Volume 4, Issue 4, April 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

# Optimization of Clustering Problem Using Population Based Artificial Bee Colony Algorithm: A Review

**Twinkle Gupta**
*Master of Technology (Dept. Of Computer Science and Engineering) GJUS&T, Hisar, Haryana, India*

**Dharmender Kumar**
*Associate Professor (Dept. Of Computer Science and Engineering) GJUS&T, Hisar, Haryana, India*

***Abstract***— ***The Artificial Bee Colony (ABC) algorithm is a population based meta-heuristic algorithm proposed by Karaboga in 2005 inspired by the intelligent foraging behaviour of honey bees. ABC as a powerful technique is easy to implement, has fewer control parameters, and could easily be integrated with other meta-heuristic algorithms due to which it continues to attract the interest of researchers from various fields around the world. Interestingly, ABC has been successfully applied to a wide variety of optimization problems. Clustering is an unsupervised classification mechanism where a data, usually multidimensional is classified into groups called clusters such that members of one group are similar according to a predefined criterion like minimizing square error criterion. The aim of this paper is to provide an extensive (not exhaustive) overview of modification to the original ABC and its application in solving Clustering problems with the expectation that it would serve as a reference material to both old and new, incoming researchers in this field.***

***Keywords— Artificial Bee Colony Algorithms, Clustering, Nature-Inspired Meta-heuristics, Optimizations, Swarm Intelligence Algorithms.***

## I. INTRODUCTION

Data mining has been widely used and unifies research in fields such as computer science, statistics, databases, machine learning and Artificial Intelligence. Different techniques that also fit in this category include association learning, data classification and clustering as well as regression [1]. Clustering is one of its main applications used to find the clusters in a set of data [2]. There is a problem of local optima in Clustering. Hence, to reach at global optima it is integrated with Swarm Optimization algorithms. Evolutionary algorithms (EA) and swarm optimization algorithms are the two optimization algorithms that have been developed based on nature-inspired concepts. EA attempts to simulate the phenomenon of natural evolution [3] in which each species search for beneficial adaptations in an ever changing environment. EA includes algorithms like Genetic algorithms (GA) and differential evolution (DE) algorithms. However, a swarm can be understood as any collection of interacting individuals like Ant colony, a swarm whose individual agents are ants. The intelligence of the swarm lies in the ways of interactions between simple agents, and between agents and the environment [4]. Particle Swarm Optimization (PSO) models on social behaviour of bird flocking or fish schooling [5], Ant colony optimization (ABO) based on a swarm of ants algorithm and Artificial Bee Colony (ABC) algorithm [6] based on foraging behaviour of honey bees are the examples of swarm optimization algorithm. Some important characteristics of ABC algorithm which makes it more attractive than other optimization algorithms:

- Employs only three control parameters (population size, MCN and limit) [15].
- It has fast convergence speed.
- Quite simple, flexible and robust [7], [8].
- It can easily be integrated with other optimization algorithms.

## II. CLUSTERING

Clustering is an unsupervised classification mechanism used to find the clusters in a set of data such that members of one cluster are similar to each other and dissimilar to members of other clusters based on some criterion. Clustering algorithms are often useful in various fields like data mining, learning theory, pattern recognition, etc. Clustering algorithms can be divided into various categories like partition-based algorithms, hierarchical-based algorithms, density-based algorithms and grid-based algorithms [2].

### *A. Partitioning Clustering Algorithm*

It splits the data objects into k partitions, where each partition represents a cluster. The partition is done based on a certain objective function like minimizing square error criterion which is computed as, $E = \Sigma \mid p - m_i \mid^2$ where, p is the data object in a cluster and $m_i$ is the mean of that cluster. It includes various methods: K-Means method, Bisecting K-Means method, Medoid method, PAM (Partitioning around Medoid), CLARA (Clustering Large Applications) and the Probabilistic Clustering [9]. Procedure of K-Means Clustering is shown in Fig.1.

 

1) Choose k data objects as initial centroid,
**2) Repeat**
3) Assign each object to the closest cluster center,
4) Recomputed the cluster centers of each cluster,
**5) Until** convergence criterion is met.

Fig. 1 K-Mean Clustering Procedure

There are several drawbacks of K-Means Algorithm like number of cluster K must be defined beforehand, sensitive to initial condition, may be trapped into local optima and can be applied only when the mean of a cluster is defined.

*B. Hierarchical Clustering Algorithm*

Hierarchical clustering is a process of dividing the similar dataset by constructing a hierarchy of clusters. This method is based on the connectivity approach based clustering algorithms. It can be classified as:

*1) Agglomerative hierarchical clustering:* It is a bottom-up clustering method in which clusters have sub-clusters, which in turn have sub-clusters and so on. It starts by letting each object has its own cluster and iteratively merges the closest clusters into larger and larger one, until all the objects form a single cluster or some termination condition is satisfied [10].

*2) Divisive hierarchical clustering:* It is a top-down clustering method, commonly not used. It works in a similar way to agglomerative method, but in the opposite direction. It starts with a single cluster containing all objects and then iteratively splits resulting clusters into smaller and smaller one until clusters of individual objects remain [10].

*C. DBSCAN Clustering Algorithm*

DBSCAN stand for Density Based Spatial Clustering of Application with Noise. It is based on the concept of density reachability and density connectability, both of which depends upon two input parameters- size of epsilon neighbourhood and minimum terms of local distribution of nearest neighbours. It starts with the randomly chosen starting point not visited previously. The point's neighbourhood is calculated, and if it contains sufficiently many points, a cluster formation is started otherwise the point is labelled as noise.

*D. Grid Density based Clustering Algorithm*

This method uses the multi-resolution grid data structure and use dense grids to form clusters. It first divided the original data space into a finite number of cells by forming the grid structure and then performed all the operations on this space. Its main advantage is the fastest processing time as data points those fall into the similar cell, will be treated as a single point. This method includes STING, Wave Cluster, and CLIQUE [11].

## III. ARTIFICIAL BEE COLONY ALGORITHM

The Artificial Bee Colony algorithm was introduced by Karaboga in 2005 for optimizing numerical problems [6], which consist of three groups: employed bees, onlooker bees and scouts. The bee carrying out search randomly is known as a scout. The bee going to the food source visited by it before and sharing its information with others bees is known as employed bee and the bee waiting on the dance area is known as onlooker bee. The onlooker bee and scout also called unemployed bee. ABC algorithm called the collective intelligence searching model consists of three essential components: Employed bees, Unemployed bees and Food sources. The employed and unemployed bees search for the rich food sources around their hive. Analogously, in the view of optimization context, the number of food sources in ABC represents the population of possible solutions. The position of a good food source indicates the position of a better solution to the given optimization problem. The quality of nectar of a food source represents the fitness cost of the associated solution (value of the benchmark function).

*A. Initialization Phase*

In this phase, a randomly distributed initial food source position of SN solutions is generated, where SN denotes the size of employed bees or onlooker bees. Each solution $x_i$ where i=1, 2. . . SN is a D-dimensional vector represents the number of optimized parameters. The initial food sources are randomly produced using the following expression:

$$x_{i,j} = x_{min,j} + rand\ (0,1)\left(x_{max,j} - x_{min,j}\right) \quad (1)$$

Where, $x_{max}$ and $x_{min}$ are the upper and lower bound of the parameter $x_i$ , respectively and j denotes the dimension.

And then evaluate each nectar amount $fit_i$.

*B. Employed bees Phase*

In this phase, each employed bee $x_i$ finds a new food source $v_i$ in its neighbourhood. The new food source is calculated using the following expression:

$$x_{i,j}(t+1) = x_{i,j}(t) + \emptyset\left(x_{i,j}(t) - x_{k,j}(t)\right) \quad (2)$$

Where, $x_i$ : Position of the onlooker bee,

t : Cycle number, $x_k$ : Randomly chosen employed bee,

j : Dimension of the solution and $\emptyset(\,)$ : A series of random variable in the range [-1, 1].

The fitness of food sources to find the global optimal is calculated by the following formula:

$$fit_m(x_m) = \begin{pmatrix} \frac{1}{1+f_m(x_m)} & ,f_m(x_m) > 0 \\ 1+|f_m(x_m)| & ,f_m(x_m) < 0 \end{pmatrix} \quad (3)$$

Where, $f_m(x_m)$ is the objective function value of $x_m$.

 

And then the new solution produced is compared with the current solution and memorizes the better one by means of a greedy selection mechanism.

*C. Onlooker bees Phase*

Employed bees share their information about food sources with onlooker bees waiting in the hive and then onlooker bees probabilistically choose their food sources depending on this information. A fitness based selection technique is used for this purpose, such as roulette wheel selection. The probability of a food source can be calculated using the following expression:

$$P_i = \frac{F(\theta_i)}{\sum_{k=1}^{S} F(\theta_k)} \quad (4)$$

Where, $P_i$ : Probability of selecting the $i^{th}$ employed bee,

S : Size of employed bees,

$\theta_i$ : Position of the $i^{th}$ employed bee and

$F(\theta_i)$ : Fitness value.

After that onlooker bees carried out randomly search in their neighbourhood. The new solution produced is compared with the current solution and the better one is memorized by means of a greedy selection mechanism similar to the employed bees.

*D. Scout bee Phase*

Employed bees whose solutions can't be improved through a predetermined number of cycles, called limit, become scouts and their solutions are abandoned. Then, they find a new random food source position using the following expression:

$$x_{i,j} = x_{j\ min} + r \cdot \left(x_{j\ max} - x_{j\ min}\right) \quad (5)$$

Where, r: A random number and lies between 0 and 1. .

These steps are repeated through a predetermined number of cycles called Maximum Cycle Number (MCN) or until a certain termination criterion is satisfied.

1) Cycle1
2) Initialize the food source positions $x_i$, 1, . . . SN
3) Evaluate the nectar amount fitness function $fit_i$ 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
End For
6) Calculate the probability values pi for the solution.
7) Onlooker Bees' Phase
For each onlooker bee choose a food source depending on pi
  Produce new food source positions $v_i$
  Calculate the value $fit_i$
  Apply greedy selection mechanism
End For
8) Scout Bee Phase
If an employed bee becomes scout, then replace it with a
new random source positions
9) Memorize the best solution achieved so far & cycle=cycle+1
**10) Until** cycle=MCN

Fig. 2 Artificial Bee Colony Procedure

## IV. VARIANTS OF ARTIFICIAL BEE COLONY ALGORITHM

ABC algorithm was introduced by Karaboga in 2005 for optimizing unconstrained numerical problems [6]. The performance of the ABC algorithm with the integration of Greedy Randomized Adaptive Search Heuristic and shift neighbourhood structures was investigated for a generalized assignment problem by Baykasoglu, Adil et al. [12]. The experimental results showed that this new ABC variant is very effective when applied to small and medium sized generalized assignment problems. Dervis Karaboga and Bahriye Basturk have used the algorithm for optimizing multivariable functions [13], multi-dimensional and multimodal numeric problems [14]. Dervis Karaboga and Bahriye Akay have used the algorithm for optimizing a large set of numerical test functions [15]. The authors have compared the results produced by ABC, Genetic Algorithm (GA), Particle Swarm Algorithm (PSO) and Particle Swarm Inspired Evolutionary Algorithm (PS-EA) and the performance of ABC was analyzed under the change of control parameter values [13] – [15]. Simulation results showed that the ABC algorithm performs better than the mentioned algorithms and can be efficiently used for solving multivariable functions and the multimodal engineering problems with high dimensionality. After implementing on simple functions, Vimal Nayak et al. [16] has described its implementation on complex benchmark functions like rastrigin, rosenbrock, sphere and schwefel. The analysis of the performance of the ABC algorithm was compared for the optimization of above benchmark functions with Particle Swarm Optimization (PSO).

Dervis Karaboga and Bahriye Basturk have extended the ABC algorithm for solving constrained optimization problems by using Deb's constrained handling method at the place of greedy selection method [17]. Milan Tuba et al. [18] proposed Adjusted ABC algorithm to handle various constrained problems in which a scout generated a new solution by adding the global experience information in the early stages of the execution and by adding the limited solution randomly between new lower and upper bounds for each cycle as the algorithm reaches to the final stage.

TSai, Pei-Wei et al. have proposed an enhanced ABC optimization algorithm named interactive ABC optimization algorithm to solve combinational optimization problem [19]. The authors have introduced the concept of universal gravitational force for the movement of onlooker bees to enhance the exploration ability of the ABC algorithm. A quick ABC algorithm was proposed by Karaboga and Gorkemli [20] so that the behavior of foragers of the ABC algorithm can be modeled more accurately. The quick ABC significantly improved the convergence speed and can be used for solving combinational optimization problems.

A hybridization of ABC with evolutionary programming called ABC programming (ABCP) for numerical benchmark functions was studied by X. Liu and Z. Cai [21]. The authors modified the selection strategy of ABC with randomized selections, hybridized a bit mutation operator into the components of ABC similar to that for differential mutation, in order to produce a new food source and the global search capability of the basic ABC was improved with Layer Noisy Crossover. Sharma and Pant were proposed a variant of ABC called RABC for solving the numerical optimization problem [22]. The authors applied three different modifications: colony size reduction mechanism during evolutionary process; modified the perturbation scheme and improvement in population diversity by using rank selection strategy. The performance was tested on various benchmark functions, where it was shown that ABCP and RABC performance was better than for traditional ABC.

In order to improve the population diversity and avoid the premature convergence, several algorithms have been developed. Li Bao et al. [23] and Malek Alzaqebah et al. [24] have used several selection strategies, such as disruptive selection strategy, tournament selection strategy and rank selection strategy and compared them. Parallelized ABC with ripple-communication strategy (PABC-RC) for solving the numerical optimization problem was proposed by Luo et al. [25] and provided more accuracy in finding the best solution. O. Kramer has combined the strategy of Powell and elements from evolutionary search in an ILS framework [26]. Chaotic search based ABC algorithm for solving the accuracy of global optimal value was proposed by Yan and Li [27]. The proposed algorithm was applied in PID control tuning. The results showed that CABC had high accuracy and the PID control systems adopted optimal parameters. Bin Wu and Cun hua Qian have proposed a differential ABC (DABC) which enhanced the diversity of populations and capability of global search [28]. H. Deng-xu and J. Rui-min [29] have proposed Cloud model-based Artificial Bee Colony Algorithm (CABC) and used positive cloud generators to generate the parameters ABC needed. This improved ABC algorithm is used to meliorate the drawback of prematurity of ABC, and applied to solve the logistics location problem.

Convergence rate was also one of the main parameters in which various researchers have done their work. To provide better convergence, Ali Hadidi et al. has proposed modifications in the neighbourhood searching method by considering the parameter of probability of changing a food source, in onlooker phase by using tournament selection, and in scout phase by using Gaussian mutation based approach [30] for the optimization of different typical truss structures under stress, displacement and buckling constraints. Anguluri Rajasekhar et al. [31] have introduced a neighbourhood search based on Cauchy and Gaussian mutation strategies to enhance the searching strategy. Junita Mohamad-Saleh et al. [32] have proposed Enhanced ABC algorithm (EABC) with two different modifications for the performance enhancement of ABC algorithm. The first modification enhanced the neighbourhood searching capability, capitalizes on the multiple global-best food source, whereas the second modification improved searching capability of the scout-bee by exploiting the so-far best-found possible-solution. The proposed algorithm was compared with various existing variants of the ABC algorithm on a wide set of high-dimensional benchmark functions. The results revealed that the proposed algorithm produced the significantly better convergence in comparison to the compared algorithms.

Milan TUBA et al. has proposed GABC which integrates ABC with self-adaptive guidance adjusted for engineering optimization problems [33] and T. K. Sharma has incorporated golden section search mechanism in the structure of basic ABC and proposed ILS-ABC algorithm [34]. T. Yong et al. [35] considered a modified ABC for numerical optimization with three modifications named $ABC_1$, $ABC_2$ and $ABC_3$. In $ABC_1$, new solutions were generated using two processes, considering the present and previous solutions. Similarly in $ABC_2$, the sensibility technique was used in selection of the solution by the onlooker while $ABC_3$ considered the integration of $ABC_1$ and $ABC_2$. When the performances and convergence rates of all these methods were compared with that for the classical ABC, it showed that the modified ABC enhanced the rate of convergence and global search capability. N. Gupta et al. [36] have added some conditions to select food sources by bees. So, if solutions were enough near to optimal solution, then further search exists around the food sources and lower and upper bounds of food sources can be replaced with smaller values related to last search. Finally, convergence speed of the MABC algorithm that is faster than the ABC algorithm was illustrated.

To deal with multi-objective optimization problems, the concept of Pareto dominance was used in MOABC by B. Zhang et al [37] for determining the flight direction of a bee. Ye Zhang et al. [38] have proposed a 'dynamic' artificial bee colony (D-ABC) algorithm with a dynamic 'activity' factor for solving multi-parameters optimization problems. The idea was that global searching should be dominant in early cycles and local searching should be primarily in the posterior cycles. Dr. Dharmender Kumar and Balwant Kumar [39] have proposed a rectangular topology based Artificial Bee Colony algorithm (RABC), significantly improved the original ABC in solving high dimensional optimization problems. The simulation results showed that the proposed RABC outperforms the original ABC in terms of convergence, robustness and solution quality.

To solve binary optimization problems, Kashan et al. have proposed a new modification to the original ABC called DisABC [40] in which a new differential expression was used that utilized a measure of dissimilarity between binary vectors and Pampara and Engelbrecht have utilized three new versions of ABC [41]. The authors proposed three adaptations: binary ABC (bin-ABC) which was based on the idea of binary PSO developed by J. Kennedy and R. C. Eberhart [42], normalized ABC (norm-ABC) based on the concept of normalized DE [43] and the angle-modulated ABC (AMABC), based on the angle-modulated PSO and DE [44], [45]. The authors evaluated the performance of these three variants on two sets of well-known binary optimization problems. The AMABC outperformed bin-ABC and norm-ABC and when compared with AMPSO and AMDE, their performances were comparable.

ABC was used in wide area of application. The application of ABC for dynamic deployment of stationary and mobile sensor networks was put forward by Ozturk et al. [46]. El-Abd has proposed the opposition based ABC algorithm for black box optimization benchmark data [47]. Discrete ABC (DABC) for blocking flow shop scheduling (BFS) with make span criterion was proposed by Zhang et al. [48]. Karaboga and Gorkemli [49] used combinatorial ABC (CABC) for traveling salesman problem (TSP). In this method, a new mutation operator from GSTM was integrated into the search phases of employed and onlooker for finding new neighborhood food sources. Karaboga et al. have been tested the performance of ABC in training of artificial neural networks [50] and also proposed the hybridization of ABC with Levenberq-Marquardt (LM) algorithm for it [51]. The authors used ABC at the initial stage of the search process, whereas the LM algorithm was used to continue the training process. S. Sobti and P. Singla [52] have proposed S-ABC and A. Rekaby et al. [53] proposed AABC algorithm which has a similar workflow to the ABC algorithm, but with different bees assignment techniques and dynamic bees' role distribution logic to solve the Travelling salesman problem.

The concept of adaptive population of food sources for ABC algorithm named ABC-SAC was proposed by Sharma et al. [54] which was further modified by including the elitism selection scheme and incorporating global-local exploration. An enhanced ABC called fast mutation ABC (FMABC) was proposed by Bi [55]. The authors modified the selection technique with pheromone and sensitivity model. The scout behavior was replaced by mutation strategy from opposition based learning. By analyzing the proposed algorithm and its results, the authors claimed that the proposed approach (ABC-SAC and FMABC) looks promising.

To achieve better performance there should be greater balance between exploration and exploitation. F. Kang et al. [56] have proposed a memetic algorithm which combines Hooke–Jeeves pattern search with ABC for numerical global optimization. Results showed that the new algorithm is promising in terms of convergence speed, solution accuracy and success rate. The integration of a Gaussian distribution with ABC, named GABC was performed by L. Dos Santos Coelho and P. Alotto [57] to obtain a balance between exploration and exploitation of the solution space. R. Murugan and M. R. Mohan [58] have proposed a Modified ABC algorithm (MABC) in which an employed bee firstly works out three new solutions by three different solution's search equations of I-ABC, ABC and GABC and then choose and determine the best one as the candidate solution. Ivona Brajevic and Milan Tuba [59] have introduced an upgraded artificial bee colony (UABC) algorithm for constrained optimization problems. The exploitation in the onlooker phase was improved by limiting new solutions only to the diagonal between two solutions and by decreasing MRO parameter, while at the same time diversity was added by a different selection probability formula that relaxes suppression of infeasible solutions. The exploration at the scout phase was increased by allowing more scouts periodically that helps to leave the local minima.

The effects of the perturbation rate that controls the frequency of a parameter change, the scaling factor that determines the magnitude of the change in parameters and the limit parameter on the performance of the ABC algorithm were investigated on real-parameter optimization by Bahriye Akay and Dervis Karaboga [60] for solving hybrid functions. Khaze, Seyyed Reza [61] has studied ABC to evaluate the efficiency of the algorithms, and concluded that it is more reliable in reaching the global optimized points in comparison to FA algorithm in solving the problems like functions of continuous optimization. Xu et al. proposed a new artificial bee colony (NABC) algorithm [62] by modifying the search pattern of both employed and onlooker bees. A solution pool was constructed by storing some best solutions of the current swarm and new candidate solutions were generated by searching the neighborhood of solutions chosen randomly from the solution pool. Brief summary of useful variants of ABC algorithm is shown in Table 1.

## V. Previous Works on Clustering Problem Using Artificial Bee Colony Algorithm

Clustering is one of the most useful applications that can be formulated as a multi-objective optimization problem. No. of researchers from different fields have analyzed clustering and its different techniques in solving Data Mining concepts, Benchmark data sets, TSP and Bioinformatics and compared their performances [11], [74], [75]. M. K. Sruthi et al. have focused on different similarity measures, view points and Document clustering [10]. Kanungo, Tapas, et al. have presented a simple and efficient implementation of the Lloyd's k-means clustering algorithm, called the filtering algorithm [76]. The algorithm is easy to implement and only requires that a kd-tree be built once for the given data points. J. Z. Huang et al. have presented W-k-means algorithm that can calculate variable weights automatically and these new weights were used in deciding the cluster memberships of objects in the next iteration [77]. The experimental results on both synthetic data and real data sets have shown that the filtering and the W-k-means algorithm outperform the k-means type algorithms in recovering clusters in data. C. Zhong et al. have proposed an improved Clustering method which combined two major stages split and merge stages with k-means method [78]. In the split stage, each cluster will be split into smaller clusters with k-mean repeatedly if they are sparse and the average distance is employed for merging standard in the merge stage. Experimental results on real and synthetic datasets demonstrated that the proposed clustering method detected clusters with different sizes, shapes and densities, robust to noises and also outperformed the traditional k-means and single-link clustering method.

R. Tang et al. [79] have found that the bio-inspired algorithms have their own virtues and could be logically integrated into K-means clustering to avoid local optima. The extended versions of clustering algorithms integrated with bio-inspired optimization methods produced improved results. X. Cui et al. have presented a Particle Swarm Optimization (PSO) document clustering algorithm [80] in order to generate more compact clustering results than the K-means algorithm. Artificial bee colony data miner for classification of data mining has been proposed by Celik et al. [81]. In recent years, ABC is one of the popular algorithms used for a variety of problems such as constrained optimization [17], in image processing [63], in Clustering [11], in engineering design [64] – [67], [100], in Bioinformatics field [68] – [70], in Scheduling [71] – [73] and many others.

Dervis Karaboga and Celal Ozturk used ABC in forming clusters of the benchmark classification problems for classification purpose [82] and Udgata et al. have presented sensor deployment in 3-d terrain using an ABC algorithm [83], [84]. This problem was modeled as data clustering problem in which centroid of each cluster represented the position of a sensor node to be deployed and optimal solution was obtained using the ABC algorithm. Results obtained showed that the ABC algorithm produced robust and good quality of solution. Abdulsalam et al. were presented a cluster deviation detection task using the ABC algorithm [85]. ABC was tested on three UCI benchmark datasets. The results showed that ABC performed well. Chunhua Ju and Chonghuan Xu have also proposed a novel collaborative filtering recommendation approach based on *K*-means clustering algorithm [86]. Y. Marinakis et al. have presented a new hybrid algorithm and combined the proposed Discrete ABC for the solution of the feature selection problem and a Greedy Randomized Adaptive Search Procedure (GRASP) algorithm for the solution of the clustering problem [87]. The performance of the algorithm showed that the clustering algorithm has the possibility to solve the clustering problem with a known or unknown number of clusters.

ABC is a very useful algorithm in solving the clustering problem, but convergence speed is yet the problem. Xin Sui et al. have presented an extended ABC algorithm, namely, the Cooperative Article Bee Colony CABC, which significantly improved the original ABC in solving clustering problems and some benchmark functions [88]. In the CABC algorithm, a super best solution vector, namely, gbest was set and its each component of D-dimensional was the best in all populations. X. Lei, X. Huang, and A. Zhang [89] added two weight factors and one random disturbance item in the iteration equation of the ABC algorithm, which changed the contractive way and adjust the searching space. The IABC algorithm was applied in gene expression data and PPI network clustering for the first time. The simulation results showed that the CABC and IABC approach performs better than the original ABC algorithm and other swarm intelligent algorithms in terms of accuracy, robustness, and convergence speed. A novel population based hybrid algorithm called Genetic Bee Tabu K-Means Clustering Algorithm (GBTKC) has been developed by Shafia, Mohammad Ali et al. [90] based on basic Honey Bee Algorithm. Also, the simplicity of K-Means, the diversity of Genetic Algorithm to find the global optimum and the advantages of Tabu Search has been combined in GBTKC. This was revealed that GBTKC is definitely a convergent optimal solution, and the quality of answers provided by the algorithm is more reliable than those of other algorithms.

Elena and Makhalova have considered the clustering based on fuzzy logic, named Fuzzy Clustering [91] and revealed that the results of the Fuzzy c-means clustering are more accurate in comparison with the results of the Hard c-means clustering as it gives the flexibility to express that data points can belong to more than one cluster. Zhexue Huang and Michael K. Ng extended the fuzzy k-means algorithm for clustering categorical data [92]. By using a simple matching dissimilarity measure for categorical objects and modes instead of a means for clusters, a new approach was developed. After that, Dervis Karaboga and Celal Ozturk tested the performance of the ABC on fuzzy clustering [93] to classify different data sets like Cancer, Diabetes and Heart from UCI database. The results indicated that the performance of Artificial Bee Colony Optimization Algorithm is successful in fuzzy clustering. Dutta et al. have also applied ABC to data obtained by an electronic nose, to distinguish between various scores of black tea [94].

Y. Zhang and L. Chen et al. firstly proposed a Variable string length Artificial Bee Colony (VABC) algorithm for variable length genotypes and then used it to make the fuzzy clustering, in order to find global optimal solutions and to automatically evolve the cluster number. Finally, the proposed VABC-FCM algorithm was used to automatically extract T–S fuzzy model [95]. Dr. Salima Ouadfel et al. [96] and Mishra et al. [97] have proposed a new fuzzy clustering algorithm based on a modified ABC, in which a new mutation strategy inspired from the DE was introduced in order to improve the exploitation process. The performance of proposed MoABC-FCM algorithm was compared with FCM, the classical ABC, MABC and PSO on a set of real images. The experimental results showed that the MoABC-FCM is effective and efficient.

M. Krishnamoorthi and A.M. Natarajan [98] have proposed an optimization method based on the swarm intelligence algorithm for the purpose of clustering. The significance of the proposed algorithm is that it used a Fuzzy C-Means (FCM) operator in the Artificial Bee Colony (ABC) algorithm and provided significant results in terms of the quality of solution. Asha Gowda Karegowda [99] has explained the application of evolutionary algorithm, namely Particle Swarm Optimization and Entropy based fuzzy clustering for identifying the initial centroid in order to enhance the performance of both k-means and fuzzy c-means clustering. Brief summary of this section is shown in Table 2.

 

## VI. DISCUSSION AND CONCLUSION

In this paper, various research articles in the domain of the ABC have studied. From the in depth literature survey, it is observed that a huge work is done by modifying the ABC algorithm to solve a variety of problems, including engineering design problems, scheduling problems, data mining problems like Clustering etc., but it was clear that some modifications to the original structure are still necessary in order to significantly improve its performance. New neighbour production mechanisms can be proposed to improve the convergence speed. New schemes for scout generation can be proposed for diversity improvement. New selection techniques can be proposed for distribution of onlookers to the food sources so that the performance of ABC can be enhanced. And also ABC can be used as an evolutionary framework into which different traditional or modern heuristic algorithmic components are integrated. In conclusion, ABC remains a promising and interesting algorithm, which would be used extensively by the researchers from different fields around the world.

## REFERENCES

[1] M.S.V.K. Pang-Ning Tan, "Data mining," in *Introduction to data mining*, Pearson International Edition, 2006, pp.2-7.

[2] O. A. Abbas, "Comparisons Between Data Clustering Algorithms", *The International Arab Journal of Information Technology (IAJIT)*, vol. 5, no. 3, 2008.

[3] P. J. Angeline, J. B. Pollack and G.M. Saunders, "An evolutionary algorithm that constructs recurrent neural networks," *Neural Networks in IEEE Transactions on*, vol. 5, no. 1, 1994, pp. 54-65.

[4] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of IEEE international conference on neural networks*, 1995, vol. 4, pp. 1942–1948.

[5] E. Bonabeau, M. Dorgio, and G. Theraulaz, "Swarm intelligence: from neural network to artificial intelligence," *NY: oxford university press*, New York, 1999.

[6] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," *Techn.Rep. TR06, Erciyes Univ. Press*, Erciyes, 2005.

[7] R. S. Rao, S. V. L. Narasimham, and M. Ramalingaraju, "Optimization of distribution network configuration for loss reduction using artificial bee colony algorithm," *International Journal of Electrical Power and Energy Systems Engineering*, vol. 1, no.2, pp. 116–122, 2008.

[8] A. Singh, "An artificial bee colony algorithm for the leaf-constrained minimum spanning tree problem," *Applied Soft Computing*, vol. 9, no. 2, pp. 625–631, Mar. 2009.

[9] Sunita Jahirabadkar and Parag Kulkarni, "Clustering for High Dimensional Data: Density based Subspace Clustering Algorithms," *International Journal of Computer Applications*, vol. 63, no.20, 2013, pp. 29-35.

[10] M. K. Sruthi and M. B. V. Reddy, "Document Clustering on Various Similarity Measures," *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*, vol. 3, no. 8, 2013, pp. 1269-1273.

[11] N. K. Kaur Mann, "Review Paper on Clustering Techniques," *Global Journal of Computer Science and Technology*, vol. 13, no. 5, 2013.

[12] Baykasoglu, Adil, Lale Ozbakir, and Pınar Tapkan, "Artificial bee colony algorithm and its application to generalized assignment problem," *Swarm Intelligence: Focus on Ant and particle swarm optimization*,2007, pp. 113-144.

[13] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm," *Journal of global optimization*, vol. 39, no. 3, pp. 459–471, 2007.

[14] D. Karaboga and B. Basturk, "On the performance of artificial bee colony (ABC) algorithm," *Applied soft computing*, vol. 8, no. 1, pp. 687–697, 2008.

[15] D. Karaboga and B. Akay, "A comparative study of artificial bee colony algorithm," *Applied Mathematics and Computation*, vol. 214, no. 1, pp. 108–132, 2009.

[16] V. Nayak, H. A. Suthar, and J. Gadit, "Implementation of Artificial Bee Colony Algorithm," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 1, no. 3, pp.112–120, September 2012.

[17] D. Karaboga and B. Basturk, "Artificial bee colony (ABC) optimization algorithm for solving constrained optimization problems," in *Foundations of Fuzzy Logic and Soft Computing*, Springer, 2007, pp. 789–798.

[18] Milan Tuba, Nebojsa Bacanin, Nadezda Stanarevic, "Adjusted artificial bee colony (ABC) algorithm for engineering problems," *WSEAS TRANSACTIONS on COMPUTERS*, vol. 11, no. 4, pp. 111-120, Apr. 2012.

[19] TSai, Pei-Wei, et al. , "Enhanced artificial bee colony optimization." *International Journal of Innovative Computing, Information and Control* ,vol. 5, no. 12, 2009, pp.5081-5092.

[20] D. Karaboga and B. Gorkemli, "A quick artificial bee colony -qABC- algorithm for optimization problems," in *2012 International Symposium on Innovations in Intelligent Systems and Applications (INISTA)*, 2012, pp. 1–5.

[21] X. Liu and Z. Cai, "Artificial Bee Colony Programming Made Faster," in *IEEE Fifth International Conference on Natural Computation, 2009. ICNC '09*, 2009, vol. 4, pp. 154–158.

[22] T. K. Sharma, M. Pant, and J. C. Bansal, "Some modifications to enhance the performance of Artificial Bee Colony," in *2012 IEEE Congress on Evolutionary Computation (CEC)*, 2012, pp. 1–8.

[23] L. Bao and J. Zeng, "Comparison and analysis of the selection mechanism in the artificial bee colony algorithm," in *Hybrid Intelligent Systems, 2009. HIS'09. Ninth International Conference on*, 2009, vol. 1, pp. 411–416.

 

[24] M. Alzaqebah and S. Abdullah, "Comparison on the selection strategies in the artificial bee colony algorithm for examination timetabling problems," *International Journal of Soft Computing and Engineering (IJSCE)* ,vol. 1, no. 5, pp. 158–163, November 2011.
[25] R. Luo, T.-S. Pan, P.-W. Tsai, and J.-S. Pan, "Parallelized Artificial Bee Colony with Ripple-communication Strategy," in *2010 Fourth International Conference on Genetic and Evolutionary Computing (ICGEC)*, 2010, pp. 350–353.
[26] O. Kramer, "Iterated local search with Powell's method: a memetic algorithm for continuous global optimization," *Memetic Comp.*, vol. 2, no. 1, pp. 69–83, Mar. 2010.
[27] Yan, Gaowei, and C. Li, "An effective refinement artificial bee colony optimization algorithm based on chaotic search and application for pid control tuning," *Journal of Computational Information Systems,* vol. 7, no.9, 2011, pp. 3309-3316.
[28] B. Wu and C. hua Qian, "Differential Artificial Bee Colony Algorithm for Global Numerical Optimization.," *Journal of Computers*, vol. 6, no. 5, May 2011.
[29] H. Deng-xu and J. Rui-min, "Cloud model-based Artificial Bee Colony Algorithm's application in the logistics location problem," in *2012 International Conference on Information Management, Innovation Management and Industrial Engineering (ICIII)*, 2012, vol. 1, pp. 256–259.
[30] A. Hadidi, S. K. Azad, and S. K. Azad, "Structural optimization using artificial bee colony algorithm," in *2nd International Conference on Engineering Optimization, Lisbon, Portugal*, September 6-9, 2010.
[31] A. Rajasekhar, S. Das, B. K. Panigrahi, and M. K. Mallick, "Neighborhood search based artificial bee colony algorithm for numerical function optimization," in *Swarm, Evolutionary, and Memetic Computing*, Springer, 2012, pp. 232–239.
[32] A. G. Abro and J. Mohamad-Saleh, "Multiple-Gbest Guided Artificial Bee Colony Optimization Algorithm," *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 7, pp. 592–600, 2013.
[33] M. Tuba, N. Bacanin, and N. Stanarevic, "Guided artificial bee colony algorithm," in *Proceedings of the 5th European conference on European computing conference*, 2011, pp. 398–403.
[34] T. K. Sharma, "Improved Local Search in Artificial Bee Colony using Golden Section Search," *Journal of Engineering*, vol. 1, no. 1, pp. 14–19, Aug. 2012.
[35] Z. Dongli, G. Xinping, T. Yinggan, and T. Yong, "Modified Artificial Bee Colony Algorithms for Numerical Optimization," in *IEEE 2011 3rd International Workshop on Intelligent Systems and Applications (ISA)*, 2011, pp. 1–4.
[36] B. Asady, P. Mansouri, and N. Gupta, "The modify version of artificial bee colony algorithm to solve real optimization problems.," *International Journal of Electrical & Computer Engineering (2088-8708)*, vol. 2, no. 4, pp. 473-480, August 2012.
[37] W. Zou, Y. Zhu, H. Chen, and B. Zhang, "Solving multi objective optimization problems using artificial bee colony algorithm," *Discrete Dynamics in Nature and Society*, vol. 2011, 2011.
[38] Y. Yan, Y. Zhang, and F. Gao, "Dynamic artificial bee colony algorithm for multi-parameters optimization of support vector machine-based soft-margin classifier," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, pp. 1–13, 2012.
[39] D. Kumar and B. Kumar, "Optimization of Benchmark Functions Using Artificial Bee Colony (ABC) Algorithm," *IOSR Journal of Engineering*, vol. 3, no. 10, pp. 09-14, October 2013.
[40] M. H. Kashan, N. Nahavandi, and A. H. Kashan, "DisABC: A new artificial bee colony algorithm for binary optimization," *Applied Soft Computing*, vol. 12, no. 1, pp. 342–352, Jan. 2012.
[41] G . Pampara and A. P. Engelbrecht, "Binary artificial bee colony optimization," in *2011 IEEE Symposium on Swarm Intelligence (SIS)*, 2011, pp. 1–8.
[42] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *1997 IEEE International Conference on Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation*, 1997, vol. 5, pp. 4104–4108.
[43] A. P. Engelbrecht and G. Pampara, "Binary differential evolution strategies," in *IEEE Congress on Evolutionary Computation, 2007. CEC 2007*, 2007, pp. 1942–1947.
[44] G. Pampara, N. Franken, and A. P. Engelbrecht, "Combining particle swarm optimisation with angle modulation to solve binary problems," in *The 2005 IEEE Congress on Evolutionary Computation, 2005*, 2005, vol. 1, pp. 89–96.
[45] G. Pampara, A. P. Engelbrecht, and N. Franken, "Binary Differential Evolution," in *IEEE Congress on Evolutionary Computation, 2006. CEC 2006*, 2006, pp. 1873–1879.
[46] Ozturk, Celal, Dervis Karaboga, and Beyza Gorkemli, "Probabilistic dynamic deployment of wireless sensor networks by artificial bee colony algorithm,"*Sensors,* vol. 11, no.6, 2011, pp. 6056-6065.
[47] M. El-Abd, "Opposition-based Artificial Bee Colony Algorithm," in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, New York, NY, USA, 2011, pp. 109–116.
[48] Y.-Y. Han, J.-H. Duan, and M. Zhang, "Apply the discrete artificial bee colony algorithm to the blocking flow shop problem with makespan criterion," in *Control and Decision Conference (CCDC), 2011 Chinese*, IEEE, 2011, pp. 2131–2135.
[49] D. Karaboga and B. Gorkemli, "A combinatorial Artificial Bee Colony algorithm for traveling salesman problem," in *IEEE 2011 International Symposium on Innovations in Intelligent Systems and Applications (INISTA)*, 2011, pp. 50–53.

 

[50] D. Karaboga and B. Akay, “Artificial Bee Colony (ABC) Algorithm on Training Artificial Neural Networks,” in *Signal Processing and Communications Applications, 2007. SIU 2007. IEEE 15th*, 2007, pp. 1–4.

[51] C. Ozturk and D. Karaboga, “Hybrid Artificial Bee Colony algorithm for neural network training,” in *2011 IEEE Congress on Evolutionary Computation (CEC)*, 2011, pp. 84–88.

[52] S. Sobti and P. Singla, “Solving Travelling Salesman Problem Using Artificial Bee Colony Based Approach,” presented at the International Journal of Engineering Research and Technology, 2013, vol. Vol.2 - Issue 6 (June - 2013).

[53] A. Rekaby, A. A. Youssif, and A. Sharaf Eldin, “Introducing Adaptive Artificial Bee Colony algorithm and using it in solving traveling salesman problem,” in *Science and Information Conference (SAI), 2013*, 2013, pp. 502–506.

[54] Sharma, Tarun Kumar, Millie Pant, and V. P. Singh, "Artificial bee colony algorithm with self adaptive colony size," *Swarm, Evolutionary, and Memetic Computing*, Springer Berlin Heidelberg, vol. 7076, 2011, pp. 593-600.

[55] X. Bi and Y. Wang, “An improved artificial bee colony algorithm,” in *2011 3rd International Conference on Computer Research and Development (ICCRD)*,IEEE 2011, vol. 2, pp. 174–177.

[56] F. Kang, J. Li, and H. Li, “Artificial bee colony algorithm and pattern search hybridized for global optimization,” *Applied Soft Computing*, vol. 13, no. 4, pp. 1781–1791, Apr. 2013.

[57] L. dos Santos Coelho and P. Alotto, “Gaussian Artificial Bee Colony Algorithm Approach Applied to Loney’s Solenoid Benchmark Problem,” *IEEE Transactions on Magnetics*, vol. 47, no. 5, pp. 1326–1329, May 2011.

[58] R. Murugan and M. R. Mohan , “Modified Artificial Bee Colony Algorithm for solving economic dispatch problem ” *ARPN Journal of Engineering and Applied Sciences*, vol. 7, no. 10, October 2012.

[59] Brajevic and M. Tuba, “An upgraded artificial bee colony (ABC) algorithm for constrained optimization problems,” *Journal of Intelligent Manufacturing*, vol. 24, no. 4, pp. 729–740, 2013.

[60] B. Akay and D. Karaboga, “A modified artificial bee colony algorithm for real-parameter optimization,” *Information Sciences*, vol. 192, pp. 120–142, 2012.

[61] Khaze, Seyyed Reza, “EVALUATION THE EFFICIENCY OF ARTIFICIAL BEE COLONY AND THE FIREFLY ALGORITHM IN SOLVING THE CONTINUOUS OPTIMIZATION PROBLEM,” *International Journal* (2013).

[62] Y. Xu, P. Fan, and L. Yuan, “A Simple and Efficient Artificial Bee Colony Algorithm,” *Mathematical Problems in Engineering*, vol. 2013, Feb. 2013.

[63] C. Chidambaram and H. S. Lopes, “A new approach for template matching in digital images using an Artificial Bee Colony Algorithm,” in *World Congress on Nature Biologically Inspired Computing, 2009. NaBIC 2009*, IEEE, 2009, pp. 146–151.

[64] S. N. Omkar, J. Senthilnath, R. Khandelwal, G. Narayana Naik, and S. Gopalakrishnan, “Artificial Bee Colony (ABC) for multi-objective design optimization of composite structures,” *Applied Soft Computing*, vol. 11, no. 1, pp. 489–499, Jan. 2011.

[65] O. Safarzadeh, A. Zolfaghari, A. Norouzi, and H. Minuchehr, “Loading pattern optimization of PWR reactors using Artificial Bee Colony,” *Annals of Nuclear Energy*, vol. 38, no. 10, pp. 2218–2226, Oct. 2011.

[66] Bhagat Singh Prajapati and Laxmi Srivastava, “Multi-Objective Reactive Power Optimization Using Artificial Bee Colony Algorithm,” *in International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 2,no.1, 2012,pp. 126-131.

[67] B. Akay and D. Karaboga, “Artificial bee colony algorithm for large-scale problems and engineering design optimization,” *J Intell Manuf*, vol. 23, no. 4, pp. 1001–1014, Aug. 2012.

[68] C. M. V. Benítez and H. S. Lopes, “Parallel Artificial Bee Colony Algorithm Approaches for Protein Structure Prediction Using the 3DHP-SC Model,” in *Intelligent Distributed Computing IV*, M. Essaaidi, M. Malgeri, and C. Badica, Eds. Springer Berlin Heidelberg, 2010, pp. 255–264.

[69] D. L. González-Álvarez, M. A. Vega-Rodríguez, J. A. Gómez-Pulido, and J. M. Sánchez-Pérez, “Finding Motifs in DNA Sequences Applying a Multiobjective Artificial Bee Colony (MOABC) Algorithm,” in *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*, C. Pizzuti, M. D. Ritchie, and M. Giacobini, Eds. Springer Berlin Heidelberg, 2011, pp. 89–100.

[70] C.-J. Lin and S.-C. Su, “Using an efficient artificial bee colony algorithm for protein structure prediction on lattice models,” *International journal of innovative computing, information and control*, vol. 8, pp. 2049–2064, 2012.

[71] L. Wang, G. Zhou, Y. Xu, S. Wang, and M. Liu, “An effective artificial bee colony algorithm for the flexible job-shop scheduling problem,” *Int J Adv Manuf Technol*, vol. 60, no. 1–4, pp. 303–315, Apr. 2012.

[72] L. P. Jiang, “An Optimization Model for Aircraft Maintenance Scheduling Based on ABC Algorithm,” *Advanced Materials Research*, vol. 490–495, pp. 147–151, Mar. 2012.

[73] S.-W. Lin and K.-C. Ying, “Increasing the total net revenue for single machine order acceptance and scheduling problems using an artificial bee colony algorithm,” *J Oper Res Soc*, vol. 64, no. 2, pp. 293–311, Feb. 2013.

[74] R. Xu and I. Wunsch, D., “Survey of clustering algorithms,” *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 645–678, May 2005.

[75] A. Joshi and R. Kaur, “A Review: Comparative Study of Various Clustering Techniques in Data Mining,” *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*, vol. 3, no. 3, March 2013, pp. 55-57.

[76] Kanungo, Tapas, et al., "An efficient k-means clustering algorithm: Analysis and implementation," *Pattern Analysis and Machine Intelligence,* IEEE Transactions on , vol.24, no.7, 2002, pp. 881-892.


 

[77] J. Z. Huang, M. K. Ng, H. Rong, and Z. Li, “Automated variable weighting in k-means type clustering,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 657–668, May 2005.

[78] Y. Lin, T. Luo, S. Yao, K. Mo, T. Xu, and C. Zhong, “An improved clustering method based on k-means,” in *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 2012, pp. 734–737.

[79] R. Tang, S. Fong, X.-S. Yang, and S. Deb, “Integrating nature-inspired optimization algorithms to K-means clustering,” in *2012 Seventh International Conference on Digital Information Management (ICDIM)*, 2012, pp. 116–123.

[80] X. Cui, T. E. Potok, and P. Palathingal, “Document clustering using particle swarm optimization,” in *Proceedings 2005 IEEE Swarm Intelligence Symposium, 2005. SIS 2005*, 2005, pp. 185–191.

[81] M. Celik, D. Karaboga, and F. Koylu, “Artificial bee colony data miner (ABC-Miner),” in *2011 International Symposium on Innovations in Intelligent Systems and Applications (INISTA)*, IEEE, 2011, pp. 96–100.

[82] D. Karaboga and C. Ozturk, “A novel clustering approach: Artificial Bee Colony (ABC) algorithm,” *Applied Soft Computing*, vol. 11, no. 1, pp. 652–657, 2011.

[83] S. K. Udgata, S. L. Sabat, and S. Mini, “Sensor deployment in irregular terrain using Artificial Bee Colony algorithm,” in *World Congress on Nature Biologically Inspired Computing, 2009. NaBIC 2009*, IEEE, 2009, pp. 1309–1314.

[84] S. Mini, S. K. Udgata, and S. L. Sabat, “Sensor Deployment in 3-D Terrain Using Artificial Bee Colony Algorithm,” in *Swarm, Evolutionary, and Memetic Computing*, B. K. Panigrahi, S. Das, P. N. Suganthan, and S. S. Dash, Eds. Springer Berlin Heidelberg, 2010, pp. 424–431.

[85] M. F. Abdulsalam and A. A. Bakar, “A Cluster-Based Deviation Detection Task Using the Artificial Bee Colony (ABC) Algorithm,” *International Journal of Soft Computing*, vol. 7, no. 2, pp. 71–78, Feb. 2012.

[86] C. Ju and C. Xu, “A New Collaborative Recommendation Approach Based on Users Clustering Using Artificial Bee Colony Algorithm,” *The Scientific World Journal*, vol. 2013, November 2013.

[87] Y. Marinakis, M. Marinaki, and N. Matsatsinis, “A hybrid discrete artificial bee colony-GRASP algorithm for clustering,” in *Computers & Industrial Engineering, 2009. CIE 2009. International Conference on*, 2009, pp. 548–553.

[88] W. Zou, Y. Zhu, H.Chen, and X. Sui, “A clustering approach using cooperative artificial bee colony algorithm, ” *Discrete Dynamics in Nature and Society*, vol. 2010, 2010.

[89] X. Lei, X. Huang, and A. Zhang, “Improved artificial bee colony algorithm and its application in data clustering,” in *2010 IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA)*, 2010, pp. 514–521.

[90] Shafia, Mohammad Ali, Mohammad Rahimi Moghaddam, and Rozita Tavakolian, "A Hybrid Algorithm for Data Clustering Using Honey Bee Algorithm, Genetic Algorithm and K-Means Method," *Journal of Advanced Computer Science & Technology Research*, vol. 1, no. 2, 2011.

[91] Elena, Makhalova, "FUZZY C-MEANS CLUSTERING IN MATLAB," *The 7th International Days of Statistics and Economics*, Prague, September 19-21, 2013.

[92] Z. Huang and M. K. Ng, “A fuzzy k-modes algorithm for clustering categorical data,” *Fuzzy Systems,IEEE Transactions on*, vol. 7, no. 4, pp. 446–452, August 1999.

[93] D. Karaboga and C. Ozturk, “Fuzzy clustering with artificial bee colony algorithm,” *Scientific research and Essays*, vol. 5, no. 14, pp. 1899–1902, 2010.

[94] A. Dutta, B. Tudu, R. Bandyopadhyay, and N. Bhattacharyya, “Black tea quality evaluation using electronic nose: An Artificial Bee Colony approach,” in *2011 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, 2011, pp. 143–146.

[95] Z. Su, P. Wang, J. Shen, Y. Zhang, and L. Chen, “Convenient T–S fuzzy model with enhanced performance using a novel swarm intelligent fuzzy clustering technique,” *Journal of Process Control*, vol. 22, no. 1, pp. 108–124, Jan. 2012.

[96] S. Ouadfel and S. Meshoul, “Handling Fuzzy Image Clustering with a Modified ABC Algorithm,” *International Journal of Intelligent Systems and Applications (IJISA)*, vol. 4, no. 12, pp. 65-74, 2012.

[97] Mishra, Nripendra. "Fuzzy Image Clustering with ABC Algorithm," *International Journal of Engineering & Management Technology (IJEMT)*, vol. 1, no. 1, pp. 110-117, Mar. 2013.

[98] M. Krishnamoorthi and A.M. Natarajan , “Artificial Bee Colony (ABC) algorithm integrated with Fuzzy C-Mean operator for data clustering ” *Journal of Computer Science*, vol. 9, no. 4, 2013, pp. 404-412.

[99] Asha Gowda Karegowda, “Particle Swarm Optimization Algorithm Based *k*-means and Fuzzy*c*-means clustering”, *International Journal of Advanced Research in Computer Science and Software Engineering,* Vol. 3, no 7, July 2013, pp. 448-451.

[100] S. Okdem, D. Karaboga, and C. Ozturk, “An application of Wireless Sensor Network routing based on Artificial Bee Colony Algorithm,” in *2011 IEEE Congress on Evolutionary Computation (CEC)*, 2011, pp. 326–330.


 

TABLE I: Summary of Variants of ABC Algorithm

| **Parameters** | **ABC** | **MABC** | **IABC** | **UABC** | **Q-ABC** | **MABC** | **DABC** | **DisABC** |
|---|---|---|---|---|---|---|---|---|
| Description | A meta-heuristic algorithm inspired by intelligent foraging behaviour of honey bees. | Incorporated Deb's rule in selection mechanism. | Introduced the concept of universal gravitational force in onlooker bees' phase. | Introduced diagonal solutions, MRO parameter, and different selection probability. | Incorporated the new definition for onlooker bees. | Incorporated three new selection mechanisms. | Incorporated differential operator in employee and onlooker phase. | Used differential expression to measure dissimilarity between binary vectors. |
| Advantages | • Can be easily integrated with other optimization algorithms.<br>• Can be used to solve unconstrained optimization problem. | • Can be used for constrained optimization problems.<br>• Provide better performance | •Enhance exploration ability.<br>•Improved global search capability of bees.<br>•Increased the accuracy and provide better solution. | • Helps in avoiding local optima.<br>• Provide better solution. | • Improved convergence speed.<br>• Improved algorithm's exploitation. | • Improved population diversity.<br>• Avoid the premature convergence<br>• Provide global optima.<br>• High accuracy.<br>• Adopted optimal parameters. | • Avoid the premature convergence<br>• Better convergence rate.<br>• Balance b/w exploration &exploiatation is good.<br>• Can be used for high dimen-sional data. | • Can be used to solve binary optimization problem.<br>• Provide better performance |
| Disadvantage | • Can't be used for constrained optimization.<br>• Can be stuck into local optima. | • Can be stuck into local optima.<br>• Poor convergence rate. | • Poor convergence rate.<br>• Poor exploitation ability. | • Poor convergence rate.<br>• Poor exploitation ability. | • Can be stuck into local optima.<br>• Less accuracy. | • Can't be used for high dimen-sional data. | • Can't be used to solve binary optimization problem. | • Poor convergence rate. |
| Problem area | • Function optimization<br>SP<br>• ANN<br>• Image processing<br>• Engineering design<br>• Scheduling and Bio informatics field. | • Constrained problems.<br>• Truss structures. | • Benchmark functions.<br>• Combinational optimization problem. | • Constrained optimization<br>• Engineering benchmark problems. | •Combinational optimization problem.<br>•High dimensional & engg. benchmark functions.<br>•Blocking flow shop scheduling. | •Benchmark examination timetabling problem.<br>•Logistics location problem.<br>•PID control tuning.<br>•Unimodal problems. | •High dimensional benchmark functions.<br>•Multi-parameters optimization<br>•Solenoid design benchmark problem. | •Binary optimization<br>•uncapacitated facility location problem |
| Variants providing same | ------ | AABC, MABC | ABCP, PABC-RC | ------ | EABC, GABC, DABC | G-CMA-ES,CABC, MABC,IL | D-ABC, RABC, ABC- | bin-ABC, norm-ABC, |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| functions | | | | | | S-ABC | SAC, MABC, GABC | AMABC |
| References | [6], [13] – [16], [46],[50], [52],[61], [63] – [73], [100] | [17],[18], [30] | [19],[21], [25] | [59] | [20],[32], [33],[48] | [23], [24], [26],[27], [29],[31], [34] | [28],[38], [39],[54], [56],[57] | [40] – [45] |

TABLE II
SUMMARY OF CLUSTERING AND ITS INTEGRATION WITH ABC ALGORITHM

| **Parameters** | **Modified Clustering** | **Lloyd's k-means** | **DABC-GRASP** | **CABC** | **Fuzzy Clustering** | **Fuzzy k-mode** | **VABC-FCM** | **MoABC-FCM** |
|---|---|---|---|---|---|---|---|---|
| Description | Proposed a variety of similarity measures, view points and two related Clustering methods. | Used the concept of kd-tree, to be built once for the given data points. | Proposed Discrete ABC for e feature selection & GRASP algorithm for clustering. | Cooperative search technique is used for exchange of information. | Integrated the concept of fuzzy logic with Clustering to give more flexibility. | Extended the fuzzy k-means algorithm by using modes instead of means of objects. | Introduced the concept of variable length genotypes. | Introduced new mutation strategy inspired from DE for better exploitation. |
| Data need to be scanned | Several times | Only once | Several times | Few times | Several times | Several times | Few times | Few times |
| Need to define no. of clusters | Yes | Yes | No | Yes | Yes | Yes | No | Yes |
| Applicable only when mean is defined | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes |
| Length of genotypes | Fixed | Fixed | Fixed | Fixed | Fixed | Fixed | Variable | Fixed |
| Quality of solutions | Can be stuck into local optimum | Can be stuck into local optimum | Provide better solution | Provide better accuracy of solution | Provide good approx. of global solution. | Provide good solution | Provide global optimum solution | Provide good solution |
| Convergence rate | Slow | Slow | Slow | High | Slow | Slow | High | High |
| Membership degree of objects | {0,1} | {0,1} | {0,1} | {0,1} | [0,1] | [0,1] | [0,1] | [0,1] |
| Problem area | Document Clustering | Synthetic data and real data sets | Data Clustering & datasets from the UCI Machine Learning Repository | Data clustering& complex optimization benchmark functions | Data Clustering | Categorical Data Clustering | Fuzzy clustering & T–S fuzzy model | Fuzzy Clustering |
| References | [10] | [76] | [87] | [88] | [91] | [92] | [95] | [96], [97] |

 