A Survey on Artificial Bee Colony Models for Numerical Optimizations and its Work In Image Segmentation and Data Classification.

Lavanya Gunasekaran and Dr. Srinivasan Subramaniam

Research Scholar, Anna University Regional Centre, Madurai
Professor & Head, Department of Computer Science and Engineering, Anna University Regional Centre, Madurai

ABSTRACT

There are many problems related to optimization and many researchers are interested in solving these problems. Literature presents several search algorithms to find an item with specified properties from a search space defined by a mathematical formula or procedure. Recent works describe several optimization problems which were then solved by many optimization algorithms. The most recently introduced Artificial Bee Colony (ABC) is one of the swarm intelligence algorithms, based on the intelligent foraging behavior of a honey bee swarm. It uses less control parameters, and it can be efficiently used for solving multimodal and multidimensional optimization problems. ABC algorithm is now applied to defined-real time problems of the clustering process, medical image segmentation is done using clustering process by means of ABC algorithm. Data classification is the categorization of data for its most effective and efficient use. Data can be classified according to any criteria, not only relative importance or frequency of use. A well-planned data classification system makes essential data to find easily. ABC algorithm is now applied to classify the data effectively by training the neural network. This paper presents the review based on ABC models for various numerical optimizations and the significant works carried out so far by ABC models for image segmentation and data classification.

INTRODUCTION

Population-based optimization algorithms find near-optimal solutions to the difficult optimization problems by motivation from nature. Swarm Intelligence (SI) is a part of Population-based optimization algorithms and is quite popular among the researchers. The swarm intelligence's algorithm is very efficient for solving problems of optimization. In recent years, swarm intelligence becomes more and more attractive for the researchers, who work in the related research field. Swarm intelligence can be defined as the measure introducing the collective behavior of social insect colonies or other animal societies to design algorithms or distributed problem-solving devices. The requirement of Optimization is quiet essential as its primary intention is to minimize the effort required and to maximize the desired benefits.

The Bee Colony Optimization-based algorithm is a stochastic Meta heuristic that belongs to the class of swarm intelligence algorithms. Many studies based on various bee colony behaviors have been developed to solve complex combinatorial or continuous optimization problems. Bee colony optimization-based algorithms are inspired by the behavior of a honeybee colony that exhibits many features that can be used as models for intelligent and collective behavior. These features include nectar exploration, mating during flight, food foraging, waggle dance, and division of labor.

The self-organized and collective behavior of colony insects enables them to solve multi-objective problems that are beyond the capability of individual members functioning alone. In the case of honey bees, this behavior helps them to explore the environment in search of flower patches and then pass the information of food source to the other bees of the colony when they return to the hive. Artificial bee colony (ABC) algorithm is a reasonably novel member of swarm intelligence. In the ABC algorithm, the colony of artificial bees contains three groups of bees: employed bees, onlookers and scouts.

In a robust search process, exploration and exploitation processes must be carried out together. In the ABC algorithm, while onlookers and employed bees carry out the exploitation process in the search space, the scouts...
control the exploration process. The survival and progress of the bee colony are dependent upon the rapid discovery and efficient utilization of the best food resources. Similarly the successful solution of difficult engineering problems is connected to the relatively fast discovery of “good solutions” especially for the problems that need to be solved in real time.

Data classification is the process of organizing data into categories for its most effective and efficient use. A well-planned data classification system makes essential data easy to find and retrieve. Data Classification can be efficiently done using swarm based optimization algorithms.

The ABC algorithm make use the advantages such as, Global optimization strategy, Local optimization strategy (in employee bee), Random selection strategy (probability in onlooker) and Feedback strategy (scout limit). On the other hand, GA is a simple algorithm and the problem can be easily solved without requiring mathematical knowledge. But, even both the algorithm seems good in optimization, they also faces some challenges in their process. In the traditional GA, solution can have tendency to converge towards local minimum since the "better" solution is only in comparison with other solutions. On the other hand, ABC algorithm does not consider the secondary information about the problem so that convergence rate of optimization problem is very slow. So, the hybridization of both the algorithm leads to an even better solution since the advantages are boosted and disadvantages are compensated with other one. The combination of both the algorithm can easily solve the local minimum problem and also speeds up the convergence rate as the new solution generation is combined with GA. The secondary information problem of ABC algorithm is solved by incorporating the fractional calculus, which is one of the effective mathematical calculations based on differential operator. This leads to having better solutions and better optimization. A hybrid optimization algorithm is suggested by applying fractional order ABC and GA for optimization to solve the existing problems.

The paper is designed as follows; second section deals with review based on various models of ABC optimization algorithm; and the third section shows the application of ABC models in image segmentation; and the fourth section shows the application of ABC models in classification; fifth section shows the data classification using neural network and last section explores the research directions to be concentrated.

**ABC Models for Optimizations:**

Teodorovic (2001) developed a new system called Bees system (BS) based on food source searching as foraging behavior of bee colonies. It was tested through many instances of the traveling salesman problem (TSP).

Bianco (2011) presented mapping paradigm which takes advantage of bee searching and traveling as an original idea allowing precise large scale navigation. This latter is performed through the use of two distinct sets of landmarks: global landmarks guide roughly the agent to a place; with local landmarks the agent performs very precise motion to the final destination. The paradigm is equivalent to a map composed of two distinct levels: the agent like bee navigates from place to place following the global potential function. When the agent is close to the place of interest then it switches to a finer map about the specific place. Tests have demonstrated that such capabilities are sufficient to get rather good precision showing that biology offers simple but powerful models.

Nakrani and Tovey (2004) proposed a Honey Bee Algorithm on self-organization of honey bee colonies to allocate foragers among food source. It was applied to the dynamic allocation of internet services. Because of the many similarities between server and nectar collection, they modeled servers as foraging bees and HTTP request queues as flower patches. This algorithm was been compared against an omniscient algorithm that computes an optimal allocation policy, a greedy algorithm that uses past history to compute allocation policy and an optimal-static algorithm that computes omnisciently the best among all possible static allocation policy. The results supported the effectiveness of the algorithm, particularly in the highly dynamic and unpredictable Internet environment.

Sato and Hagiwara (1997) used bee foraging behavior to improve genetic algorithm. The proposed Bee System must carry out a global search to obtain superior chromosomes with pretty high fitness using the simple genetic algorithm. After that, it is the local search. All population chromosomes will be crossover with superior chromosomes in concentrated crossover manner. Next, the populations will migrate between them to transfer generation individual to the neighboring population. The authors enhanced the local search ability of their proposed algorithm using a pseudo-simplex method. This work was been commented on by noticing that the concentrated crossover and the pseudo-simplex method ensured the high ability for local search as lack of the genetic algorithm without degrading global search ability. Bee system possessed better performance against conventional GA especially, for high complex multivariate functions.

Quijano and Passino (2007) introduced a model of honey bee social foraging to solve the resource allocation problems as a numerical optimization. They reached to an ideal free distribution and globally optimal allocation strategy. There are many potential applications based on the marriage phenomenon among bees which represent the second class of bee colony algorithms and it is called as MBO.
Yang et al. (2007) based on MBO, proposed FMBO a faster algorithm which had as a main objective to make better the results. They introduced the generation in each time of drone as mating agent in randomly manner. Moreover, they mate with finite quantity of queens with the absent of the energy or speed factors using in MBO. Results presented more performance of FMBO against MBO and genetic algorithm. Particularly, it is the global convergence proved by the use of Markov chain theory, easy implementation, and adjustment of a few parameters and jumping of the local optimum.

The efficiency of the ABC with GA, PSO and PS-EA, which were swarm intelligence and population related methods as the ABC algorithm were suggested by Dervis Karaboga and Bahriye Basturk (2007). To obtain the performance of the ABC algorithm, PSO, PS-EA, GA and ABC algorithms, they examined on five high dimensional numerical benchmark functions that have multimodality. From the result it was proved that the introduced theory can get out of a local minimum and was efficiently utilized for multivariable, multimodal function optimization.

A hybrid optimization method related to Cat Swarm Optimization (CSO) and ABC was suggested by Pei-Wei Tsai et al. (2011). CSO, which was an optimization algorithm, its job was to solve numerical optimization problems, where as ABC was an optimization designed by stimulating the behavior of bees finding foods. By combining those two algorithms, a hybrid algorithm known as Hybrid Particle Cat Swarm Optimization Artificial Bee Colony (PCSOABC) was presented. Five benchmark functions were applied in order to calculate the accuracy, convergence, speed, and the stabilization of the Hybrid PCOSOABC.

Min-Yuan Cheng and Li-Chuan Lien (2012) introduced an optimization hybrid swarm algorithm the particle bee algorithm (PBA), that take knowledge from the intelligent swarming behavior of honeybees and birds, to integrate BA global search ability along with the local search advantages of PSO. Their experiment finds similarities regarding the performance of PBA with that of GA, DE, bee algorithm (BA) and PSO for multi-dimensional benchmark numerical problems. Also, their theory compared the performance of PBA with that of BA and PSO for practical construction engineering of Construction Site Layout (CSL) problem. The outcome proved that the performance of PBA was equally efficient as the above mentioned theories in the benchmark functions and was quiet well employed to solve hypothetical CSL problem with high dimensionality.

Subotic (2011) introduced a modification of artificial bee colony (ABC) algorithm for constrained optimization problems. Normally, more than one onlooker bee goes to a promising food source reported by employed bee. The proposed modification forms a mutant solution in onlooker phase utilizing three onlookers. Their procedure obtains better results than the original artificial bee colony algorithm. Their multiple onlooker modified approach was tested on the full set of 24 well known benchmark functions known as g–functions and proved to be more efficient than the pure ABC algorithm in most cases.

Gerhardt and Gomes (2012) proposed a method in which the performance of the Artificial Bee Colony (ABC) algorithm in engineering optimization problems is compared against those of other methods listed in the literature. The classic spring design optimization problem, and truss optimization on size and shape with frequency constraint problems were chosen for the numerical experiments. It is well known that algorithm performance is problem dependent. Taking advantage of its flexibility, and based on related works, some modifications were implemented in the ABC algorithm. The results presented here indicate that the ABC algorithm is an effective global optimizer with relative high computational cost. However, the performance is comparable to the state of the art metaheuristics algorithms. Therefore, the applicability of the ABC algorithm in engineering optimization problems is compromised with its cost-benefit function, by weighing the advantages against the disadvantages of its characteristic features.

Gao et al. (2008) proposed an upgraded Artificial Bee Colony (ABC) optimization algorithm, which is called the Interactive Artificial Bee Colony (IABC) optimization, for numerical optimization problems. The onlooker bee is designed to move in order to the picked coordinate indicated by the employed bee and evaluates the fitness values near it in the original Artificial Bee Colony algorithm in order to reduce the computational complexity. Hence, the exploration capacity of the ABC is constrained in a zone. Based on the framework of the ABC, the IABC introduces the concept of universal gravitation into the consideration of the affection between employed bees and the onlooker bees. By assigning different values of the control parameter, the universal gravitation should be involved for the IABC when there are various quantities of employed bees and the single onlooker bee. Therefore, the exploration ability is redeemed about on average in the IABC. Five benchmark functions are simulated in the experiments in order to com-pare the accuracy/quality of the IABC, the ABC and the PSO. The experimental results manifest the superiority in accuracy of the proposed IABC to other methods.

Dervis Karaboga and Bahriye Akay (2011) suggested a technique of a modified ABC algorithm for constrained optimization problems and compare the functionality of the modified ABC algorithm against those of state-of-the-art algorithms for a set of constrained test problems. Artificial Bee Colony (ABC) algorithm was firstly presented for unconstrained optimization problems, where that ABC algorithm showed superior performance. For constraint handling, ABC algorithm uses Deb’s rules consisting of three simple heuristic rules and a probabilistic selection scheme for feasible solutions based on their fitness values and infeasible solutions based on their violation values. ABC algorithm is tested on thirteen well-known test problems and the results
obtained are compared to those of the state-of-the-art algorithms and discussed. Furthermore, a statistical parameter analysis of the modified ABC algorithm was conducted and appropriate values for each control parameter are acquired using analysis of the variance (ANOVA) and analysis of mean (ANOM) statistics.

Guopu Zhu and Sam Kwong (2010) proposed a theory inspired by PSO, an improved ABC algorithm called gbest -guided ABC (GABC) algorithm by incorporating the information of global best (gbest) solution into the solution search equation to enhance the exploitation. Artificial bee colony (ABC) algorithm which was invented by Karaboga is a biological-inspired optimization algorithm, which has been shown to be competitive with some conventional biological-inspired algorithms, such as genetic algorithm (GA), differential evolution (DE) and particle swarm optimization (PSO). However, there is still an insufficiency in ABC algorithm regarding its solution search equation, which is good at exploration but poor at exploitation. Hence, experimental results of this new approach were tested on a set of numerical benchmark functions show that GABC algorithm can outperform ABC algorithm in most of the experiments.

**ABC Models in Image Segmentation:**

Taherdangkoo et al. (2010) suggested a method in which two parameters described below and used the artificial bee colony (ABC) algorithm to reduce the time and to reach a higher quality than that obtained by previous researches. Now, segmentation of medical images, particularly magnetic resonance images of brain is complex and it is considered as a huge challenge in image processing. Among the numerous algorithms presented in the context, the fuzzy C-mean (FCM) algorithm is widely used in MR images segmentation. Also, researchers have introduced two new parameters in order to improve the performance of FCM algorithm, which are calculated using neural network in a complex and time consuming manner. These two parameters have been then calculated by other researchers using genetic algorithm (GA) and particle swarm optimization (PSO) algorithm, which although it has reduced the time but no change obtained in the resulted segmentation quality. Finally, they segment real MR images with their proposed algorithm and compare it with previous presented algorithms.

Ming-Huwi Horng (2011) presented an upgraded multilevel MEt algorithm based on the technology of the artificial bee colony (ABC) algorithm; the maximum entropy based artificial bee colony thresholding (MEABCT) method. Multilevel thresholding is an important technique for image processing and pattern recognition. The maximum entropy thresholding (MET) has been widely applied in the literature. Four different methods are compared to their proposed method: the particle swarm optimization (PSO), the hybrid cooperative-comprehensive learning based PSO algorithm (HCOCLPSO), the Fast Otsu’s method and the honey bee mating optimization (HBMO). The experimental results demonstrate that the proposed MEABCT algorithm can search for multiple thresholds which are very close to the optimal ones examined by the exhaustive search method. Compared to the other four thresholding methods, the segmentation results of using the MEABCT algorithm is the most, however, the computation time by using the MEABCT algorithm is shorter than that of the other four methods.

**ABC Models in Classification:**

Fathian et al. (2007) developed an application of honeybee mating optimization in clustering (HBMK-means) based on (Abbass (2001) applied to data mining clustering. This study was tested on several datasets such as Iris dataset, wine dataset, and Wisconsin breast cancer database. It was compared against several typical stochastic algorithms: ant colony optimization, simulated annealing, genetic algorithm and tabu search algorithm. As result, authors concluded that is a viable and an efficient heuristic to find optimal or near optimal solutions and their results are very encouraging in terms of quality of solutions, average number of function evaluations and the processing time required.

Dervis Karaboga et al. (2007) suggested a method in which Artificial Bee Colony (ABC) Algorithm which has good exploration and exploitation capabilities in searching optimal weight set was used in training neural networks. Training an artificial neural network is an optimization task since it is desired to find optimal weight set of a neural network in training process. Orthodox training algorithms has some drawbacks such as getting stuck in local minima and computational complexity. Therefore, evolutionary algorithms were considered to train neural networks to overcome these issues.

Ramakrishnan Ramanathan (2006) suggested a simple classification scheme using weighed linear optimization. Inventory classification using ABC analysis is one of the most widely employed techniques in organizations. The need to consider multiple criteria for inventory classification has been stressed in the literature.

Peng Zhou et al. (2007) presented an extended version of the R-model for multi-criteria inventory classification. Ramanathan recently proposed a weighted linear optimization model for multi-criteria ABC inventory classification. Despite its many advantages, Ramanathan’s model (R-model) could lead to a situation where an item with a high value in an unimportant criterion is inappropriately classified as a class A item. This
model provides more reasonable and encompassing index since it uses two sets of weights that are most favorable and least favorable for each item.

Ching-Wu Chu et al. (2008) proposed a new inventory control approach called ABC–fuzzy classification (ABC–FC) that can handle variables with either nominal or non-nominal attribute, incorporate manager’s experience, judgment into inventory classification, and can be implemented easily. The objective of inventory management is to make decisions regarding the appropriate level of inventory. In practice, all inventories cannot be controlled with equal attention. The most widespread used inventory system is the ABC classification system, but the limitation of the ABC control system is that only one criterion is considered. Their ABC–FC approach is implemented based on the data of the Keelung Port. The results of their research show that 59 items are identified as very important group, 69 items as important group, and the remaining 64 items as unimportant group. When compared the results of ABC–FC with the original data, they found that our ABC–FC analysis shows a high accuracy of classification.

Data Classification using Neural Network:

Nii et al. (2007) presented Nursing-care Data Classification using Neural Networks. Nursing-care data in their work are Japanese texts written by nurses which consist of answers for questions about nursing-care. The nursing-care data are collected via WWW application from many hospitals in Japan. The collected data are stored into the database. The nursing-care experts evaluated the collected data to improve nursing-care quality. Currently, the collected data are evaluated by experts reading all texts carefully. It is difficult, however, for experts to evaluate the data because there are huge numbers of nursing-care data in the database. In their work, to reduce workloads for the evaluation of nursing-care data, neural networks are used for classifying nursing-care data instead of fuzzy classification system. They used standard three-layer feed forward neural networks with back-propagation type learning. First, they extracted attribute values (i.e., training data) from texts written by nurses. And then, they trained a neural network using the training data. From computer simulations, they showed the effectiveness of their proposed system using the leaving-one out method.

McIntire et al. (2002) classified the arctic sea ice, cloud, water and lead using neural networks and 1.6-μm data. Polar sea ice plays a critical role in regulating the global climate. Seasonal variation in sea ice extent, however, coupled with the difficulties associated with in situ observations of polar sea ice, makes remote sensing the only practical way to estimate this important climatic variable on the space and time scales required. Unfortunately, accurate retrieval of sea ice extent from satellite data is a difficult task. Sea ice and high cold clouds have similar visible reflectance, but some other types of clouds can appear darker than sea ice. Moreover, strong atmospheric inversions and isothermal structures, both common in winter at some polar locations, further complicate the classification. In their technique, they used a combination of feed-forward neural networks and 1.6-μm data from the new Chinese Fengyun-1C satellite to mitigate these difficulties. The 1.6-μm data are especially useful for detecting illuminated water clouds in polar regions because 1) at 1.6 μm, the reflectance of water droplets is significantly higher than that of snow or ice and 2) 1.6-μm data are unaffected by atmospheric inversions. Validation data confirmed the accuracy of their classification technique.

Jung-Hsien Chiang et al. (2008) presented a combination of rough-based feature selection and RBF neural network to classify gene expression data. Their rough-based feature selection can find the relevant features without requiring the number of clusters to be known a priori and identify the centers that approximate to the correct ones. In their technique, they attempted to introduce a prediction scheme that combines the rough-based feature selection method with radial basis function neural network. For further consider the effect of different feature selection methods and classifiers on this prediction process, they used the Naive Bayes and linear support vector machine as classifiers, and compared the performance with other feature selection methods, including information gain and principle component analysis. They demonstrated the performance by several published datasets and their results showed that their proposed method can achieve high classification accuracy rate.

Chakraborty et al. (2003) presented a hybrid artificial neural network (ANN) architecture, comprised of the self-organizing feature map and learning vector quantization (LVQ) to classify sea-floor roughness. The presently studied numerical model, e.g., composite roughness, is successful for the purpose of seafloor classification employing processed multibeam angular backscatter data from manganese-nodule-bearing locations of the Central Indian Ocean Basin. They implemented their technique as an alternative technique for sea-floor roughness classification, giving comparative results with the aforesaid numerical model for processed multibeam angular backscatter data. However, the composite-roughness model approach is protracted due to the inherent need for processed data including system-gain corrections. In order to establish that tedious processing of raw backscatter values is unessential for efficient classification, hybrid ANN architecture has been attempted here due to its nonparametric approach. In this technical communication, successful employment of LVQ algorithm for unprocessed (raw) multibeam backscatter data indicates true real-time classification application.

Wutao Chen et al. (2009) proposed a technique to classify gene expression data using artificial neural network ensembles based on samples filtering. Bioinformatics analysis based on microarray technology is
facing serious challenges, due to the extremely high dimensionality of the gene expression data comparing to the typical small number of available samples. Single artificial neural network was unstable and inaccurate for classification. In their technique, they introduced classifying gene expression data using artificial neural network ensembles based on samples filtering. Simulation tests were carried out to verify their proposed strategy using Leukemia data sets, and the test results were compared with those of single artificial neural network, bagging artificial neural network ensembles and support vector machine. Their results indicated that their method is more stable and more accurate.

Wani (2014) proposed a technique to classify microarray data using subspace grids with synergistic and distributed neural network models. Synergistic and distributed neural network models are employed in their work for Microarray data classification. Their proposed approach used subspace grids as input to synergistic and distributed neural network models. Their technique first describes projection of multidimensional Microarray data to a number of lower dimensional subspaces. Their technique used two algorithms to define lower dimensional subspaces. The range of value associated with each vector of a subspace is divided into a number of equal parts to define subspace grids. The resulting subspace grid data is used with their proposed synergistic and distributed neural network models to classify patterns associated with multidimensional Microarray data. Their results showed that the use of subspaces grids with synergistic and distributed neural network models produced good results to classify patterns in multidimensional Microarray data.

Tian et al. (2010) proposed a technique to classify micro-array gene expression data using neural networks. Classification of yeast genes based on their expression levels obtained from micro array hybridization experiments is an important and challenging application domain in data mining and knowledge discovery. Over the past decade, neural networks and support vector machines (SVMs) have achieved good results for genes classification. In their technique, they used two neural networks to classify unseen genes based on their expression levels. In order to remove some of the noise and deal with the imbalanced class distribution of the dataset, data pre-processing is firstly performed before data classification in which data cleaning, data transformation and data over-sampling using SMOTE algorithm are undertaken. Thereafter, two neural networks with different architectures are trained using Scaled Conjugate Gradient in two different ways: 1) the training-validation-testing approach and 2) 10-fold cross-validation. Experimental results showed that their methodology outperformed the previous best-performing SVM for this problem and 8 other classifiers: 3 SVMs, C4.5, Bayesian network, Naive Bayes, K-NN and JRip.

**Research Directions:**

The area to be concentrated is to consider the secondary information of ABC, and more mathematical validation of ABC is required. Application of ABC in image segmentation has to be increased. Training an artificial neural network is an optimization task since it is desired to find optimal weight set of a neural network in training process. Training neural networks by ABC algorithm can be increased, so that classification tasks would be an easier job.

Multi-objective optimization has been a difficult problem and focus for research in fields of science and engineering. ABC could be applied to the multi objective problems, since most of the real world problems are multi objective in nature.

**Conclusion:**

This paper deals with the Artificial Bee Colony and the various models of the bee colony used in the literature. This paper shows the review based on ABC models for different optimizations, the application of ABC models in image segmentation and shows the application of ABC models in classification and finally data classification using neural network. This survey concludes that implementation of ABC models in clustering process for image segmentation and ABC models trained Neural Network for classification can be concentrated in future research work.

NOTE: The high-lightened texts are included according to the reviewer’s suggestions.

**REFERENCES**


