International Journal of Parallel Programming (2020) 48:310–328
https://doi.org/10.1007/s10766-018-0594-6

A Novel Artificial Bee Colony Optimization Algorithm with SVM for Bio-inspired Software-Defined Networking

Hsiu-Sen Chiang[1] · Arun Kumar Sangaiah[2] · Mu-Yen Chen[1] · Jia-Yu Liu[1]

Received: 19 March 2018 / Accepted: 10 August 2018 / Published online: 16 August 2018

Abstract
In recent years, artificial intelligence and bio-inspired computing methodologies have risen rapidly and have been successfully applied to many fields. Bio-inspired network systems are a field of biology and computer science, it has the high relation to the bio-inspired computing and bio-inspired system. It has the self-organizing and self-healing characteristics that help them in achieving complex tasks with much ease in the network environment. Software-defined networking provides a breakthrough in network transformation. However, increasing network requirement and focus on the controller for determining the network functionality and resources allocations aims at self-management capabilities. More recently, the artificial bee colony (ABC) algorithm has been used to solve the issues of parameter optimization. In this paper, a discretized food source for an artificial bee colony (DfABC) optimization algorithm is proposed and applied to optimize the kernel parameters of a support vector machine (SVM) model, creating a new hybrid. In order to further improve prediction accuracy, the proposed DfABC algorithm is applied to six popular UCI datasets. We also compare the DfABC algorithm to particle swarm optimization (PSO), the genetic algorithm (GA), and the original ABC algorithm. The experimental results show that the proposed DfABC-SVM model achieves better classification accuracy with a shorter convergence time, outperforming the other hybrid artificial intelligence models.

Keywords Parametric optimization · Bio-Inspired computing · Support vector machine · Artificial bee colony algorithm · Software-defined networking

1 Introduction

In recent years, predictive models based on artificial intelligence methods have rapidly risen in popularity and have been successfully applied to many fields. Although arti-

✉ Mu-Yen Chen
mychen.academy@gmail.com

Extended author information available on the last page of the article

ficial intelligence methods can be widely applied to the fields of decision support and data mining, a few issues in the model development process remain. Such issues include parameter settings which are not objective, parameters which cannot satisfy different types of datasets, and the inability of the learning and prediction rates of the model to converge. Therefore, how to develop an artificial intelligence model that converges rapidly and is applicable to different types of datasets has always been an important issue.

Parameter settings are critical to model development. In the past, the design of experiment (DOE) approach has been proposed to solve related problems. The goal of DOE is to find the important variables through experimentation, enhance performance accuracy, and further decrease both the frequency and duration of the experiments [9]. Scholars conducted DOE studies covering everything from the traditional Trial and Error and Rule of Thumb to the later Taguchi method [29, 30]. The concept of Trial and Error is similar to that of the Exhaustive-Attack Method: determining a better parameter combination over time. Although the Trial and Error method incurs a greater cost, the method finds the best parameter combination relatively easily. The Trial and Error method is most often used when the number of parameters to be optimized is not large. Under such conditions, the use of the Trial and Error method to adjust slowly can sometimes achieve unexpectedly good results [2, 16]. With the Rule of Thumb, parameter settings are based on personal experience or a setting mode recommended by past scholars in order to establish an effective and stable model. The experimental model can thus be constructed with a lower time cost, but the model may not necessarily be stable. The main reason for this is that different types of datasets have their own unique features. If we use the same method to set the model parameters for different types of datasets, there will be a big gap between the experimental result and the stability of the model. Therefore, the Rule of Thumb method is more suitable for a mature model [6, 7]. The Taguchi method is an experimental design in which each factor is set to determine a relatively good result combination by exploring the interaction between the factors [29, 30]. Although the Taguchi method can effectively get relatively good results with fewer experiments, the both number of experimental factors used and how the experimental level is set will influence model stability. The Taguchi method is applicable in well-developed artificial intelligence methods, and many scholars use it to assist in the adjustment of neural network parameters [10, 25, 26, 35].

Bonabeau et al. [5] defined swarm intelligence (also called bio-inspired algorithms) as "…algorithms or distributed problem-solving devices inspired by the collective behavior of social insect colonies and other animal societies…" Developed by Eberhart and Kennedy in 1995 [13], the particle swarm optimization (PSO) algorithm is a population-based stochastic optimization technique inspired by the social behavior of flocking birds or schooling fish. In PSO, potential solutions are called particles. The algorithm searches for the optimal value by sharing historical information and social information between individual particles [8]. Moreover, the cooperation and competition among the population often allows the PSO algorithm to efficiently and effectively obtain solutions. However, the main drawback of the PSO algorithm is how easily it is impacted by partial optimism, which decreases accuracy when speed and direction are regulated [22]. After the introduction of PSO, some scholars proposed to model the intelligent behaviors of honeybee swarms in an effort to solve heterogeneous dataset problems [36–34, 37]. Karaboga [18] presented the artificial bee colony (ABC) algorithm and the empirical results showed that ABC can overcome the common issue of falling into local optimization. However, the results also showed that the ABC cannot quickly select the global optimum and requires an increase in the number of iterations during the probability choice stage [20, 21], significantly lengthening the convergence time at the cost of efficiency.

In recent years, novel bio-inspired algorithms such as PSO and ABC have been used to solve parameter optimization issues. To address the above problem, the novel approach proposed by this study is a discretized food source for an artificial bee colony (DfABC) optimization algorithm, which, along with parameter adjustments in the core program of an optimal support vector machine (SVM), which proposed by Vapnik and Cortes [36], allows us to construct a quickly converging classification model that has excellent prediction performance and can handle different types of datasets. In addition, this study uses a number of popular benchmark datasets from UCI Machine Learning Repository with different dimensions, sizes, and dimensional classifications in order to validate the proposed DfABC-SVM model. The experimental results show that the proposed DfABC algorithm can be applied to the SVM model and used to handle different types of datasets, and that the classification accuracy and convergence rate are better than other artificial intelligence or bio-inspired methods.

The rest of this paper is organized as follows. In Sect. 2, the proposed model is briefly presented with a step by step explanation. Section 3 analyzes the experimental results involving several UCI datasets. Section 4 discusses the results about the proposed methods. Finally, our findings and future recommendations are discussed in Sect. 5.

2 Methodology

2.1 Bio-inspired Software-Defined Networking

Software Defined Networks (SDN) raise a variety of opportunities for network evolution. The key feature of SDNs is the decoupling of data and control planes, removing the control plane from the network hardware. This provides remarkable resilience in programming, along with a broad range of opportunities to optimize network resource utilization. In fact, SDNs can address numerous IoT (Internet of Things, IoT) challenges, from serving massive numbers of service requests and responses; processing the enormous data flow of IoT sensors, devices, and appliances; accommodating unique identification schemes, allowing for autonomous network and service management, optimizing resources, and enabling service virtualization to seamless communication.

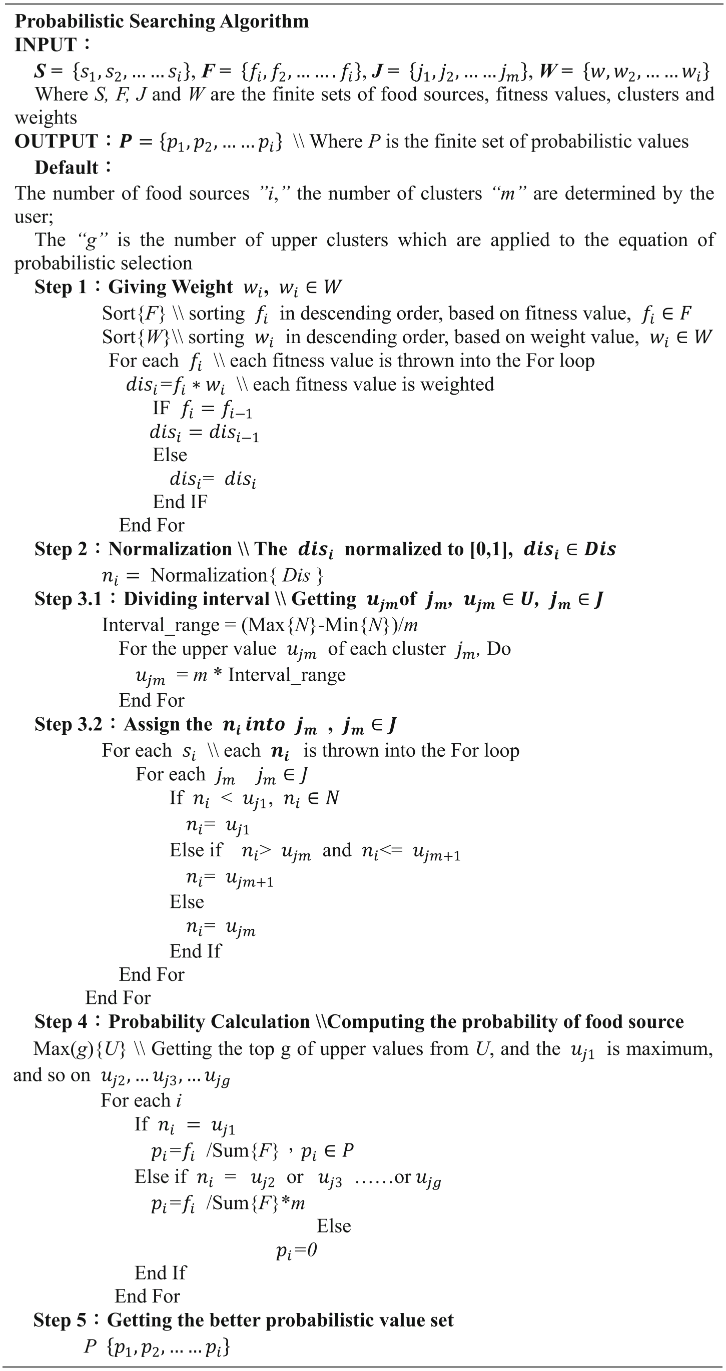
Bio-inspired networking techniques have been proposed and applied into the practices for a long time. The previous research contributions promoted the networking developments and related applications growth especially in large scale networks. Dressler and Akan [12] figured out the three different bio-inspired researches, including bio-inspired computing, bio-inspired systems, and bio-inspired networking. The bio-inspired computing is focus on the efficient algorithms deployment and reach the optimization network processes and network-level traffic pattern recognition. The bio-inspired systems is focus on the technical developments for the ideal structure and function of the biological systems. Finally, the bio-inspired networking is focus on the strategies developments to guarantee the efficient and scalable networks under the complex and uncertain conditions. However, there are many issues need to deal with and then can achieve the next generation network advantages [12]. In such a network, link or node failures in the SDN-enabled physical infrastructure may disrupt not only its own connectivity and capacity, but may also disrupt user requests for cloud applications and services in network abstraction. Considering their dynamic nature and lack of infrastructure, to ensure continuous service, such networks must be capable of self-organization [11], self-evolution [24], and survivability [38].

2.2 The Proposed DfABC Algorithm

In this paper, a discretized food source for an artificial bee colony (DfABC) algorithm is proposed and improve the traditional ABC bottleneck of food source fitness values (Fig. 1). The notation of the proposed DfABC is listed in Table 1.

2.3 The Illustration of the DfABC-SVM Model

The main purpose of this research is to optimize the core parameters of the SVM model using the DfABC algorithm. By optimizing the two parameters r and γ in the SVM model, a better cutting line is drawn to classify different types of datasets and minimize the prediction error rate. The performance of SVM mainly depends on model selection, including the selection of the kernel function type and kernel parameter [1]. In the research, polynomial kernel function was selected and implemented in the SVM model. The first consideration of choosing polynomial kernel function is the because it has fewer parameters to select and has significant performance to other kernel functions for the hybrid SVM model with stochastic optimization technique [15]. Second, Fujibuchi and Kato [14] found the different kernel function - linear, polynomial, Gaussian radial basis function kernel function (RBF) showed the similar accuracies for the heterogeneous microarray data when no noise was added. However, the standard linear and RBF kernels decreased the accuracies with increasing noise levels in the training datasets. Third, Sebtosheikh and Salehi [27] claimed that the prediction by SVMs using small training datasets, they recommended to use the polynomial kernel function. The empirically results showed the polynomial kernel function outperformed than the RBF kernel function. Fourth, Alibrandi et al. [3] used the second-order polynomial SVM model and found it can obtain a good approximation of the limit state without high computational complexity. Finally, Liu et al. [23] used SVM to learn the dynamic kinematics relationships between the legs and the trunk of the biped robots. The empirically results demonstrated the polynomial kernel superiority of the RBF kernel. They claimed that polynomial kernel is typical representative of global kernels and RBF kernel is the typical representative of local kernels. Therefore, we decided to choose the polynomial kernel function based on the above reasons and it will be suitable for the heterogeneous data classification. The



```
Probabilistic Searching Algorithm
INPUT：
  $S = \{s_1, s_2, \ldots\ldots s_i\}$, $F = \{f_i, f_2, \ldots\ldots.f_i\}$, $J = \{j_1, j_2, \ldots\ldots j_m\}$, $W = \{w, w_2, \ldots\ldots w_i\}$
  Where S, F, J and W are the finite sets of food sources, fitness values, clusters and weights
OUTPUT：$P = \{p_1, p_2, \ldots\ldots p_i\}$ \\ Where P is the finite set of probabilistic values
  Default：
The number of food sources "i," the number of clusters "m" are determined by the user;
  The "g" is the number of upper clusters which are applied to the equation of probabilistic selection
  Step 1：Giving Weight $w_i$, $w_i \in W$
        Sort{F} \\ sorting $f_i$ in descending order, based on fitness value, $f_i \in F$
        Sort{W}\\ sorting $w_i$ in descending order, based on weight value, $w_i \in W$
        For each $f_i$ \\ each fitness value is thrown into the For loop
          $dis_i = f_i * w_i$ \\ each fitness value is weighted
            IF $f_i = f_{i-1}$
            $dis_i = dis_{i-1}$
            Else
              $dis_i = dis_i$
            End IF
          End For
  Step 2：Normalization \\ The $dis_i$ normalized to [0,1], $dis_i \in Dis$
        $n_i =$ Normalization{ Dis }
  Step 3.1：Dividing interval \\ Getting $u_{jm}$of $j_m$, $u_{jm} \in U$, $j_m \in J$
        Interval_range = (Max{N}-Min{N})/m
          For the upper value $u_{jm}$ of each cluster $j_m$, Do
            $u_{jm} = m *$ Interval_range
          End For
  Step 3.2：Assign the $n_i$ into $j_m$ , $j_m \in J$
        For each $s_i$ \\ each $n_i$ is thrown into the For loop
          For each $j_m$   $j_m \in J$
            If $n_i < u_{j1}$, $n_i \in N$
              $n_i = u_{j1}$
            Else if  $n_i > u_{jm}$ and $n_i <= u_{jm+1}$
              $n_i = u_{jm+1}$
            Else
              $n_i = u_{jm}$
            End If
          End For
      End For
  Step 4：Probability Calculation \\Computing the probability of food source
  Max(g){U} \\ Getting the top g of upper values from U, and the $u_{j1}$ is maximum, and so on $u_{j2}, \ldots u_{j3}, \ldots u_{jg}$
        For each i
          If $n_i = u_{j1}$
            $p_i = f_i$ /Sum{F} ， $p_i \in P$
          Else if $n_i = u_{j2}$ or $u_{j3}$ ……or $u_{jg}$
            $p_i = f_i$ /Sum{F}*m
                                Else
                        $p_i = 0$
          End If
        End For
  Step 5：Getting the better probabilistic value set
      $P$ $\{p_1, p_2, \ldots\ldots p_i\}$
```

Fig. 1 The pseudo code for the DfABC algorithm

Table 1 The notation of DfABC

Parameter name	Parameter description
S	The finite set of food source, $S=\{s_1, s_2, \ldots\ldots s_i\}$
F	The finite set of fitness values, $F=\{f_i, f_2, \ldots\ldots. f_i\}$
J	The finite set of clusters, $J=\{j_1, j_2, \ldots\ldots j_m\}$
N	The finite set of the normalized fitness values, $N=\{n_1, n_2, \ldots\ldots n_i\}$
Dis	The finite set of the weighted fitness values, $Dis=\{dis_1, dis_2, \ldots\ldots dis_i\}$
W	The finite set of weights, $W=\{w, w_2, \ldots\ldots w_i\}$
P	The finite set of probabilistic values, $P=\{p_1, p_2, \ldots\ldots p_i\}$
m	The number of clusters, determined by domain experts
U	The finite set of each cluster's upper value, $U=\left\{u_{j1}, u_{j2}, \ldots\ldots u_{jm}\right\}$
Interval_range	The range of each interval, determined by domain expert
i	The total number of food sources
g	Getting the top g of upper values from U
Max{ }	The maximum function
Min{ }	The minimum function
Sort{ }	The sorting function
Sum{ }	The summing function
Normalization{ }	The normalization function

kernel function of the SVM model adopted by this research is the polynomial equation shown below:

$$K_{(x_i,x_j)} = \left(\gamma X_i^T X_j + r\right)^d \tag{1}$$

where d is the dimensions of one problem to be solved, T is the transposition of a matrix, and γ and r are parameters. DfABC is used to optimize parameters γ and r so that each food source has both parameter values γ and r. Through the steps of discretization, grouping, and probabilistic choice, a better food source is found. Next, after repeated cycles, the optimal solution for γ and r are found, and this solution is applied to the SVM model to obtain the best classification result. The entire parameter optimization process is described in the following steps.

Step 1. Randomly Generate 10 Food Sources and Food Source Fitness Values in the First Iteration: e.g. fitness value r, γ.

$$\text{Food source sequence}[i] = 0.65_{0.1,0.2}, 0.7_{0.3,0.8}, 0.2_{0.2,0.4}, 0.65_{0.1,0.3}, 0.3_{0.1,0.4}, 0.3_{0.1,0.5}, 0.25_{0.1,0.9}, 0.4_{0.25,0.3}, 0.35_{0.1,0.6}, 0.75_{0.3,0.76}$$

With food source initialization, the Employ bee generates one new food source with two parameter solutions, r and γ. After comparing the new food source with the current food source, the Employ bee will choose the one with a greater fitness value to be the guarded food source. Below is an example of the first food source.

$$v_{ij} = x_{ij} + \varphi_{ij}\left(x_{ij} - x_{kj}\right) = 0.65_{0.1,0.2} + 0.5\left(0.65_{0.1,0.2} - 0.61_{0.15,0.23}\right)$$

$$v_{ij} = 0.65_{0.1,0.2} + 0.5\left(0.04_{-0.05,-0.03}\right) = 0.67_{0.05,0.17}$$

Once v_{ij} is obtained, the Employed bee executes the greedy selection and chooses v_{ij}.

Step 2. Onlooker Bee Executes the Probabilistic Choice

The food source fitness values have been successfully divided into *m* groups after performing the DfABC algorithm. In this step, the *m*–groups of food sources are applied to different probabilistic choice methods as a reference for the Onlooker bee. The threshold value of this study is 0.1, i.e., only the food source with a probabilistic choice greater than 0.1 can be selected by the Onlooker bee.

Step 3. Onlooker Bee Executes a Comparison of Fitness Values

The Onlooker bee finds one food source near to the selected food source and compares the fitness values of these two food sources. The Onlooker be chooses the food source with the greater fitness value. This step is the same as Step 1, and we assume that the Onlooker bee selects the second food source in the descending sort sequence.

Food source in the descending sort sequence [*i*]=

$$0.75_{0.3,0.76}, 0.7_{0.3,0.8}, 0.65_{0.1,0.2}, 0.65_{0.1,0.3}, 0.4_{0.25,0.3},$$

$$0.35_{0.1,0.6}, 0.3_{0.1,0.4}, 0.3_{0.1,0.5}, 0.25_{0.1,0.9}, 0.2_{0.2,0.4}$$

$$v_{ij} = x_{ij} + \varphi_{ij}\left(x_{ij} - x_{kj}\right)$$

$$= 0.7_{0.3,0.8} + 0.7\left(0.73_{0.3,0.8} - 0.73_{0.25,0.7}\right)$$

$$v_{ij} = 0.7_{0.3,0.8} + 0.7\left(-0.3_{0.05,0.1}\right) = 0.679_{0.35,0.9}$$

Once v_{ij} is obtained, the Onlooker bee executes the greedy selection and chooses x_{ij}.

Step 4. Scout Bee Looks for a New Food Source

This phase may or may not be executed. The triggering condition is that the fitness value of one food source has not changed after the number of iterations exceeds its restriction. When that happens, the Employed bee will discard the food source and generate a new food source. Assuming the fitness value of the 5th food source after descending sorting has not changed after the restricted limit of iterations, a new food source is generated.

$$\text{Set parameter } j = [-5, 5] \quad X_i^j = X_{min}^j + rand(0, 1)\left(X_{max}^j - X_{min}^j\right)$$
$$= -5 + 0.7(10) = 2$$

Table 2 UCI dataset descriptions

Dataset name	Instances	Dimensions	Categories
Iris	150	5	3
Glass	214	10	7
Spam based	4601	57	2
MAGIC Gamma	19,020	11	2
German	1000	20	2
Australian	690	14	2

By repeatedly performing the above steps, the best parameter values of γ and r in the SVM model can be obtained.

3 Experiment and Result Methodology

In order to evaluate the effectiveness of the proposed model, this study uses several datasets to determine if the DfABC-SVM model can be applied to different dataset types. In addition, we also compare the DfABC-SVM model with decision tree algorithms such as the classification and regression tree (CART) and the conventional SVM model, and with other hybrid SVM models which use EA or bio-inspired algorithms, such as GA, PSO, and ABC. The results show that the model developed by this study has superior convergence speed and a high level of classification accuracy, and does not fall to local optimization.

3.1 Dataset Descriptions

This study adopts six datasets of different types to evaluate the accuracy and convergence speed of DfABC-SVM. To evaluate if DfABC-SVM has consistent effect in heterogeneous datasets with different features, we chose datasets which vary in the number of instances, dimensions, and subject column categories. The six datasets (i.e., Iris, Glass, Spam Based, MAGIC Gamma, German, and Australian) were collected from the UCI Machine Learning Repository. Table 2 describes these datasets.

The Iris dataset is one of the most popular datasets in the data mining field and is frequently used to prove model stability. This dataset has 5 dimensions and only 150 instances to classify three different kinds of irises. What distinguishes this dataset from the others is that it has few dimensions, a small number of instances, and only three categories of subject columns. The four major dimensions (sepal length, sepal width, petal length, and petal width) are the main base by which to determine to which of the three iris breeds any data belongs (i.e., Iris Setosa, Iris Versicolour, and Iris Virginica).

The creation of the Glass dataset was motivated by criminological investigations. The glass left behind at the scene of a crime can be used as evidence, if it is correctly identified. This dataset uses 10 dimensions and only 214 instances to classify seven different kinds of glass. The number of dimensions of the Glass dataset is not high. The instances of data are also not too many. However, the subject columns for classification are divided into seven glass categories. This provides a challenge for a classification model because the classification error rate is higher when more types of classifications are available. Therefore, among the seven datasets, the Glass dataset is the most difficult to process.

The Spam Based dataset has the highest number of dimensions and a relatively large number of dataset instances. The dimensions record information regarding e-mail content, such as how many commas and spaces are in the message. Most of the dimensions indicate whether a particular word or character occurs frequently in the e-mail. This dataset has 57 dimensions and 4601 instances to classify only two different kinds of e-mail, i.e., to determine whether a given message is or is not junk mail.

The MAGIC Gamma Telescope observes the cosmic rays striking Earth's atmosphere. The resulting dataset includes both gamma rays and hadronic showers. This dataset has 11 dimensions and 19,020 instances to classify future readings from this telescope as either Signal (gamma) or Background (hadron). This dataset has the largest number of instances, so it is used to evaluate whether the model is capable of handling a large amount of data.

The German dataset consists of client data from a single German bank. The dataset contains 20 dimensions and 1000 instances. The German dataset is considered to be of moderate complexity and has a large number of instances. The main purpose of this dataset is to classify credit card applicants into one of two types to determine whether or not to issue a credit card to the client. Nineteen dimensions of client information (salary, whether or not the client has children, etc.) allow the model to determine if the client's credit status is good. The subject column for classification discriminates the records into only two classes: good and bad.

The Australian dataset also classifies clients into two categories: those with good credit and those with poor credit. The complexity and number of instances are low and the entire dataset is continuous. This dataset uses 14 dimensions and 690 instances to classify credit as good or bad. The subject column for classification is divided into two types, so the dataset can be used to analyze and explore the classification model, and makes it easier to obtain better predictive results.

3.2 Experiment Environment and Parameter Setting

In order to make a fair comparison, the values of colony size and maximum cycle number of the ABC algorithm are chosen same as or less than the values of swarm size and maximum iteration number used in PSO case, respectively. The ABC algorithm has a few control parameters: the maximum number of cycles (MCN) equals the maximum number of iterations and the colony size equals the population size (i.e., 125), as in Srinivasan and Seow [28], Karaboga and Basturk [20]. Onlooker bees and employed bees each made up 50% of the colony, while the number of scout bees was set to one. The increased number of scouts encourages exploration as increasing the number of onlookers on a food source increases the degree of exploitation. In the SVM model, the searching range of C and γ were set between 1 and 100 [17]. Then, the PSO-SVM model, the cognition learning factor c_1 was set to 2, and social learning factor c_2 was set to 1. The number of particles was set to 125, and maximum number of iterations was set to 1000. The inertia weight and maximum velocity value were set

Table 3 Classification accuracy of different datasets (unit:%)

Dataset	CART	SVM	GA-SVM	PSO-SVM	ABC-SVM	DfABC-SVM
Iris	86	95	97.7	97.7	100	100
Glass	45.3	48.1	54.9	50.8	59.0	65.5
Spam base	60.6	83.5	84.2	84.5	84.9	85.7
MAGIC Gamma	64.8	73.4	83.5	80.6	85.1	88.1
German	55.4	60.3	59.5	61.2	60.9	62.5
Australia	82.2	85	85	83.8	85	86.2
Average	65.71	74.21	77.46	76.43	79.15	81.33

to 0.9 and 2. In the GA setting, the single point crossover rate and mutation rate were set to 0.8 and 0.01 [4, 19].

3.3 Evaluating Classification Accuracy

As noted, above, 6 distinctly different datasets are used by this study to evaluate the applicability, accuracy and stability of the DfABC-SVM model. The model is also compared with other classification methods.

The ratio of training to testing datasets was 7:3. In order to determine the best clusters (m) of cutting intervals in DfABC-SVM, the experimental design set the clusters from 2 to 10 and assessed the classification accuracy of the six datasets. The results are shown in Table 3.

Table 3 shows that the SVM model is better than CART, and the SVM model which was optimized with the bio-inspired algorithm is better than conventional SVM. This means that the bio-inspired algorithm has a certain degree of effect on finding the relative optimal solution for parameter settings. Moreover, the method proposed by this study proved to be the most stable, had the best predictive results, and performed better than the others when handling different types of datasets.

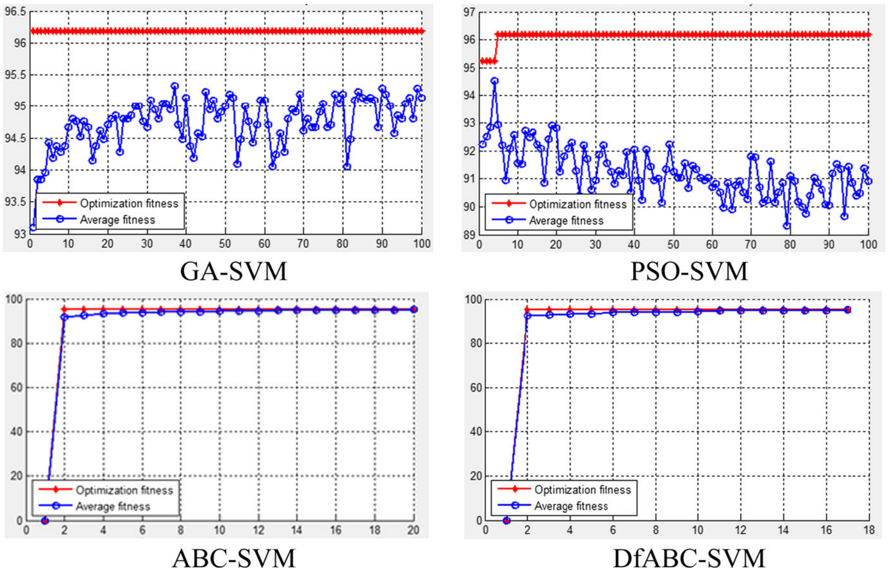
Table 4 shows that applying the concept of discrete food sources to m clusters is feasible. When the subject column for classification has less than three categories, setting the number of clusters m to a value greater than 8 produces a less effective result. However, if the subject column for classification has more than three categories, setting the clusters m to a value greater than 8 has a better result. Under general conditions, setting the number of clusters m to between 4 and 7 provides a preferable, more stable experimental result which is not impacted by the features of the dataset.

3.4 Evaluating Convergence Effectiveness

First, the global optimum was determined and used as the standard line. The number of iterations, convergence speed, and global optimum of DfABC-SVM was then compared with those of other SVM models which had been combined with a bio-inspired algorithm. In Figs. 2, 3, 4, 5, 6 and 7, the Y-axis represents the classification accuracy as a percentage, and the X-axis represents the number of iterations. Optimization

Table 4 Classification accuracy of DfABC-SVM in different numbers of clusters

Clusters number (m)	Iris	Glass	Spam based	MAGIC Gamma	German	Australian
2	97.7	60.6	85.1	87.1	61.2	85.02
3	100	62.2	84.2	87.1	60.9	85.02
4	100	60.6	84.03	87.5	61.2	84.5
5	100	62.2	85.7	88.1	61.5	86.2
6	100	55.7	84.0	87.5	60.9	84.5
7	97.7	65.5	84.4	88.1	61.2	85.02
8	100	62.2	84.0	86.9	62.5	66.3
9	100	65.5	83.3	86.3	61.5	66.3
10	100	62.2	84.3	86.3	62.5	68.5



GA-SVM

PSO-SVM

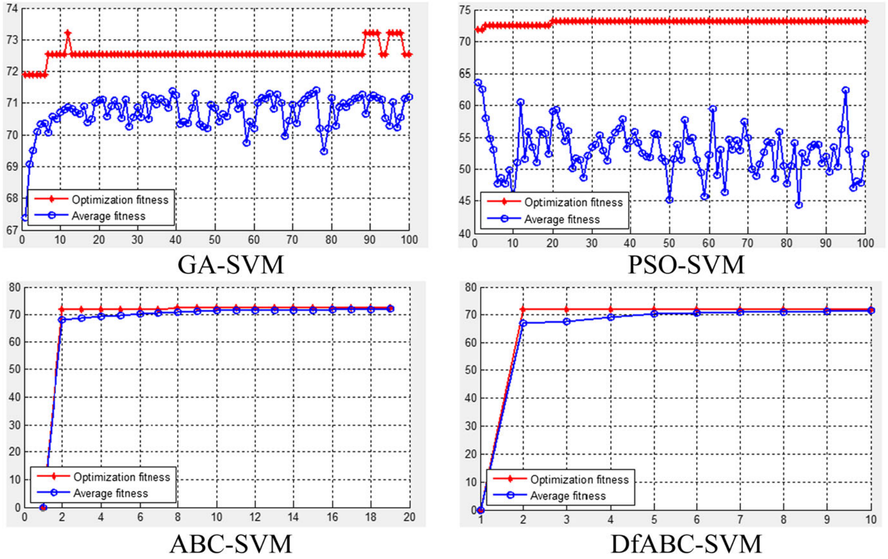
ABC-SVM

DfABC-SVM

Fig. 2 Convergence result of various bio-inspired SVM models on the iris dataset

Fitness is the average accuracy after five rounds of cross-validation substituting the global optimal parameter setting generated by the Exhaustive Attack Method in the model. Average Fitness is the average accuracy after five rounds of cross-validation substituting the relative better parameter setting by various bio-inspired algorithms in the SVM model. When the two lines are very close or matching, it means that the classification model is completely trained and has achieved convergence.

Figures 2, 3, 4, 5, 6 and 7 show that the convergence statuses of various bio-inspired SVM models are quite different. It is obvious that the convergence of the GA- and PSO-SVM models is not complete, and, in some datasets, those models failed to achieve convergence. According to the convergence figures, both the ABC-SVM and



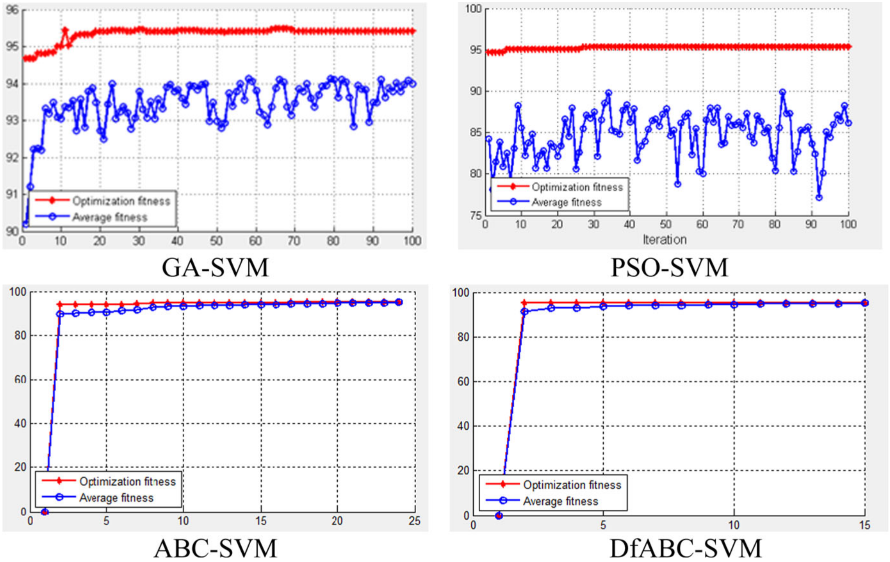
GA-SVM

PSO-SVM

ABC-SVM

DfABC-SVM

Fig. 3 Convergence result of various bio-inspired SVM models on the glass dataset



GA-SVM

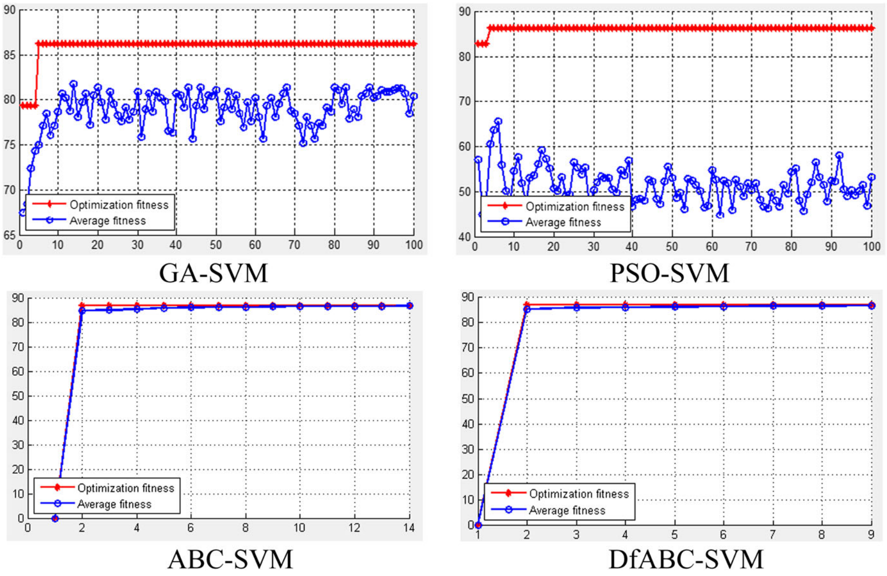
PSO-SVM

ABC-SVM

DfABC-SVM

Fig. 4 Convergence result of various bio-inspired SVM models on the spam based dataset

DfABC-SVM models reached a stable state within a very few iterations. This proves that the ABC parameter optimization model is indeed more capable than the other bio-inspired algorithms. In addition, the convergence speed of the DfABC model is faster than that of the ABC model.



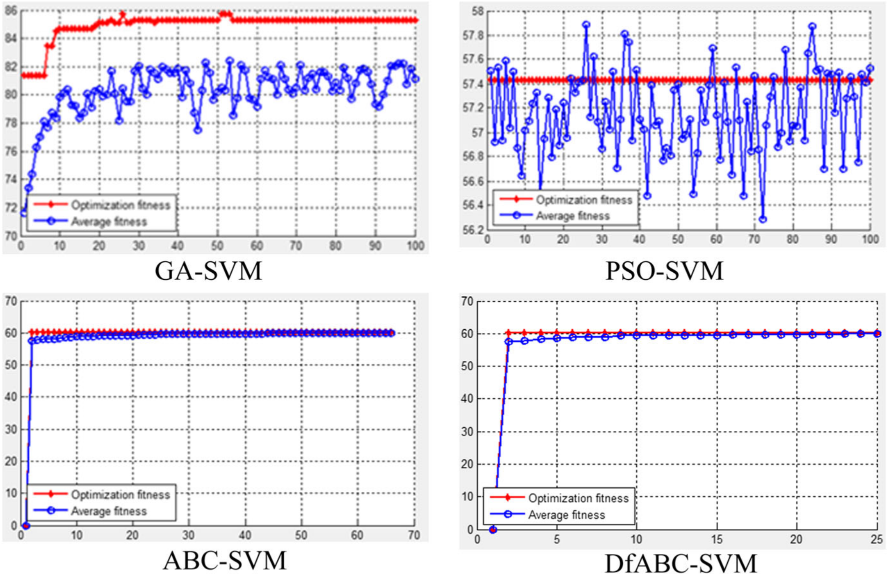
GA-SVM

PSO-SVM

ABC-SVM

DfABC-SVM

Fig. 5 Convergence result of various bio-inspired SVM models on the MAGIC Gamma dataset



GA-SVM

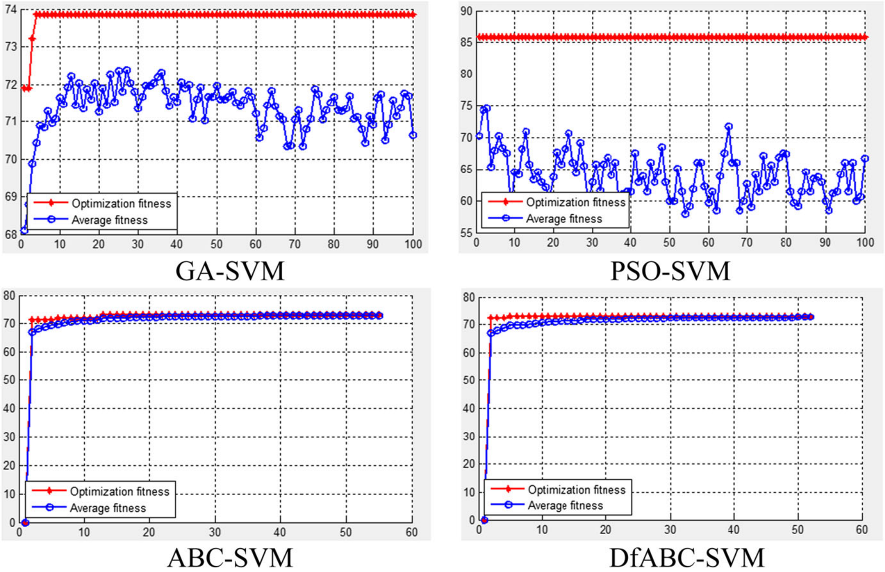
PSO-SVM

ABC-SVM

DfABC-SVM

Fig. 6 Convergence result of various bio-inspired SVM models on the German dataset

Table 5 shows that the ABC and DfABC optimized models reach convergence more quickly. The definition of "reach convergence" means that the gap between Optimization Fitness and Average Fitness is no greater than 0.3 after three continuous iterations. Although GA and PSO spent less time processing each iteration, they did not reach



GA-SVM

PSO-SVM

ABC-SVM

DfABC-SVM

Fig. 7 Convergence result of various bio-inspired SVM models on the Australian dataset

convergence after 100 iterations. Besides, such models are very unstable. Therefore, although the time cost for the proposed model is higher, the accumulated convergence time is, in fact, the shortest among the four models. Its classification effectiveness is better than that of the other models. Thus, we conclude that the proposed model is a stable classification model which reaches convergence quickly.

4 Discussion

This study modifies the ABC algorithm based on a bio-inspired concept to assist in the parameter optimization of an SVM model. According to the experimental results, the classification effectiveness of DfABC-SVM model proposed by this study is indeed better than that of the other models. The convergence speed is fast, and the model is capable of handling different types of datasets. This study also explored six datasets in three aspects to understand the degree to which the DfABC algorithm could be applied to different types of datasets, and how the number of clusters *m* should be set.

1. Datasets with few instances, few dimensions, and fewer classification categories

In datasets with a relatively small number of instances, few dimensions, and with classification dimensions fewer than four categories (e.g., the Iris and Australian datasets), setting the number of clusters *m* between 3 and 7 can obtain good results, as shown in Table 4. This is showed the DfABC algorithm, in datasets with few instances, few dimensions, and fewer classification categories, the discretization of the food source can improve the fitness value of each food source. Dividing the food

Table 5 The time cost of each method (Unit: second)

Dataset	GA-SVM	PSO-SVM	ABC-SVM	DfABC-SVM
Iris	3.097,300 (not converged)	4.707064 (not converged)	2.277551 (the 20th iteration converged)	2.093585 (the 17th iteration converged)
Glass	16.891995 (not converged)	16.191765 (not converged)	6.996393 (the 19th iteration converged)	4.016722 (the 10th iteration converged)
Span base	2751.207885 (not converged)	3784.791565 (not converged)	2087.174705 (the 24th iteration converged)	1193.579092 (the 15th iteration converged)
MAGIC gamma	5872.547110 (not converged)	7305.404091 (not converged)	2686.644140 (the 14th iteration converged)	1713.750846 (the 9th iteration converged)
German	438.25 (not converged)	507.83 (not converged)	604.35 (the 66th iteration converged)	241.14 (the 25th iteration converged)
Australia	66.02 (not converged)	90.38 (not converged)	18.88 (the 55th iteration converged)	18.39 (the 53rd iteration converged)

sources into different groups results in faster convergence and increase in classification effectiveness.

2. Datasets with large instances, many dimensions, and fewer classification categories

In datasets with a large number of instances, many dimensions, and with fewer than four categories for the classification dimensions (e.g., the Spam Based and MAGIC Gamma datasets), the classification effectiveness is higher in comparison to other models. In addition, the classification effectiveness does not differ when the number of clusters *m* is set differently, indicating that the number of clusters of food sources does not affect classification accuracy and the accuracy rate is very stable than other approaches. Therefore, the DfABC algorithm proposed by this study has a certain level of effectiveness in handling this type of dataset but the difference of clusters *m* is not so obvious.

3. Datasets with many classification categories

In datasets with more than four categories in their classification dimensions (e.g., the Glass and German datasets), good predictive results are difficult to obtain (see Table 4) because there are too many categories. However, the results show that the proposed DfABC algorithm can obtain a more stable result than do the other models when the number of clusters is higher. This proves that the proposed algorithm can handle complex datasets with a high number of dimensions.

According to Tables 3 and 4, the classification model of DfABC-SVM achieves better results with all six different types of datasets, when compared to other artificial intelligence models. For the setting of the cluster of food source parameter, the DfABC-SVM model achieves a better result when the number of clusters *m* is set between 5 and 7 (with the exception of the Glass dataset which has more classification dimensions). The convergence speed of the DfABC algorithm is better than that of the GA, PSO, and ABC optimized models. This study proves that the proposed DfABC algorithm can effectively perform parameter optimization and reduce the number of iterations.

5 Conclusion

This study proposed the DfABC algorithm to improve the food source fitness value selection and probabilistic calculation of the original ABC, and applied the proposed DfABC algorithm to an SVM model to improve its parameter adjustment capabilities. After being applied to six different types of datasets, and compared to CART, conventional SVM, and hybrid SVM models with bio-inspired optimized algorithms, the DfABC-SVM model was found to perform better in terms of classification accuracy and convergence speed. Thus, we conclude that DfABC-SVM is a stable model with good classification effectiveness and generalizability.

In addition, the following four findings can be obtained from the experimental results: (1) the modified DfABC algorithm has a faster convergence speed than does the original ABC, and has better classification results. (2) What affects the classification effectiveness most in one particular dataset (the Glass dataset) is not the number of dimensions or instances, but the number of classification dimensions. (3) Of the bio-inspired algorithms, GA and PSO showed poor convergence effectiveness and typically required hundreds of iterations. (4) The experimental results showed suitable values for the cluster of food source parameter for the DfABC-SVM model for different types of datasets. Future studies should investigate whether the DfABC algorithm can be used for all types of optimization problems, i.e., binary, combinatorial, and integer optimization problems. We also suggest that further research be done to determine whether the proposed algorithm can be used in an online data stream or in large scale data environments.

Acknowledgements The authors would like to thank the Ministry of Science and Technology of Taiwan for Grants MOST106-2410-H-025-007, MOST-103-2410-H-025-022-MY2 and National Science Council of Taiwan for Grants NSC-101-2410-H-025-004-MY2 which supported part of this research.

References

1. Aborokbah, M.M., Al-Mutairi, S., Sangaiah, A.K., Samuel, O.W.: Adaptive context aware decision computing paradigm for intensive health care delivery in smart cities-a case analysis. Sustain. Cities Soc. (2017). https://doi.org/10.1016/j.scs.2017.09.004
2. Ahn, H., Kim, K.J.: Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach. Appl. Soft Comput. **9**, 599–607 (2009)
3. Alibrandi, U., Alani, A.M., Ricciardi, G.: A new sampling strategy for SVM-based response surface for structural reliability analysis. Probab. Eng. Mech. **41**, 1–12 (2015)
4. Asogbon, M.G., Samuel, O.W., Omisore, M.O., Awonusi, O.: Enhanced neuro-fuzzy system based on genetic algorithm for medical diagnosis. J. Med. Diagn. Methods (2016). https://doi.org/10.4172/2168-9784.1000205
5. Bonabeau, E., Dorigo, M., Theraulaz, G.: Swarm Intelligence: From Natural to Artificial Intelligence. Oxford University Press, New York, NY (1999)
6. Boritz, J.E., Kennedy, D.B.: Effectiveness of neural network types for prediction of business failure. Expert Syst. Appl. **9**, 503–512 (1995)
7. Carlos, S.C.: Self-organizing neural networks for financial diagnosis. Decis. Support Syst. **17**, 227–238 (1996)
8. Chen, M.Y.: A hybrid ANFIS model for business failure prediction utilizing particle swarm optimization and subtractive clustering. Inf. Sci. **220**, 180–195 (2013)
9. Chen, M.Y., Fan, M.H., Chen, Y.L.: Wei HM (2013) Design of experiments on neural network's parameters optimization for time series forecasting in stock markets. Neural Netw. World **23**(4), 369–393 (2013)
10. Chen, W.C., Hsu, Y.Y., Hsieh, L.F., Tai, P.H.: A systematic optimization approach for assembly sequence planning using Taguchi method, DOE, and BPNN. Expert Syst. Appl. **37**, 716–726 (2010)
11. Dressler, F.: Self-Organization in Sensor and Actor Networks. Wiley, Hoboken (2007)
12. Dressler, F., Akan, O.B.: Bio-inspired networking: from theory to practice. IEEE Commun. Mag. **48**(11), 176–183 (2010)
13. Eberhart, R.C., Kennedy, J.: A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science (Nagoya, Japan), IEEE Service Center, Piscataway, NJ, pp. 39–43 (1995)
14. Fujibuchi, W., Kato, T.: Classification of heterogeneous microarray data by maximum entropy kernel. BMC Bioinform. **8**, 267 (2007). https://doi.org/10.1186/1471-2105-8-267
15. Harish, N., Mandal, S., Rao, S., Patil, S.G.: Particle swarm optimization based support vector machine for damage level prediction of non-reshaped berm breakwater. Appl. Soft Comput. **27**, 313–321 (2015)
16. Hsieh, T.J., Hsiao, H.F., Yeh, W.C.: Mining financial distress trend data using penalty guided support vector machines based on hybrid of particle swarm optimization and artificial bee colony algorithm. Neuro Comput. **82**, 196–206 (2012)
17. Hsu, C.W, Chang, C.C., Lin, C.J.: (2010) A practical guide to support vector classification, Technical Report. Department of Computer Science, National Taiwan University, 2010. http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf/

18. Karaboga, D.: An idea based on honey bee swarm for numerical optimiza-tion, Technical Report-TR06, Erciyes University, Engineering Faculty, Computer Engineering Department (2005)
19. Karaboga, D., Akay, B.: A comparative study of artificial bee colony algorithm. Appl. Math. Comput. **214**, 108–132 (2009)
20. Karaboga, D., Basturk, B.: A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J. Glob. Optim. **39**, 459–471 (2007)
21. Karaboga, D., Basturk, B.: On the performance of artificial bee colony (ABC) algorithm. Appl. Soft Comput. **8**, 687–697 (2008)
22. Kennedy, J., Eberhart, R.: Particle swarm optimization. Proc. IEEE Int. Conf. Neural Netw. **4**, 1942–1948 (1995)
23. Liu, Z., Wang, L., Zhang, Y., Chen, C.L.P.: A SVM controller for the stable walking of biped robots based on small sample sizes. Appl. Soft Comput. **38**, 738–753 (2016)
24. Rathore, H.: Bio-inspired software-defined networking. In: Mapping Biological Systems to Network Systems, pp. 107–115. Springer, Cham (2016)
25. Samuel, O.W., Asogbon, G.M., Sangaiah, A.K., Fang, P., Li, G.: An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction. Expert Syst. Appl. **68**, 163–172 (2017)
26. Sangaiah, A.K., Samuel, O.W., Li, X., Abdel-Basset, M., Wang, H.: Towards an efficient risk assessment in software projects-Fuzzy reinforcement paradigm. Comput. Electr. Eng. (2017). https://doi.org/10.1016/j.compeleceng.2017.07.022
27. Sebtosheikh, M.A., Salehi, A.: Lithology prediction by support vector classifiers using inverted seismic attributes data and petrophysical Logs as a new approach and Investigation of training data set size effect on its performance in a heterogeneous carbonate reservoir. J. Pet. Sci. Eng. **134**, 143–149 (2015)
28. Srinivasan, D, Seow, T.H.: Evolutionary Computation, CEC'03, 8-12 Dec. 2003, 4, Canberra, Australia, pp. 2292–2297 (2003)
29. Taguchi, G.: Taguchi Methods, Research and Development. American Suppliers Institute Press, Dearborn (1991)
30. Taguchi, G.: Taguchi Methods, Signal-to-Noise Ratio for Quality Evaluation. American Suppliers Institute Press, Dearborn (1991)
31. Teodorovicˇ, D.: Transport modeling by multi-agent systems: a swarm intellgence approach. Transp. Plan. Technol. **26**(4), 289–312 (2003)
32. Teodorovicˇ, D., Dell'orco, M.: Bee colony optimisation—a cooperative learning approach to complex transportation problems. In: 10th EWGT Meeting, Poznan, pp. 13–16 (2005)
33. Tereshko, V.: Reaction-diffusion model of a honeybee colony's foraging behaviour. In: Schoenauer, M., et al. (eds.) Parallel Problem Solving from Nature VI. Lecture Notes in Computer Science, vol. 1917, pp. 807–816. Springer, Berlin (2000)
34. Tereshko, V., Lee, T.: How information mapping patterns determine foraging behaviour of a honey bee colony. Open. Syst. Inf. Dyn. **9**, 181–193 (2002)
35. Tzeng, Y.F., Chen, F.C.: Multi-objective optimisation of high-speed electrical discharge machining process using a Taguchi fuzzy-based approach. Mater. Des. **28**, 1159–1168 (2007)
36. Vapnik, V., Cortes, C.: Support-vector networks. Mach. Learn. **20**, 273–297 (1995)
37. Wedde, H.F., Farooq, M., Zhang, Y.: BeeHive: an efficient fault-tolerant routing algorithm inspired by honey bee behavior, ant colony, optimization and swarm intelligence. In: 4th International Workshop, ANTS 2004, Brussels, Belgium, pp. 5–8 (2004)
38. Zhou, Z., Lin, T.: Survivable probability of SDN-enabled cloud networking with random physical link failure. CoRR abs/1709.03096 (2017)

Affiliations

Hsiu-Sen Chiang[1] · Arun Kumar Sangaiah[2] · Mu-Yen Chen[1] · Jia-Yu Liu[1]

Hsiu-Sen Chiang
hschiang@nutc.edu.tw

Arun Kumar Sangaiah
sarunkumar@vit.ac.in

Jia-Yu Liu
s1800b112@nutc.edu.tw

[1] Department of Information Management, National Taichung University of Science and Technology, 129 Sec. 3, Sanmin Rd, Taichung 404, Taiwan, ROC

[2] School of Computing and Engineering, VIT University, Vellore, India