Wind Turbine Blade Fault Diagnosis Using Support Vector Machine Optimized by Dynamic Cauchy Bee Swarm Algorithm: A Postprint

Authors: Wang Yupeng, Wang Zhijie, Liu Qi, Xu Lili, Wang Hong, Cheng Yali

Date: 2019-03-05T00:00:00+00:00

Abstract

To improve the fault recognition rate of wind turbine blades, support vector machine (SVM) is utilized to establish a nonlinear relationship between blade faults and characteristic parameters. A dynamic Cauchy factor is introduced into the bee colony algorithm to dynamically adjust the search step size during the optimization process, thereby enhancing the disturbance capability of the algorithm and preventing it from becoming trapped in local search. This dynamic Cauchy bee colony algorithm is employed to optimize the parameters of the support vector machine, establishing a support vector machine model optimized by the dynamic Cauchy bee colony algorithm. Characteristic data under four operating conditions of wind turbine blades from a wind farm in southern China were collected to train this model and conduct fault diagnosis. The diagnostic results demonstrate that the improved bee colony algorithm-optimized support vector machine model can enhance the fault recognition rate of wind turbine blades and possesses certain engineering reference significance.

Full Text

Fault Diagnosis of Wind Turbine Blade Based on Dynamic Cauchy Artificial Bee Colony Algorithm Optimized Support Vector Machine

Wang Yupeng[1], Wang Zhijie[1], Liu Qi[2], Xu Lili[2], Wang Hong[3], Cheng Yali[1]

[1]School of Electrical Engineering, Shanghai Dianji University, Shanghai 201306, China

[2]School of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China

Abstract: To improve the recognition rate of wind turbine blade fault diagnosis, this paper employs Support Vector Machine (SVM) to establish the nonlinear relationship between blade faults and characteristic parameters. A dynamic Cauchy factor is introduced into the Artificial Bee Colony (ABC) algorithm to dynamically adjust the search step size during the optimization process, thereby enhancing the algorithm' s perturbation capability and preventing the colony from falling into local optima. This Dynamic Cauchy Artificial Bee Colony (DCABC) algorithm is used to optimize the parameters of SVM, establishing a DCABC-optimized SVM model. Feature data collected under four operating conditions from wind turbine blades at a southern wind farm are used to train this model for fault diagnosis. The diagnostic results demonstrate that the improved ABC algorithm optimizing the SVM model can enhance the fault recognition rate of wind turbine blades, offering significant engineering reference value.

Keywords: Dynamic Cauchy factor; Artificial Bee Colony algorithm; Support Vector Machine; wind turbine; blade fault diagnosis

1 Introduction

With the continuous increase of wind power installed capacity in China, the reliability of wind turbines has become particularly critical, with blade failures accounting for a high proportion of all faults [1]. The key to wind turbine blade fault diagnosis lies in the identification and classification of blade faults. Support Vector Machine (SVM) not only describes the nonlinear relationship between faults and features, capable of solving nonlinear and high-dimensional problems, but also offers fast diagnosis speed, making it a primary research direction in electrical fault diagnosis [2]. Reference [3] achieved good recognition results by using harmonic wavelet packet decomposition for signal processing and SVM for classification. However, in the SVM classification process, the selection of model parameters significantly affects the final classification results.

D. Karaboga proposed the Artificial Bee Colony (ABC) algorithm [4], which has few parameters and strong global search capability, leading to extensive research and application in various optimization problems [5]. Reference [6] studied the application of ABC algorithm to optimize SVM for electrical fault diagnosis, solving the difficulty of SVM parameter optimization and improving the reliability of electrical fault diagnosis. This paper proposes a Dynamic Cauchy Artificial Bee Colony (DCABC) algorithm to optimize SVM and uses DCABC-SVM for wind turbine blade fault diagnosis, verifying the feasibility of this method through experiments.

2 Support Vector Machine Theory

Given a sample set $T = \{x_i, y_i | i = 1, 2, ..., m\}$, where $x_i \in \mathbb{R}^n$ represents the input vector and $y_i \in \{+1, -1\}$ represents the corresponding expected output vector, with m being the number of samples. Through a nonlinear mapping function, the input data is mapped to a high-dimensional linear space where an optimal classification hyperplane is constructed:

$$\min \Phi(w) = \frac{1}{2}\|w\|^2 + C\sum_i \xi_i$$

subject to $y_i = [(wx_i + b)] \geq 1 - \xi_i$ for $i = 1, 2, 3, ..., l$, where ξ_i is the slack variable and C is the penalty factor.

A kernel function $K(x_i, x_j) = \phi(x_i)\phi(x_j)$ satisfying Mercer' s condition [8] is introduced. This paper selects the Radial Basis Function (RBF) kernel:

$$K(x_i, x_j) = \exp(-\|x_i - x_j\|^2/2\sigma^2)$$

By introducing Lagrangian multipliers a_i, the problem is transformed into a quadratic programming optimization problem:

$$\max L(a) = \sum_{i,j=1}^{l} a_i a_j y_i y_j K(x_i, x_j)$$

subject to $\sum_i a_i y_i = 0$ and $0 \leq a_i \leq C$ for $i = 1, 2, 3, ..., l$. The points corresponding to $a_i > 0$ become support vectors. Typically, the number of support vectors is less than the number of training samples. The classification decision function is obtained by solving the above:

$$f(x) = \operatorname{sgn}\left(\sum_{i,j=1}^{l} a_i a_j K(x_i, x_j) + b\right)$$

In modeling, the selection of penalty parameter C and kernel function parameter γ affects the SVM diagnosis performance [9]. The parameter C regulates the trade-off between misclassification rate and model complexity, while the kernel function parameter γ primarily influences the complexity of sample data distribution in high-dimensional space. This paper employs the improved bee colony search algorithm to optimize C and γ to obtain optimal values for the classification model in practical applications.

3 Dynamic Cauchy Bee Colony Algorithm

3.1 Basic Bee Colony Algorithm

The Artificial Bee Colony (ABC) algorithm simulates the honey-gathering behavior of real bees, comprising employed bees (leaders), onlooker bees (followers), and scout bees [10]. During initialization, the number of food sources equals the number of employed bees. Food source positions represent candidate solutions to the optimization problem. An initial population is randomly generated, and search is conducted around the better half of individuals based on fitness values using a competitive selection mechanism to retain superior individuals. Onlooker bees then search using a roulette wheel selection strategy to choose better individuals and perform greedy search around them to generate the other half. Finally, scout bees search; if a food source is abandoned, a scout bee is generated to find a new food source.

Assuming the total number of food sources is N, the initial positions are randomly generated as:

$$X_{ij} = X_j^{\min} + \mathrm{rand}(0,1)(X_j^{\max} - X_j^{\min}) \tag{6}$$

where $X_j^{\max}$ and $X_j^{\min}$ are the upper and lower bounds of the search space, and $\mathrm{rand}(0,1)$ is a random number in $[0,1]$.

Employed bees randomly select food sources for crossover search according to equation (7), updating food sources to generate new solutions. When calculating fitness values of new solutions, a greedy criterion is applied to select between new and optimal solutions.

Onlooker bees select food sources with probability P_i using the roulette wheel method. The corresponding fitness value from neighborhood search is calculated using equation (9), and the greedy criterion is applied to select between new and optimal solutions.

When the exploitation limit of a food source X_i is reached (i.e., after lim cycles without improvement), the food source is abandoned, and the employed bee becomes a scout bee that randomly generates a new solution to replace the original one in the next iteration.

3.2 Dynamic Adjustment of Search Step Size

Due to the randomness of search step size, onlooker bees searching in the neighborhood of selected food sources cannot guarantee global search in the initial stage and may fall into local search during subsequent iterations, compromising overall algorithm performance. Therefore, during the optimization process, the algorithm should expand its search range in the initial stage and narrow it near the optimal solution. Reference [11] made linear adjustments to the search step size, reducing random steps but still falling into local search near optimal solutions. This paper introduces a dynamic factor w, adjusted as:

$$w = w_{\min} + (w_{\max} - w_{\min})e^{-t} \tag{10}$$

where $w_{\min}$ is the minimum inertia weight, $w_{\max}$ is the maximum inertia weight, and t represents the iteration number.

From equation (10), in the initial stage when t is small, w has a larger value. By expanding the dynamic factor, the search range of the bee colony is broadened, enabling the algorithm to escape local search, maintain solution diversity, and avoid missing the global optimum. In the later stages of search when t is large, w becomes smaller, narrowing the dynamic factor and strengthening the local search capability of employed bees near the optimal solution, thereby improving the algorithm' s dynamic search efficiency.

3.3 Introduction of Cauchy Factor

When scout bees search for new food sources after abandonment, the randomness is strong but the perturbation ability is poor [12]. This paper introduces the Cauchy distribution into the scout bee' s search formula to enhance algorithm perturbation and help the colony escape local optima. The Cauchy distribution probability density function is:

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1 + x^2} \tag{11}$$

From equation (11), when the location parameter is 0 and scale parameter is 1, $f(x)$ becomes the standard Cauchy distribution $\text{Cauchy}(0, 1)$. The random variable generation function is $\beta = \tan[(\alpha - 1/2)\pi]$, where α is a random number in $[0, 1]$.

As shown in [Figure 1: see original paper], the probability density functions of standard normal distribution and standard Cauchy distribution are compared. The Cauchy distribution approaches zero at a slower rate on both tails than the normal distribution, which helps avoid premature convergence and facilitates escaping local extrema. The peak of the standard Cauchy distribution is lower than that of the standard normal distribution, which can improve perturbation capability.

The scout bee' s search formula becomes:

$$X_{ij} = X_j^{\min} + \text{Cauchy}(0, 1)(X_j^{\max} - X_j^{\min}) \tag{12}$$

The position update formulas for employed bees and onlooker bees become:

$$v_{ij} = x_{ij} + w(x_{ij} - x_{kj}) \tag{13}$$

where $k \in \{1, 2, ..., N/2\}$, $j \in \{1, 2, ..., n\}$, and $k \neq i$; ϕ_{ij} is a random number in $[-1, 1]$.

3.4 Performance Analysis of Dynamic Cauchy Bee Colony Algorithm

To test the performance of the Dynamic Cauchy Bee Colony Algorithm, this paper selects four benchmark test functions: Rastrigin, Sphere, Ackley, and Griewank, comparing DCABC with the standard ABC algorithm. The mathematical models and variable ranges are:

(1) Rastrigin Function

$$f(x) = \sum_{i=1}^{n} [x_i^2 - 10\cos(2\pi x_i) + 10]$$

where $x_i \in [-5.12, 5.12]$, $i = 1, 2$, with minimum value 0 at $(0, 0)$.

(2) Sphere Function

$$f(x) = \sum_{i=1}^{n} x_i^2$$

where $x_i \in [-5.12, 5.12]$, $i = 1, 2$, with minimum value 0 at $(0, 0)$.

(3) Ackley Function

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

where $x_i \in [-10, 10]$, $i = 1, 2$, with minimum value 0 at $(0, 0)$.

(4) Griewank Function

$$f(x) = 1 + \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right)$$

where $x_i \in [-600, 600]$, $i = 1, 2$, with minimum value 0 at $(0, 0)$.

Simulation parameters are set as: total food sources = 50, maximum iterations = 200, exploitation limit = 15, maximum inertia weight = 2.0, minimum inertia weight = 0.4, with each experiment run independently 50 times. Simulation curves are shown in [Figure 2: see original paper] through [Figure 5: see original paper].

As demonstrated by the figures and , DCABC achieves smaller optimal values and average values than ABC across all four test functions, with shorter convergence time. Therefore, DCABC demonstrates better optimization performance and improved efficiency in optimizing SVM parameters (penalty parameter C and kernel function parameter γ).

4 DCABC Algorithm Optimizing SVM Parameters for Wind Turbine Blade Fault Diagnosis

4.1 Using DCABC Algorithm for SVM Parameter Optimization

The specific steps are:

(1) **Algorithm Initialization**: Set exploitation limit lim, maximum iterations N, and dimension $D = 2$ (for optimizing penalty parameter C and kernel function parameter γ). Randomly generate initial solutions, write initial values of C and γ into the SVM model, train on sample data, and use classification accuracy as the fitness value. Record initial solutions and calculate corresponding fitness values.

(2) **Employed Bee Phase**: According to equation (13), employed bees search for new food sources and calculate their fitness values. If a new solution yields better fitness, it replaces the current solution; otherwise, the current solution is retained.

(3) **Probability Calculation**: Calculate fitness values for all solutions and compute corresponding probabilities using the fitness-based selection formula.

(4) **Onlooker Bee Phase**: Onlooker bees follow employed bees based on roulette wheel selection, search new solutions using equation (13), and apply the greedy criterion to update optimal solutions.

(5) **Scout Bee Phase**: If a food source’ s fitness fails to improve after multiple iterations, the employed bee becomes a scout bee and searches using equation (12).

(6) **Record Optimal Solution**: After each iteration, record the best solution found so far.

(7) **Termination Check**: If maximum iterations are reached or error conditions are satisfied, output the optimal solution and complete the optimization; otherwise, return to step (2).

The fault diagnosis process using DCABC-SVM is illustrated in [Figure 6: see original paper].

4.2 Using DCABC-SVM for Wind Turbine Blade Fault Diagnosis

After collecting and normalizing sample feature data, the DCABC algorithm parameters are initialized. DCABC is used to optimize C and γ to obtain optimal parameters. The SVM model is then trained using sample data with these optimized parameters, and finally tested using experimental sample data for diagnostic verification, as shown in the flowchart in [Figure 6: see original paper].

5 Experimental Data and Results

5.1 Experimental Data Selection and Quantification

This experiment selects a 1.25MW wind turbine operating at a coastal wind farm in southern China as the test platform. When blade faults occur, temperature, humidity, brightness, and sound levels at the leading and trailing edges change to varying degrees. These data are measured and recorded, noise is removed, feature values are extracted and normalized to form datasets. The distribution of each condition dataset is shown in , with partial test sample feature parameters listed in serving as SVM model inputs. The model identifies and diagnoses faults by recognizing different features under normal conditions versus skin cracking, leading edge cracking, and trailing edge cracking.

5.2 Training and Diagnosis of Wind Turbine Blade Samples

After training the SVM with training samples, test samples are input into the trained SVM for verification. Normal condition is coded as 1, skin cracking as 2, leading edge cracking as 3, and trailing edge cracking as 4. Both ABC-SVM and DCABC-SVM methods are applied for fault diagnosis, with results shown in [Figure 7: see original paper] and [Figure 8: see original paper] respectively.

The diagnostic results comparison in shows that DCABC-SVM improves fault diagnosis accuracy by 2.36% compared to ABC-SVM. This improvement primarily results from DCABC avoiding local search during optimization, enhancing perturbation ability, finding better parameters, and constructing a more accurate SVM fault recognition model. Additionally, training time is reduced, the number of support vectors is decreased, and classification speed is accelerated. The comparison demonstrates the effectiveness of the proposed DCABC-SVM fault diagnosis model.

6 Conclusion

This paper proposes a wind turbine blade fault diagnosis method based on DCABC-SVM. The dynamic factor introduced in the ABC algorithm expands the search range in the early stage and narrows it near the optimal solution. The Cauchy operator improves algorithm perturbation ability, helping the colony escape local search and enhancing optimization efficiency. Using this DCABC algorithm to optimize SVM' s penalty factor C and kernel function parameter

γ yields optimal parameters. Feature parameters collected from wind turbine blades form training samples for SVM model training. Verification using test samples shows that DCABC-SVM achieves higher diagnostic accuracy than traditional ABC-SVM, demonstrating significant engineering value.

References

[1] Yu Zhanlong. A review of condition monitoring and fault diagnosis techniques for wind turbine [J]. Engineering Technology: Digest, 2016, 16(10): 297-297.

[2] Xue Haoran, Zhang Keheng, Li Bin, et al. Transformer fault diagnosis based on cuckoo algorithm and support vector machine [J]. Power System Protection and Control, 2015(8): 8-13.

[3] Rao Jingen. Research on wind turbine blade damage identification based on harmonic wavelet and support vector machine [D]. Lanzhou: Lanzhou Jiaotong University, 2014.

[4] Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm [J]. Journal of Global Optimization, 2007, 39(3): 459-471.

[5] Hu W, Yu Y, Zhang S. A hybrid artificial bee colony algorithm for parameter identification of uncertain fractional-order chaotic systems [J]. Nonlinear Dynamics, 2015, 82(3): 1-16.

[6] Wu Guocheng, Fan Liangzhong. Study on electrical fault diagnosis of artificial bee colony optimization support vector machine [J]. Power System and Clean Energy, 2016, 32(7): 87-91.

[7] Ye Dongyi, Chen Zhaojiong. An effective combination of artificial bee colony algorithm for minimal attribute reduction [J]. Acta Electronica Sinica, 2015, 43(5): 1014-1020.

[8] Xu Jianhua, Zhang Xuegong. Nonlinear kernel form of classical linear algorithm [J]. Control & Decision, 2006, 21(1): 1-6.

[9] Gu Guimei, Hu Rang, Li Yuanyuan. Description of damage identification of blower vane based on drosophila optimization algorithm combined with SVM [J]. Process Automation Instrumentation, 2016, 37(2): 9-12.

[10] Xiao Jian, Zhou Jianzhong, Zhang Xiaoyuan, et al. A fault diagnosis method based on Levy-ABC SVM for hydroelectric units [J]. Journal of Vibration, Testing and Diagnosis, 2013, 33(5): 839-844.

[11] Zhang Yinxue, Tian Xuemin, Cao Yuping. Artificial bee colony algorithm with improved search strategy [J]. Journal of Computer Applications, 2012, 32(12): 3326-3330.

[12] Zhang Xin, Chen Guochu, Gong Weixiang. An adaptive cauchy bee colony and its convergence analysis [J]. Computer Systems & Applications, 2015, 24(12): 118-124.

Note: Figure translations are in progress. See original paper for figures.