Knowledge-Based Systems 67 (2014) 278–289

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Tuning extreme learning machine by an improved artificial bee colony to model and optimize the boiler efficiency

Guoqiang Li [a,*], Peifeng Niu [a,b], Yunpeng Ma [a], Hongbin Wang [a,1], Weiping Zhang [a]

[a] Key Lab of Industrial Computer Control Engineering of Hebei Province, Yanshan University, Qinhuangdao CO 066004, China
[b] National Engineering Research Center for Equipment and Technology of Cold Strip Rolling, Qinhuangdao, China

ARTICLE INFO

Article history:
Received 16 January 2014
Received in revised form 26 March 2014
Accepted 25 April 2014
Available online 9 May 2014

Keywords:
Artificial bee colony
Extreme learning machine
Greedy selection mechanism
Opposition-based learning
Coal-fired boilers

ABSTRACT

In this paper, a novel optimization technique based on artificial bee colony algorithm (ABC), which is called as PS-ABCII, is presented. In PS-ABCII, there are three major differences from other ABC-based techniques: (1) the opposition-based learning is applied to the population initialization; (2) the greedy selection mechanism is not adopted; (3) the mode that employed bees become scouts is modified. In order to illustrate the superiority of the proposed modified technique over other ABC-based techniques, ten classical benchmark functions are employed to test. In addition, a hybrid model called PS-ABCII-ELM is also proposed in this paper, which is combined of the PS-ABCII and Extreme Learning Machine (ELM). In PS-ABCII-ELM, the PS-ABCII is applied to tune input weights and biases of ELM in order to improve the generalization performance of ELM. And then it is applied to model and optimize the thermal efficiency of a 300 MW coal-fired boiler. The experimental results show that the proposed model is very convenient, direct and accurate, and it can give a general and suitable way to predict and improve the boiler efficiency of a coal-fired boiler under various operating conditions.

1. Introduction

The artificial bee colony algorithm (ABC) [1] proposed by Karaboga has been become available and promising techniques for real-world optimization problems. Due to the simple concept, easy implementation and quick convergence, ABC has attracted much attention and wide applications in various fields [2–11]. In ABC, the exploration process refers to the ability of seeking for the global optimum in the solution space of various unknown optimization problems, while the exploitation process refers to the ability of applying the knowledge of previous solutions to look for better solutions. However, there are still some insufficiencies, namely, ABC is good at exploration but poor at exploitation and its convergence speed is also an issue in some cases, and sometimes traps into local optimum solutions. For these insufficiencies, a few modified or improved algorithms based on ABC are presented in recent years, such as best-so-far ABC [12], GABC [13], BSO [14], IABC and PS-ABC [15]. These improved ABCs have better performances than the original ABC, especially that the PS-ABC could achieve very good search results in need of less iterations, however, it needs more time to calculate and compare the fitness of candidate solutions.

Recently, Extreme Learning Machine (ELM) proposed in the literature [16] is a kind of single hidden layer feedforward neural network (SLFN). Compared with BP, RBF and SVM, the ELM has a much faster learning algorithm and better reliability and generalization capability. In addition, the ELM overcomes many difficulties encountered by gradient-based learning methods, such as stopping criteria, learning rate, number of epochs and local minima. Due to these advantages, ELM has been successfully applied to various domains, such as classification [17], online learning [18], function approximation [19,20], Sales forecasting [21], nontechnical loss analysis [22], terrain reconstruction [23] and protein structure prediction [24]. An ELM is usually formed by an input layer, a hidden layer and an output neuron. Each neuron in the hidden and output layers is defined as a traditional neuron node, with a given activation function defined by the user. In ELM, the input weights and bias of hidden neurons are randomly generated, and the output weights are analytically calculated by means of the Moore–Penrose generalized inverse. The minimum norm least squares solution of a general linear system plays an important role in fastening the SLFN learning process. Although ELM is fast and presents good generalization performance, there may exist some non-optimal or unnecessary input weights and biases. Also, ELM may require more

* Corresponding authors. Tel.: +86 1393 3628751; fax: +86 0335 8072979.
E-mail addresses: zhihuiyuang@163.com (G. Li), hb_wang@ysu.edu.cn (H. Wang).
[1] Co-corresponding author.

http://dx.doi.org/10.1016/j.knosys.2014.04.042
0950-7051/

hidden neurons than conventional tuning-based learning algorithms in some applications, which may make ELM respond slowly to unknown testing data. For these problems, many modified or improved ELMs have been proposed in literatures. In order to further enhance the performance of ELM, there are four types improved ELM models: ensemble type, optimization type, incremental type, replacement type. In the ensemble type, different ELMs are trained by disjoint subsets of data, but share the same hidden layer neurons [25–28]. In the optimization type, various optimization methods are employed to adjust input weights and hidden layer bias of ELM and optimize the network structure [29,30]. In the incremental type, the improved ELMs create new hidden layer neurons one by one according to certain criteria.[31,32]. In the replacement type, modified ELMs substitute the activation functions of ELM (sigmoid and RBF) for sine and cosine functions (or other functions), which is very helpful to improve the accuracy and the convergence rate for the problem of function approximation [33].

In this paper, for this shortage of PS-ABC, an improved PS-ABC algorithm (PS-ABCII) is proposed to achieve faster convergence speed and better solution accuracy with a minimum incremental computational burden which is more proper for real-world optimization problems. In PS-ABCII, there are three differences from the PS-ABC: (1) the greedy selection mechanism is adopted in PS-ABC but not in PS-ABCII. Namely, a new storage unit and a sort algorithm are introduced into PS-ABCII to replace the greedy selection mechanism to search candidate solutions in order to fasten the convergence speed and the running time of the PS-ABCII; (2) the initial population is generated based on the opposition-based learning method to further enhance the global convergence; (3) the mode that employed bees become scouts is modified in order to further keep the diversity of the population. Compared with ABC and PS-ABC, PS-ABCII needs much less convergence time and running time, and its search performance is more outstanding than the other two algorithms.

In addition, a novel PS-ABCII-ELM model, which adopts the PS-ABCII algorithm to tune input weights and bias of ELM, is proposed. In order to illuminate the feasibility of the proposed model, it is applied to identify the thermal efficiency of a 300 WM coal-fired boiler. Experimental results show that the PS-ABCII-ELM model can achieve good prediction effects under various operating conditions, and then it could be adopted to predict the boiler efficiency of a coal-fired boiler under various operating conditions.

The rest of this paper is arranged as follows. In Section 2, some basic concepts and related works are reviewed. The PS-ABCII is proposed in Section 3. Experimental study shows the PS-ABCII validity in Section 4. Section 5 presents PS-ABCII-ELM model and applies it to model and predict the boiler efficiency. Finally, Section 6 concludes this paper.

2. Basic concepts and related works

2.1. Extreme learning machine

Extreme Learning Machine (ELM) proposed by Huang et al. [16]is a novel single hidden layer feedforward neural network (SFLN) learning algorithm, whose structure is given in Fig. 1. In ELM, one key principle is that the input weight values and hidden layer biases are randomly assigned. After that, the SLFN would become a linear system and its output weights could be analytically calculated by the least square method. Here, we briefly describe the learning algorithm of ELM.

Given a training data with N training samples $(\mathbf{x}_i, \mathbf{t}_i)$ where $\mathbf{x}_i = [x_1^i, x_2^i, \ldots, x_n^i]$ is an n-dimensional feature vector of the ith sample and $\mathbf{t}_i = [t_1^i, t_2^i, \ldots, t_L^i]$ is the target vector. Here, the matrix

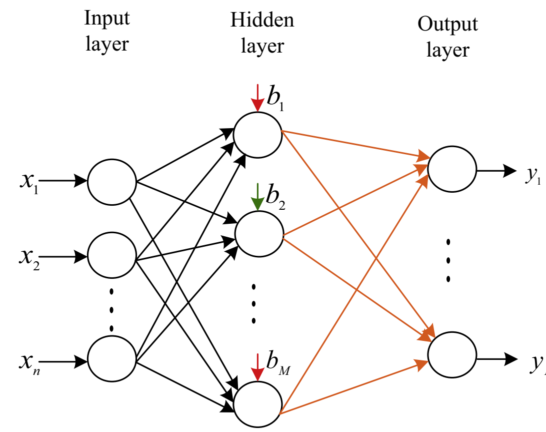

Fig. 1. Structure of extreme learning machine.

$\mathbf{W}$ and $\mathbf{B}$ are randomly assigned. Then the output ($\mathbf{T}$) of the ELM with M hidden neurons could be calculated by the following form:

$$\mathbf{t}_k^i = \sum_{j=1}^{M} \beta_{kj} g_j(\mathbf{W}, \mathbf{B}, \mathbf{X}), \quad k = 1, 2, \ldots, L \tag{1}$$

where $g_j(\cdot)$ is the activation function, $\mathbf{W}$ is $M \times n$ input weight matrix, $\mathbf{B}$ is $M \times 1$ hidden layer bias matrix and $\boldsymbol{\beta}$ is $L \times M$ output weight matrix.

The Eq. (1) could be rewritten in matrix form as

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \tag{2}$$

where $\mathbf{H}$ is the hidden layer output matrix and defined as:

$$\mathbf{H}(\mathbf{W}, \mathbf{B}, \mathbf{X}) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_M \cdot \mathbf{x}_1 + b_M) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_M \cdot \mathbf{x}_N + b_M) \end{bmatrix}_{N \times M} \tag{3}$$

$$\boldsymbol{\beta} = [\boldsymbol{\beta}_1 \quad \cdots \quad \boldsymbol{\beta}_M]_{L \times M}^T \tag{4}$$

and

$$\mathbf{T} = [\mathbf{t}_1 \quad \cdots \quad \mathbf{t}_N]_{L \times N}^T. \tag{5}$$

The output weight matrix $\boldsymbol{\beta}$ could be estimated analytically by the minimum norm least-square solution:

$$\tilde{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\| = \mathbf{H}^{+}\mathbf{T} \tag{6}$$

where $\mathbf{H}^{+}$ is the Moore–Penrose generalized inverse of $\mathbf{H}$. If the $\mathbf{H}$ is nonsingular, the Eq. (6) could be rewritten as

$$\tilde{\boldsymbol{\beta}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{T} \tag{7}$$

The learning algorithm of ELM can be summarized as the following three steps:

1) Randomly assign the input weight values $\mathbf{W}$ and hidden layer biases $\mathbf{B}$.
2) Calculate the output matrix $\mathbf{H}$ of the hidden layer.
3) Determine the output weight matrix $\boldsymbol{\beta}$.

2.2. Opposition-based learning

Opposition-based learning (OBL), which is a new concept in computational intelligence, was first proposed by Tizhoosh in 2005. OBL has been proved to be an effective method to enhance various optimization techniques [34,35].