}x_i^2})-\exp(\frac{1}{D}\sum_{i=1}^{D}\cos(2\pi x_i))$$

$$x\in[-32.768,32.768]$$

Schwefel function:

$$f_6(x)=D*418.9829+\sum_{i=1}^{D}-x_i\sin(\sqrt{|x_i|})$$

$$x\in[-500,500]$$

4.2 Simulation results

In the experiment, all functions are tested on 30 dimensions; and the population sizes of VABC, ABC and PSO algorithms were 100. The PSO algorithm we used is the standard PSO. In PSO algorithm, inertia weight ω varies from 0.9 to 0.7 linearly with the iterations and the acceleration factors c_1 and c_2 were both 2.0 [17]. The experimental results, including the mean and standard deviation of the function values found in 30 runs are proposed in Table II and the two algorithms were terminated after 1,000 generation.

From Table II, the VABC algorithm is better than the ABC and PSO algorithms on all functions. On f_2 and f_6 functions, both ABC and VABC algorithms are almost the same. On f_4 functions, in the experiment, we found that the best value of the fitness functions for ABC algorithm is as good as VABC algorithm. However, ABC algorithm is easy trapped at local optimum, so the average value is worse than VABC. This means that the ability of VABC algorithm to get rid of local minima is very strong. On the other functions, the VABC algorithm is much better than the ABC algorithm. The VABC algorithm can increase the mean and the standard deviation of the functions by almost two orders of magnitude than ABC algorithm. The PSO converges very slowly and its performance is very bad on f_3, f_4, f_5 and f_6.

In order to show the performance of the VABC algorithm more clearly, the graphical representations of the results in Table II are reproduced in Figures 2-7. From the figures, we are concluded that the speed of convergence of VABC is much faster than ABC and PSO algorithms on all functions. From Figure 4, we can observe that the ABC algorithm is easy trapped at local optimum and the VABC algorithm is able to continue improving its solution on these two functions. For PSO algorithm, we can see that PSO algorithm is easy trapped in local optimum on f_3, f_4, f_5 and f_6. The performance PSO algorithms deteriorate in optimizing these functions

TABLE II
Units for Magnetic Properties Results comparison of different optimal algorithms for 30 runs

30D		VABC	ABC	PSO
f_1	Mean	6.7374e-016	1.1396e-014	2.8575e-008
	Std	9.4347e-017	8.0826e-015	3.8261e-008
f_2	Mean	1.7060e-001	3.3325e-001	2.6555e+001
	Std	1.7242e-001	2.3784e-001	1.7080e+001
f_3	Mean	1.0191e-008	6.0208e-007	4.1956e+002
	Std	5.3151e-008	3.2419e-006	2.9238e+001
f_4	Mean	9.1437e-007	1.8603e-001	4.6671e+001
	Std	4.8177e-006	3.6710e-001	1.2656e+001
f_5	Mean	4.1871e-008	6.0643e-006	4.2520e+000
	Std	1.6534e-008	3.5254e-006	8.3370e-001
f_6	Mean	9.4874e+001	1.9897e+002	9.5252e+003
	Std	8.4636e+001	1.1697e+002	3.7111e+002

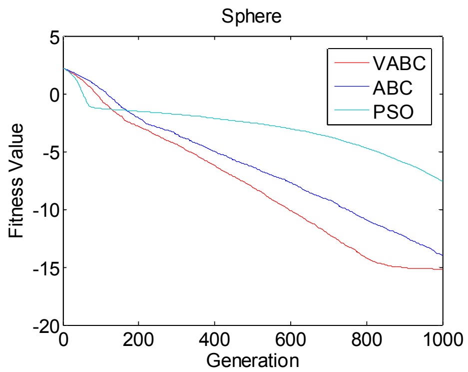

Fig. 2. The median convergence characteristics of Sphere function.

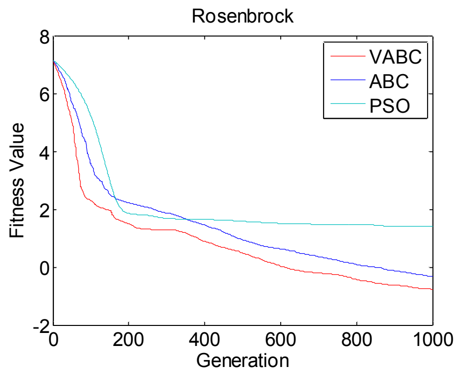

Fig. 3. The median convergence characteristics of Rosenbrock function.

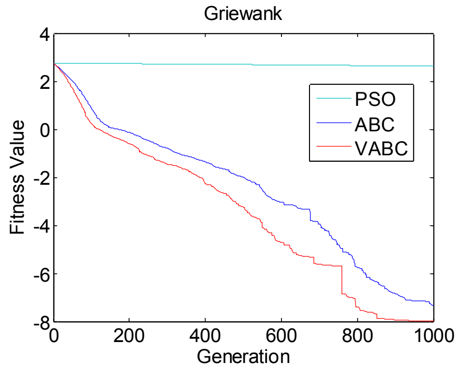

Fig. 4.The median convergence characteristics of Griewank function.

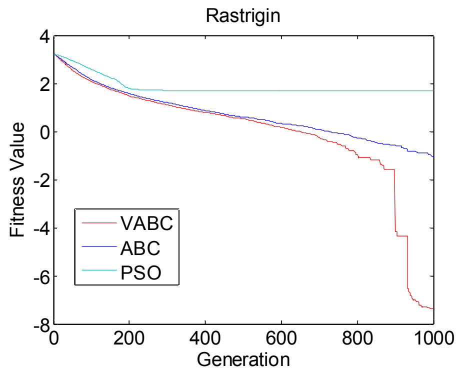

Fig. 5.The median convergence characteristics of Rastrigin function.

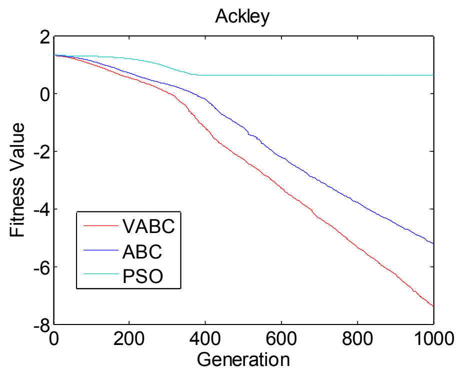

Fig. 6.The median convergence characteristics of Ackley function.

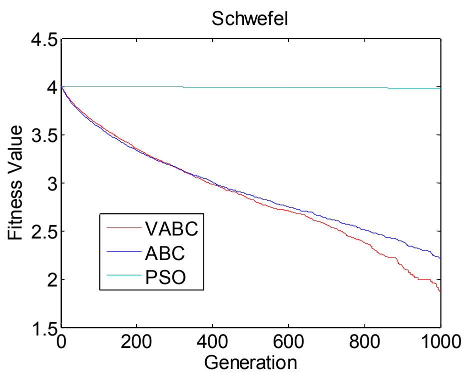

Fig. 7.The median convergence characteristics of Schwefel function.

5 Data clustering

5.1 K-means algorithms

As mentioned above, the goal of data clustering is grouping data into a number of clusters and k-means algorithm is the most popular clustering algorithm. In this section, we briefly describe the k-means algorithm. Let X=(x_1, x_2,⋯, x_n) be a set of n data and each data vector is a p-dimensional vector. Let C={c_1, c_2,⋯, c_k} be a set of K clusters and K denotes the number of cluster centroids which is provided by the user. In k-means algorithm, firstly, randomly initialize the K cluster centroid vectors and then assign each data vector to the class with the closest centroid vector. In this study, we will use Euclidian metric as a distance metric. The expression is given as follows:

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^{P} (x_{ik} - c_{jk})^2} \quad (4)$$

After all data being grouped, recalculate the cluster centroid vectors by using:

$$c_j = \frac{1}{n_j} \sum_{\forall x_i \in c_j} x_i \quad (5)$$

where n_j is the number of data vectors which belongs to cluster j. After the above process, reassign the data to the new cluster centroids and repeat the process until a criterion is satisfied. In this study, the criterion is when the maximum number of iterations has been exceeded. To know whether the partition is good or not, a measure for partition must be defined. A popular performance function for measuring goodness of the partition is the total within-cluster variance or the total mean-square quantization error (MSE) [18][19], which is defined as follows:

$$Perf(X,C) = \sum_{i=1}^{N} Min\{\| X_i - C_l \|^2 | l = 1,...,K\} \quad (6)$$

Because k-means algorithm is sensitive to the initial states and always converges to the local optimum solution, more population-based stochastic search algorithms are presented. In this paper, we will use VABC algorithm to solve clustering problem.

5.2 VABC algorithms on clustering

In the VABC algorithm, each individual represents a solution in K dimensional space. The number of dimension is equal to the number of clusters. Each component of an individual represents a cluster centroid and each cluster centroid is a p-dimensional vector. In the initialization phase, we use maximum and minimum value of each component of the data set (which is to be grouped) as VABC algorithm individuals' initialization range. And initial solution is randomly generated in this range. We use the expression (6) to calculate the fitness function of individuals. Here the main steps of the fitness function are given below:

Main steps of the fitness function

For data vector x_i
    Calculate the Euclidean distance by using (4)
    Assign x_i to the closest centroid cluster c_j.
    Calculate the measure function using equation (6)
EndFor.
Return value of the fitness function.

5.3 Clustering experimental results

To evaluate performance of the proposed VABC approach for clustering, we compare the results of the ABC and PSO clustering algorithms using five different data sets. They are Iris, Wine, Contraceptive Method Choice, Wisconsin breast cancer and Ripley's glass, which are selected from the UCI machine learning repository [14]. The detailed description of the test datasets can be seen [20].

For every data set, each algorithm is applied 30 times individually with random initial solution. The parameters of all algorithms are set like the section 4. Table III summarizes the intra-cluster distances, as defined in Eq (6), obtained from all algorithms for the data sets. The average and standard deviation are presented.

TABLE III
COMPARISON OF INTRA-CLUSTER DISTANCES FOR THE THREE CLUSTERING ALGORITHMS

30D		VABC	ABC	PSO
Iris	Mean	**9.4607e+001**	9.4607e+001	9.7526e+001
	Std	1.0551e-003	7.7734e-003	4.4576e+000
Wine	Mean	1.6302e+004	**1.6298e+004**	1.6372e+004
	Std	1.4904e+000	6.2411e+000	1.0718e+002
CMC	Mean	**5.6952e+003**	5.6954e+003	5.7293e+003
	Std	1.6994e+000	1.3824e+000	4.0245e+001
Cancer	Mean	**2.9644e+003**	2.9644e+003	2.9656e+003
	Std	2.2817e-003	1.0731e-002	2.2730e+000
Glass	Mean	**2.2403e+002**	2.2539e+002	2.5881e+002
	Std	7.4652e+000	1.2685e+001	1.4607e+001

From the values in Table III, we can conclude that the results obtained by VABC are as good as ABC algorithm and it is clearly much better than the PSO algorithm for all data sets. For all test data set, except for Wine, the VABC found solutions smaller. Especially in Cancer and Glass data set, we can see from Table III the mean value of VABC is one order of magnitude better than that of ABC. From the mean value of VABC, we can conclude that the VABC algorithm is able to converge to the global optimum in all of runs.

6 Conclusion

This paper, based on the Von Neumann topology structure, a novel Article Bee Colony (ABC) algorithm is presented, namely Von Neumann Neighborhood Article Bee Colony (VABC). This resulted in a significant improvement in the performance in terms of solution quality, convergence speed and robustness.

In order to demonstrate the performance of the VABC algorithm, we compared the performance of the VABC with those of ABC and PSO algorithms on a set of benchmark functions. From the simulation results, it is concluded that the proposed algorithm has the ability to attain the global optimum and get rid of local minima, moreover it definitely outperforms the original ABC.

Because the VABC algorithm can be efficiently used for multivariable, multimodal function optimization, we apply it to solve clustering problems. The algorithm has been tested on several well-known real data sets. To evaluate the performance of the VABC algorithm on clustering problems, we compare it with the original ABC and PSO. From the experimental results, we can see that the proposed optimization algorithm is better than the other algorithms in terms of average value and standard deviations of fitness function. Therefore VABC can be considered a viable alternative to solve multivariable, multimodal optimization problems.

7 Acknowledgment

This work is supported by the 863 Hi-Tech research and development program of China under contract No.2008AA04A105.

8 References

[1] D. Karaboga, An Idea Based on Honey Bee Swarm for Numerical Optimization, Technical Report-TR06, Erciyes University, Engineering Faculty, Computer Engineering Department, 2005.

[2] D.Karaboga and B.Basturk, On the performance of artificial bee colony(ABC) algorithm, Applied Soft Computing 8(2008), pp.687-697, 2008

[3] D. Karaboga, B. Basturk, Artificial Bee Colony (ABC) Optimization Algorithm for Solving Constrained Optimization Problems, LNCS: Advances in Soft Computing: Foundations of Fuzzy Logic and Soft

Computing, Vol: 4529/2007, pp: 789-798, Springer-Verlag, 2007, IFSA 2007.

[4] D. Karaboga, B. Basturk Akay, Artificial Bee Colony Algorithm on Training Artificial Neural Networks, Signal Processing and Communications Applications, 2007. SIU 2007, IEEE 15th. 11-13 June 2007, Page(s):1 - 4, doi: 10.1109/SIU.2007.4298679

[5] C. Ozturk, D. Karaboga, Classification by Neural Networks and Clustering with Artificial Bee Colony (ABC) Algorithm,6th INTERNATIONAL SYMPOSIUM on INTELLIGENT and MANUFACTURING SYSTEMS "FEATURES, STRATEGIES AND INNOVATION", October 14-17, Sakarya, Turkey.

[6] R. Mendes, J. Kennedy, and J. Neves, " The fully informed particle swarm: Simpler, maybe better", IEEE Trans. Evol. Comput, vol. 8, no. 3, pp. 204-210, 2004.

[7] J. Kennedy and R. Mendes. "Population structure and particle swarw performance." in Proceedings of the IEEE Congress on Evolutionary Computation. (Honolulu. Hawaii USA). May 2002.

[8] J.Han, M.Kamber, Data Mining: Concepts and Techniques, Academic Press, 2001.

[9] P.S.Shelokar, V.K.Jayaraman, B.D.Kulkarni, An ant colony approach for clustering, Analytica Chimica Acta 509 (2) (2004) 187–195.

[10] Y.Kao, K.Cheng, An ACO-based clustering algorithm, in:M.Dorigo, etal. (Eds.), ANTS, LNCS4150, Springer, Berlin, 2006, pp.340–347.

[11] M.Omran, A.Engelbrecht, A.Salman, Particle swarm optimization method for image clustering, Int.J.PatternRecogn. Artif. Intell. 19 (3) (2005) 297–322.

[12] V.D. Merwe, A.P. Engelbrecht, Data clustering using particle swarm optimization, in: Proceedings of IEEE Congress on Evolutionary Computation 2003 (CEC2003), Canbella, Australia, 2003, pp. 215–220

[13] D.Karaboga and C.Ozturk, A novel clustering approach: Artificial Bee Colony(ABC) algorithm, Applied Soft Computing 2010

[14] C.L. Blake and C.J. Merz. UCI Repository of Machine Learning Databases, Available from: http://archive.ics.uci.edu/ml/datasets.html.

[15] Akay, B., Karaboga, D.: Parameter tuning for the artificial bee colony algorithm. In Proceeding of the First International Conference, ICCCI 2009, Wroclaw, Poland (2009)

[16] CHEN Zi yu, HE Zhong shi, ZHANG Cheng,: "Image multi-level thresholds based on PSO with Von Neumann neighborhood," Application Research of Computers, vol 26, no 5, pp.1977-1979, 2009

[17] Y.Shi and R.C.Eberhart, "Empirical study of particle swarm optimization," in Proceedings of the IEEE Congress on Evolutionary Computation(CEC'99), vol.3, pp.1945–1950, Piscataway, NJ, USA, 1999.

[18] K.R.Zalik, An efficient k-means clustering algorithm, Pattern Recognition Letters 29(2008) 1385–1391.

[19] Zulal Gungor and Alper Unler, K-harmonic means data clustering with simulated annealing heuristic, Appl. Math. Comput. (2006).

[20] T.Niknam, B.Bahmani Firouzi, M.Nayeripour, An efficient hybrid evolutionary algorithm for cluster analysis, World Applied Sciences Journal4(2) (2008)300–307.