Wind Turbine Blade Fault Diagnosis Based on Support Vector Machine Optimized by Dynamic Cauchy Bee Colony Algorithm - Postprint

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

Date: 2018-06-15T00:00:00+00:00

Abstract

To improve the fault recognition rate of wind turbine blades, support vector machine is utilized to establish the 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 algorithm' s disturbance capability and avoiding local search entrapment. This dynamic Cauchy bee colony algorithm is employed to optimize the parameters of 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 are 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, holding certain engineering reference significance.

Full Text

Fault Diagnosis of Wind Turbine Blade Based on 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, 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 SVM parameters, 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 fault recognition rates for wind turbine blades, offering valuable engineering reference significance.

Keywords: Dynamic Cauchy factor, artificial bee colony algorithm, support vector machine, wind turbine, blade fault diagnosis

---

1 Introduction

With the continuous increase in wind power installed capacity in China, the reliability of wind turbines has become particularly critical, with blade failures accounting for a significant 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] utilized harmonic wavelet packet decomposition for signal processing and SVM for classification, achieving good recognition results. However, in SVM classification, parameter selection significantly impacts the final classification outcome.

D. Karaboga proposed the Artificial Bee Colony (ABC) algorithm [4], which features few parameters and strong global search capabilities, leading to extensive research and application in various optimization problems [5]. Reference [6] investigated the application of ABC-optimized SVM for electrical fault diagnosis, addressing the challenge 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 applies DCABC-SVM to wind turbine blade fault diagnosis, with experiments validating the feasibility of this approach.

---

2 Support Vector Machine Theory

Given a sample set $T = \{x_i, y_i | i = 1, 2, \dots, m\}$, where $x_i \in \mathbb{R}^n$ represents the input vector and $y_i \in \{+1, -1\}$ represents the corresponding expected output vector, and m is 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:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\xi_i$$

subject to:

$$y_i[(w \cdot x_i) + b] \geq 1 - \xi_i, \quad i = 1, 2, \dots, l$$

where ξ_i is the slack variable and C is the penalty factor.

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

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$$

By introducing Lagrangian multipliers a_i, the problem transforms 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:

$$\sum_{i=1}^{l} a_i y_i = 0, \quad 0 \leq a_i \leq C, \quad i = 1, 2, \dots, l$$

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:

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

In modeling, the selection of penalty parameter C and kernel function parameter γ affects SVM diagnostic performance [9]. The C value 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 an improved bee colony search algorithm to optimize C and γ to obtain optimal classification model values.

---

3 Dynamic Cauchy Bee Colony Algorithm

3.1 Basic Bee Colony Algorithm

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

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

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

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 Eq. (7), generating new solutions. When calculating fitness values, a greedy criterion selects between new and optimal solutions.

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj})$$

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

Onlooker bees select food sources with probability P_i using the roulette wheel method, calculate corresponding fitness values from neighborhood search using Eq. (9), and apply the greedy criterion to select between new and optimal solutions.

$$P_i = \frac{\mathrm{fit}_i}{\sum_{i=1}^{N} \mathrm{fit}_i}$$

where fit_i is the fitness function value of X_i.

$$\text{fit}_i = \begin{cases} \frac{1}{1+f_i} & \text{if } f_i \geq 0 \\ 1 + |f_i| & \text{if } f_i < 0 \end{cases}$$

where f_i is the objective function value of the i-th solution.

When food source X_i remains unchanged after lim continuous cycles, it is abandoned, and the employed bee becomes a scout bee that generates a new solution through Eq. (6) to replace the original food source.

3.2 Dynamic Adjustment of Bee Colony Search Step

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 colony optimization, it is desirable to expand the search range in the initial stage and narrow it near the optimal solution. Reference [11] made linear adjustments to search step size, reducing random steps but still falling into local search near optimal solutions. This paper introduces a dynamic factor w:

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

$$t = f_i$$

where $w_{\min}$ is the minimum inertia weight, $w_{\max}$ is the maximum inertia weight, and f_i is the current fitness value.

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

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

From Eq. (10), in the initial stage when t is small, w is large. The expanded dynamic factor broadens the colony' s search range, enabling the algorithm to escape local optima, ensuring solution diversity and avoiding missing the global optimum. In later search stages when t is large, w becomes smaller, narrowing the dynamic factor and strengthening local search capability near the optimal solution, thereby improving algorithm efficiency.

3.3 Introduction of Cauchy Factor

When the colony exploits the same food source reaching the exploitation limit, employed bees become scout bees that randomly generate new solutions with strong randomness but poor perturbation [12]. This paper introduces the Cauchy distribution into the scout bee search formula to enhance perturbation capability and help the colony escape local optima. The Cauchy distribution probability density function is:

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

From Eq. (13), when $t = 1$, $f(x)$ becomes the standard Cauchy distribution Cauchy(0,1). The function generating this Cauchy random variable is $\beta = \tan[(\alpha - 1/2)\pi]$, where α is a random number in $[0, 1]$.

[Figure 1: see original paper] shows the probability density functions of standard normal distribution and standard Cauchy distribution. The Cauchy distribution' s tails approach zero slower than the normal distribution' s, preventing premature convergence and facilitating escape from local extrema. The peak of the standard Cauchy distribution is lower than that of the standard normal distribution, enhancing perturbation capability.

The scout bee search formula becomes:

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

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—to compare DCABC with the standard ABC algorithm. The mathematical models and variable ranges are:

(1) Rastrigin Function

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

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

(2) Sphere Function

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

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

(3) Ackley Function

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

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

(4) Griewank Function

$$f(x) = 1 + \frac{1}{200}\sum_{i=1}^{2} x_i^2 - \prod_{i=1}^{2}\cos\left(\frac{x_i}{\sqrt{i}}\right)$$

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

Simulation parameters: food source population = 50, iterations = 200, exploitation limit = 15, maximum inertia weight = 2.0, minimum inertia weight = 0.4, with 50 independent runs per experiment. Simulation curves are shown in [Figure 2: see original paper] through [Figure 5: see original paper].

[Figure 2: see original paper] shows the simulation curves of both algorithms on the Rastrigin test function. [Figure 3: see original paper] shows the simulation on the Sphere test function. [Figure 4: see original paper] shows the simulation on the Ackley test function. [Figure 5: see original paper] shows the simulation on the Griewank test function.

As seen in [Figure 2: see original paper]–[Figure 5: see original paper] and , DCABC achieves smaller optimal values and mean values than ABC across all four test functions, with shorter convergence time. Therefore, DCABC demonstrates superior optimization performance, more efficiently optimizing SVM parameters C and γ.

---

4 DCABC Algorithm for Wind Turbine Blade Fault Diagnosis

4.1 SVM Parameter Optimization Using DCABC Algorithm

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 parameter γ). Generate initial solutions randomly, write initial C and γ values into the SVM model, train on samples, and use classification accuracy as the fitness value. Record initial solutions and calculate their fitness.

2. **Employed Bee Phase**: According to Eq. (12), employed bees search for new food sources and calculate their fitness. If a new solution yields better fitness, replace the old solution; otherwise, retain the current solution.

3. **Probability Calculation**: Calculate fitness values for all solutions and compute selection probabilities using Eq. (8).

4. **Onlooker Bee Phase**: Onlooker bees follow employed bees using roulette wheel selection, search new solutions via Eq. (12), and calculate fitness. Replace old solutions if new ones are better.

5. **Scout Bee Phase**: If a food source's fitness fails to improve after multiple iterations, employed bees become scouts and search using Eq. (14).
6. **Record Best Solution**: After each iteration, record the best solution found so far.
7. **Termination**: If maximum iterations are reached or error conditions are satisfied, output the best solution and terminate; otherwise, return to step (2).

4.2 Wind Turbine Blade Fault Diagnosis Using DCABC-SVM

The diagnostic process involves collecting sample feature data, performing normalization, initializing DCABC parameters, optimizing C and γ using DCABC, training the SVM model with optimal parameters, and finally testing with experimental samples. The flowchart is shown in [Figure 6: see original paper].

---

5 Experimental Data and Results

5.1 Data Selection and Quantification

This experiment uses a 1.25 MW 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, recorded, denoised, and feature-extracted, then normalized to form datasets. The distribution of each condition dataset is shown in , with partial test sample features in . These serve as inputs to the SVM model for fault identification and diagnosis by recognizing different features under normal conditions versus skin cracking, leading edge cracking, and trailing edge cracking.

5.2 Blade Sample Training and Diagnosis

After training the SVM with training samples, test samples are input for validation. 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 are used for diagnosis, with results shown in [Figure 7: see original paper] and [Figure 8: see original paper].

Diagnostic results are summarized in . Compared with ABC-SVM, DCABC-SVM improves fault diagnosis accuracy by 2.36%, primarily because DCABC avoids local optima during parameter optimization, enhances perturbation capability, finds better parameters, and builds a more accurate SVM fault recognition model. It also reduces training time and support vector count, accelerating classification speed. 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 early iterations and narrows it near the optimal solution. The Cauchy operator improves perturbation capability, helping the colony escape local optima and enhancing optimization efficiency. Using DCABC to optimize SVM penalty factor C and kernel parameter γ yields optimal parameters. Feature parameters collected from wind turbine blades form training samples for SVM model development. Validation with test samples shows that DCABC-SVM achieves higher diagnostic accuracy than traditional ABC-SVM, offering 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.