Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 9718345, 9 pages
https://doi.org/10.1155/2020/9718345

*Research Article*

# Fault Detection of the Wind Turbine Variable Pitch System Based on Large Margin Distribution Machine Optimized by the State Transition Algorithm

**Mingzhu Tang ,[1] Jiahao Hu ,[1] Zijie Kuang,[1] Huawei Wu ,[2] Qi Zhao,[1] and Shuhao Peng[1]**

[1] *School of Energy and Power Engineering, Changsha University of Science & Technology, Changsha 410114, China*
[2] *Hubei Key Laboratory of Power System Design and Test for Electrical Vehicle, Hubei University of Arts and Science, Xiangyang 441053, China*

Correspondence should be addressed to Huawei Wu; whw_xy@hbuas.edu.cn

Received 31 July 2020; Revised 31 August 2020; Accepted 8 September 2020; Published 13 October 2020

Academic Editor: Yong Chen



Aiming at solving the problem that the parameters of a fault detection model are difficult to be optimized, the paper proposes the fault detection of the wind turbine variable pitch system based on large margin distribution machine (LDM) which is optimized by the state transition algorithm (STA). By setting the three parameters of the LDM model as a three-dimensional vector which was searched by STA, by using the accuracy of fault detection model as the fitness function of STA, and by adopting the four state transformation operators of STA to carry out global search in the form of point, line, surface, and sphere in the search space, the global optimal parameters of LDM fault detection model are obtained and used to train the model. Compared with the grid search (GS) method, particle swarm optimization (PSO) algorithm, and genetic algorithm (GA), the proposed model method has lower false positive rate (FPR) and false negative rate (FNR) in the fault detection of wind turbine variable pitch system in a real wind farm.

## 1. Introduction

With the rapid development of the wind power industry, the installed capacity and quantity of wind turbines are growing continuously. The World Wind Energy Association predicts that by the end of 2020, the global installed capacity will reach $1.9 \times 10^6$ mw [1]. However, the availability of wind turbines is not ideal due to the increasing failure rate and maintenance cost of wind turbines along with the development of wind farms. The wind turbine variable pitch system is one of the important parts of the wind turbine, which has a complex internal mechanical structure and operated in a harsh environment that will lead to its failure rate significantly higher than other wind turbine subsystems. Since the safe and stable operation of the variable pitch system directly affects the operation efficiency of wind turbines, fault detection of the variable pitch system is of great significance for stable and efficient generation of wind turbines [2, 3].

The fault detection method is generally divided into the model-based method and the data-driven method [4]. The model-based fault detection method needs to establish an accurate mathematical model for the diagnosis object through mathematical and physical knowledge and detect faults by observing the change of the residual value [5]; the residual value of an equipment under normal state should be zero or close to zero, and it is not zero when the equipment is disturbed or malfunctioned. This method can be divided into the parameter estimation method [6], state estimation method [7], and equivalent space method [8]. The model-based fault detection method can quickly get a more accurate mathematical model and detect faults accurately for the system with simple structure. However, for the fault detection of large-scale wind turbines, the modeling process is easy to be affected by various parameters which will influence the robust performance and the accuracy of the fault detection and even makes it difficult to locate the wind turbine internal fault causes.

The data-driven fault detection method extracts useful information through various data processing and analysis methods based on the collected data, compares the collected historical data with the real-time data of the system, and analyzes their potential relationship so as to carry out fault detection. Being capable of detecting the fault of the equipment through data analysis, this method does not need to establish an accurate mathematical model. It does not depend on the complexity and uncertainty of the system, so a good detection performance is obtained of it. As the method meets the needs of the industrial big data era, it is widely used in the industrial field. Artificial neural networks (ANNs), SVM, LDM, and other models are usually used for fault detection of equipment in the data-driven fault detection method.

The ANN is a classic data-driven model based on mimicing the biological nervous system. It can automatically analyze and infer the input information to detect faults by simulating the physiological structure and thinking mode of the human brain. The application of ANNs in wind turbine fault detection has a good detection performance. Concerning the problem of sensor fault of the wind turbine, Qiu et al. proposed a damage prediction method for the offshore wind turbine tower structure based on ANNs, which can improve the accuracy of fault prediction [9]. In the case of gearbox fault, Chen et al. proposed a fault diagnosis method based on wavelet analysis and neural networks to diagnose the wind turbine gearbox and predict the early fault signs and obtained good results [10]. The ANN has the ability of self-study which is similar to the human brain. It has good robustness to the interference and noise of the system. However, due to the nature of the black box, this method is difficult to make a good explanation for specific faults. Moreover, it has high requirements for data in actual use and requires high running cost.

SVM is another classic data-driven model based on global optimization. It has good performance and can solve the problems of multiclassification recognition and regression prediction [11, 12], which has been widely recognized in the field of wind turbine fault research. Hang et al. proposed a wind turbine fault diagnosis method based on a multiclass fuzzy SVM classifier to improve the accuracy of fault diagnosis [13]. Rotating parts in wind turbines are one of the key objects in fault diagnosis of wind turbines. However, in fact, the vibration signals collected from the rotating parts are generally non-Gaussian and nonstationary, and the fault samples are very limited. Liu et al. proposed a wind turbine fault diagnosis method based on a diagonal spectrum and clustering binary tree SVM, which achieved good results [14]. Although having a good performance on simple binary classification problems, SVM is ineffective in dealing with large-scale data problems and sensitive to model parameters and data integrity.

The distribution machine supported by large margin theory can find the distribution model according to the sample distribution characteristics while considering the sample mean value and sample variance. Compared with the former two models, LDM has higher fault detection performance. In the fault detection of wind turbines, Tang et al. proposed a cost-sensitive large margin distribution machine (CLDM) to solve the problems of class imbalance data and misclassification unequal cost of large wind turbine data sets, which has effectively improved the fault detection performance [15].

The data-driven fault detection model has good practicability in actual wind turbine fault detection and fault diagnosis. However, most of these models depend on the selection of parameters, so it is necessary to use the parameter optimization algorithm to quickly and accurately find the global optimal model parameters. The GS, PSO, and GA are most commonly used to optimize the parameters of the fault detection model in wind turbine fault detection. Aguilar et al. proposed a multiobjective particle swarm optimization (MOPSO) algorithm for the electrical fault of variable-speed wind turbines, which improved the stability of wind turbines [16]. Concerning the problem that the traditional threshold setting is difficult to identify the abnormal operation of wind turbines, Zhang et al. put forward a new backpropagation neural network (BPNN) anomaly identification model combined with GA, which provides good performance effect for abnormal identification of wind turbines [17]. Yan et al. optimized the parameters of SVM by the GS method in wind turbine fault detection to improve the diagnostic accuracy [18]. PSO, GA, and GS can achieve approximate global optimal solution for parameter optimization of a simple model, but it is easy to fall into local optimum when used in fault detection of large and complex wind turbines.

The STA is a parameter optimization algorithm with four state transition operators, facing the complex fault detection problem; the global optimal value can be quickly and accurately found by the four state transformation operators alternately, which is suitable for detecting the complex fault of the wind turbine variable pitch system. Because of its strong performance and practicability, the STA has solved many problems in the industry and other fields [19, 20].

It is of great significance to choose a fault detection model with proper performance. However, in the fault detection model based on machine learning, parameter optimization is an important process, and how to select appropriate parameters to enable the detection model to meet the fault detection standard of the wind turbine variable pitch system is the key and difficult problem of all machine learning models. Therefore, an improved LDM model based on the STA is studied with an aim to effectively finding out the optimal model parameters, making it meet the fault characteristics of the variable pitch system, and improving the accuracy of fault detection.

## 2. Large Margin Distribution Machine

If the traditional machine learning algorithm based on margin theory for optimization is adopted, attention should be paid to find the minimum margin between samples, such as SVM, which can be adopted to find the hyperplane that maximizes the minimum margin between two kinds of samples in the optimization process [21]. However, the method only focuses on the support vectors that only account for a small proportion in a large number of samples, while the rest of the sample information is not considered in the learning process. The above method will lead to the loss of some samples of useful information as well as reduction of the learning ability of the algorithm for samples; in addition, the learning effect remains to be improved.

The LDM proposed by Zhang and Zhou is used to find the separation hyperplane according to the distribution characteristics of samples under the premise of considering the margin distribution of the whole sample [22]. Compared with the support vector machine which only optimizes the minimum margin, it has stronger generalization performance. Figure 1 shows the different results of the final classification hyperplane due to different margin considerations in the classification process.

In Figure 1, the triangle icon refers to the first type of sample, the square icon refers to the second type of sample, the elliptical dotted line shows the potential distribution of the two types of samples, and the red triangle and red square indicate the distribution mean of the two types of samples. If the classification hyperplane is searched based on the minimum margin between the two types of samples as $h_{\min}$ in the figure, it can be found to intersect with the potential distribution range of the right sample, and there is the possibility of misclassification; if the overall distribution of two types of samples has been considered in the classification hyperplane with the classification plane as $h_{\text{dist}}$ in the figure, the conclusion can be drawn that it has better classification performance and stronger robustness for two types of samples [23].

For LDM, we should set $X = [\phi(x_1), \ldots, \phi(x_m)]$ and set the algorithm $y = [y_1, \ldots, y_m]$ as the $m$-dimensional column vector, where $\phi(\cdot)$ is the feature mapping through a positive definite function $\kappa(\cdot, \cdot)$, $Y$ is an $m$-order square matrix, and the diagonal is $y_1, \ldots, y_m$; therefore, the mean value of the margin can be defined as follows:

$$\gamma_m = \frac{1}{m} \sum_{i=1}^{m} y_i \omega^{\mathrm{T}} \phi(x_i) = \frac{1}{m} (Xy)^{\mathrm{T}} \omega. \tag{1}$$

The margin variance is as follows:

$$\gamma_v = \frac{1}{m} \sum_{i=1}^{m} \left( y_i \omega^{\mathrm{T}} \phi(x_i) - \gamma_m \right)^2 = \omega^{\mathrm{T}} X \frac{mI - yy^{\mathrm{T}}}{m^2} X^{\mathrm{T}} \omega. \tag{2}$$

It is important to make a linear combination of margin mean and margin variance into an optimization problem, introduce L2-norm as the regularization term, and select hinge loss for the loss function; therefore, the formalization of LDM is as follows:

$$\begin{aligned} \min \quad & _{\omega,\xi_i} \frac{1}{2} \left\| \omega^2 \right\| + \lambda_1 \gamma_v - \lambda_2 \gamma_m + \frac{C}{m} \sum_{i=1}^{m} \xi_i \\ \text{s.t.} \quad & y_i \omega^{\mathrm{T}} \phi(x_i) \geq 1 - \xi_i, \\ & \xi_i \geq 0, \ \forall i \in [m], \end{aligned} \tag{3}$$

where parameters $\lambda_1$ and $\lambda_2$ are trade-off parameters and are used to adjust the weight of the margin mean and margin variance in the objective function, while $C$ is a loss function parameter. Although the theory of large margin distribution has achieved good results in theory and practice, the classification surface may show unbalanced tendency in the face of the number of unbalanced margins and samples with noise, and the robustness to noise is not strong. Therefore, the model needs further development and improvement.

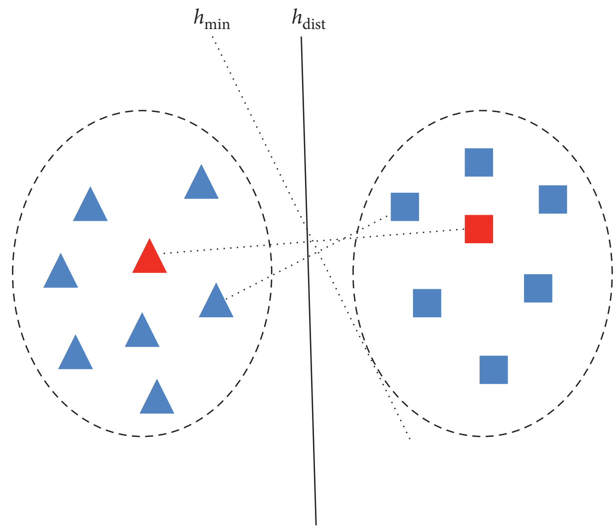


Figure 1: Minimum margin hyperplane and margin distribution hyperplane.

## 3. The State Transition Algorithm

Being a global optimization method proposed by Zhou et al. [24] the STA is an individual-based intelligent stochastic global optimization method. It uses different state transformation operators to operate independently through the given current solution, thus generating the candidate solution set and finding out the solution better than the current candidate solution in the candidate solution set, which serves as a new solution of the update iteration. The process should be repeated till the certain termination condition is met.

In brief, the STA is based on the state space of modern control theory, which treats the solving process of the optimization problem as the process of state transition and treats the generation and update of the solution as the formation and update of the state.

*3.1. The State Transformation Operator.* The state-space expression in modern control theory is used as the unified framework of the candidate set, and the state transformation operators are designed for the framework. The unified framework of candidate solutions for the STA is as follows:

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k, \\ y_{k+1} = f(x_{k+1}), \end{cases} \tag{4}$$

where $x_k = [x_1, x_2, \ldots, x_n]^T$ is the current state and represents a candidate solution in the optimization problem, $A_k$

and $B_k$ are the state transition matrices, which are random matrices and equivalent to state transformation operators, $u_k$ is a function of the historical state and current state, which is equivalent to a control variable, and $f(\cdot)$ is the objective function, that is, the fitness function.

The four state transition operators in the STA correspond to four search functions, and each state transformation operator can form a regular geometric neighborhood with unique shape and adjustable size. State transformation operators mainly include rotation transformation operator, translation transformation operator, expansion transformation operator, and axesion transformation operator.

(1) The rotation transformation operator:

$$x_{k+1} = x_k + \alpha \frac{1}{n\|x_k\|_2} R_r x_k, \tag{5}$$

where $\alpha > 0$ is the rotation factor; $R_r \in R^{n\times n}$ is a random matrix with its element values evenly distributed between $[-1, 1]$; $\|\cdot\|_2$ is the vector L2-norm, and the function of the rotation transformation operator is to search in the hypersphere with $\alpha$ as the radius.

(2) The translation transformation operator:

$$x_{k+1} = x_k + \beta R_t \frac{x_k - x_{k-1}}{\|x_k - x_{k-1}\|_2}, \tag{6}$$

where $\beta > 0$ is the translation factor; the value range of $R_t \in R$ is $[0, 1]$, meeting the uniform distribution. As a heuristic search operator, the translation transformation operator can search with $\beta$ as the maximum length from point $x_{k-1}$ to point $x_k$ along the line.

(3) The expansion transformation operator:

$$x_{k+1} = x_k + \gamma R_e x_k, \tag{7}$$

where $\gamma > 0$ is the expansion factor and $R_e \in R^{n\times n}$ is a diagonal matrix, with its element value of nonzero, complying with the Gaussian distribution. As the global search operator, the expansion transformation operator can expand each element in $x_k$ to the whole range of $(-\infty, +\infty)$, thus realizing the search of the whole space.

(4) The axesion transformation operator:

$$x_{k+1} = x_k + \delta R_a x_k, \tag{8}$$

where $\delta > 0$ is the axesion factor and $R_a \in R^{n\times n}$ is a sparse random diagonal matrix, with its element value of nonzero, complying with the Gaussian distribution. Being a heuristic search operator with relatively strong single-dimensional search ability, the axesion transformation operator can search along the axesion axis direction.

## 4. Large Margin Distribution Machine Optimized by the State Transition Algorithm

Fitness function is a main factor affecting the convergence speed and finding the optimal solution of the parameter optimization algorithm, and it is an evaluation criterion to select and update the optimal solution in the process of parameter optimization. In the STA, the mean accuracy of the LDM optimization model which was verified by 10-fold cross-validation is used as the fitness function to judge the selection and update of the current parameter state; if the accuracy is higher than that of the current optimal state, the new parameter will be used as a better solution to update the current state, and if the accuracy is lower than that of the current optimal state, the parameter will be abandoned for the next iteration. The fitness function is as follows:

$$\text{fitness} = \frac{\sum_{i=0}^{k_{cv}} \text{accuracy}\left(\text{LDM}\left(\lambda_1, \lambda_2, C\right)\right)}{k_{cv}}, \tag{9}$$

where $k_{cv} = 10$ is the number of cross-validations, $\lambda_1$, $\lambda_2$, and $C$ are the three parameters in LDM, which are the margin variance parameter, margin mean parameter, and loss function parameter, respectively. The meaning and value range of three parameters are shown in Table 1.

LDM parameters are adjusted by the STA, and the three parameters in LDM are taken as a three-dimensional vector form, a state in the STA. The new candidate solution set is generated by alternately using the four transformation operators of rotation, expansion, axesion, and translation.

The use of the fitness function of the improved LDM and the selection and updated pseudocode of the current optimal state solution are given in Algorithm 1.

However, $\text{Best}_0(\lambda_{10}, \lambda_{20}, C)$ refers to the initial state, and the three parameters of LDM are assigned from Step 6 to Step 8; the training set is adopted to train the adjusted LDM algorithm to establish the learning model in Step 9; the testing set is used to predict the model in Step 10; the classification accuracy of the predicted results is used as the evaluation criterion of fitness function in Step 11; the rotation transformation, expansion transformation, axesion transformation, and the function of selection and update are realized from Steps 12 to 14, and the discriminant rules for selection and update follow the fitness function Fitness based on predicted accuracy of LDM. If the specified termination criterion is met, the output solution $\text{Best}(\lambda_1, \lambda_2, C)$ will be the global optimal parameters to improve the LDM. Figure 2 shows the specific process of the STA selecting the optimal parameters of LDM by the fitness function.

## 5. Experimental Results and Analysis

The experimental data used the wind turbine variable pitch system fault data of one year's SCADA data set collected by a wind farm in East China, including variable pitch main power supply fault, variable pitch blade server drive temperature over-limit fault, and variable pitch system emergency stop fault. The number of fault samples and the number of fault features of the three fault data are shown in Table 2.

According to the different fault detection of the wind turbine variable pitch system, the sample set in normal operation should be classified as normal, and the sample set in failure should be classified as a fault. It is important to

Table 1: Meaning and value range of LDM parameters.

| Parameter | Meaning | Value range |
|---|---|---|
| $\lambda_1$ | The trade-off parameter of margin variance, which is adopted to adjust the weight of margin variance | $[2^{-1}, 2^{10}]$ |
| $\lambda_2$ | The trade-off parameter of margin mean, which is adopted to adjust the weight of margin mean | $[2^{-1}, 2^{10}]$ |
| $C$ | The loss function parameter, which is adopted to adjust the weight of the loss function in the objective function | $[2^{0}, 2^{20}]$ |

**Best**$\longleftarrow$**Best**$_0$ $(\lambda_{10}, \lambda_{20}, \mathbf{C})$
repeat
if $\alpha < \alpha_{\mathbf{min}}$ then
$\alpha \longleftarrow \alpha_{\mathbf{max}}$
end if
$\lambda_1 \longleftarrow$ **Best** (1)
$\lambda_2 \longleftarrow$ **Best** (2)
**C**$\longleftarrow$**Best** (3)
**L DM**$\longleftarrow$ $(\lambda_1, \lambda_2,$ **C**, **training set**)
**accuracy** (**L DM**)$\longleftarrow$**testing set**
**Fitness**$\longleftarrow$**accuracy** (**L DM**)
**Best**$\longleftarrow$**rotation transformation** (Fitness, **Best**, **SE**, $\beta$, $\alpha$)
**Best**$\longleftarrow$**expansion transformation** (Fitness, **Best**, **SE**, $\beta$, $\gamma$)
**Best**$\longleftarrow$**axesion transformation** (Fitness, **Best**, **SE**, $\beta$, $\delta$)
$\alpha \longleftarrow \alpha/\mathbf{f_c}$
Until the specified termination criterion is met
Output **Best**

Algorithm 1: Optimal parameters of the improved LDM.

divide the whole sample set into two parts with each part containing normal data and fault data, which are used as a training set and testing set, respectively. The training set is mainly used to train the fault detection model, and the testing set is used to predict the model. The parameters of the STA are set as $\alpha_{\max} = 1$, $\alpha_{\min} = 1e-4$, $\beta = 1$, $\gamma = 1$, $\delta = 1$, SE = 30, and $f_c = 2$.

In order to verify that the STA can be adopted to improve the parameter adjustment of LDM and the improved LDM is effective for fault detection of the wind turbine variable pitch system, measures should be taken to introduce GS, PSO, and GA into the model parameter optimization method for comparison. The evaluation indexes were four indexes produced by the confusion matrix, including accuracy, F1-score, FPR, and FNR:

$$\begin{aligned} \text{accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}}, \\ F1 - \text{score} &= \frac{2}{(1/(\text{TP}/(\text{TP} + \text{FP}))) + (1/(\text{TP}/(\text{TP} + \text{FN})))}, \\ \text{FPR} &= \frac{\text{FP}}{\text{TN} + \text{FP}}, \\ \text{FNR} &= \frac{\text{FN}}{\text{TP} + \text{FN}}, \end{aligned} \tag{10}$$

where TP means the actual sample is positive and the prediction is positive; FP means the actual sample is negative and the prediction is positive; TN means the actual sample is negative and the prediction is negative; FN means the actual sample is positive and the prediction is negative.

In terms of the fault of wind turbine variable pitch main power supply, the boxplot of accuracy is shown in Figure 3. The comparison results of F1-score, FPR, and FNR are shown in Table 3.

The results indicated that the detection accuracy and F1-score of the improved LDM based on the STA for the variable pitch main power supply fault were higher than the values in terms of the methods of parameter adjustment through PSO, GA, and GS. FNR and FPR were the lowest among the four parameter adjustment methods.

For wind turbine variable pitch blade server drive temperature over-limit fault situation, Figure 4 shows the accuracy boxplot. The comparison results of F1-score, FPR, and FNR are shown in Table 4.

The results indicated that the detection accuracy and F1-score of the improved LDM based on the STA for wind turbine variable pitch blade server drive temperature over-limit fault were the highest while FNR and FPR were lower than the other three parameter adjustment methods.

For the wind turbine variable pitch system emergency stop fault situation, Figure 5 shows the accuracy boxplot. The comparison results of F1-score, FPR, and FNR are shown in Table 5.

The results indicated that the detection accuracy and F1-score of the improved LDM based on the STA for wind turbine variable pitch system emergency stop fault were

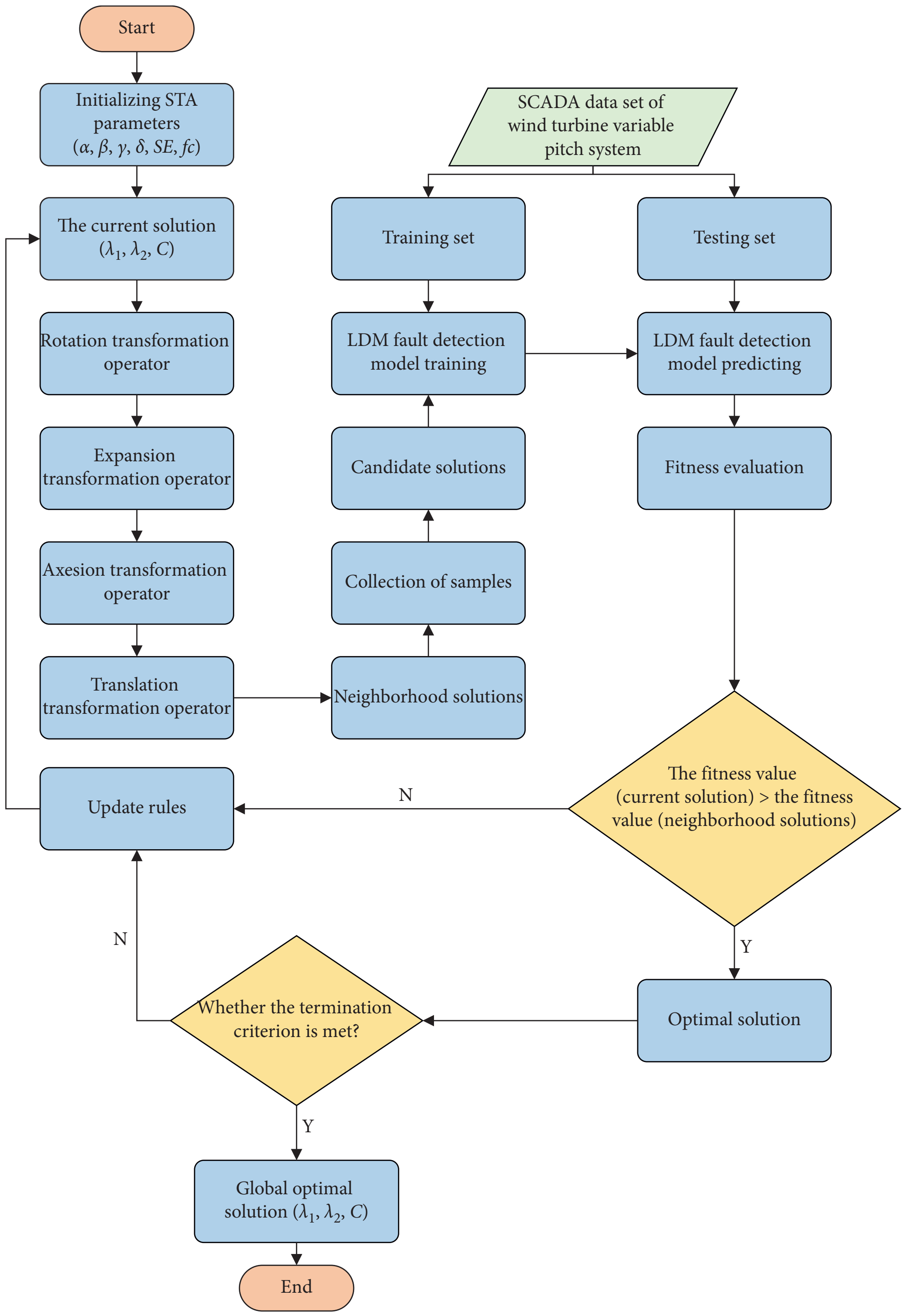


Figure 2: The specific process of the STA selecting optimal parameters of LDM by fitness function.

Table 2: The number of fault samples and the number of fault features of the three variable pitch system fault data.

| Fault type | Number of fault samples | Number of fault features |
|---|---|---|
| Variable pitch main power supply fault | 2902 | 212 |
| Variable pitch blade server drive temperature over-limit fault | 4864 | 212 |
| Variable pitch system emergency stop fault | 5893 | 212 |

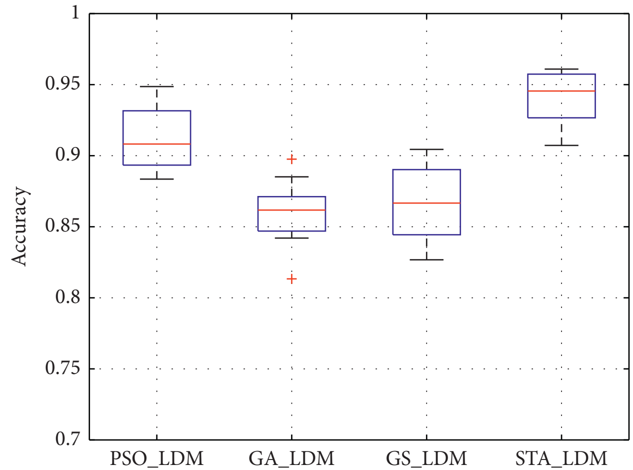


FIGURE 3: Boxplot of variable pitch main power supply fault detection accuracy.

TABLE 3: Performance comparison of variable pitch main power supply fault detection.

| Fault detection model | F1-score | FPR | FNR |
|---|---|---|---|
| PSO_LDM | 95.54% (±0.0057) | 9.39% (±0.1086) | 3.43% (±0.0296) |
| GA_LDM | 89.38% (±0.0779) | 13.19% (±0.0868) | 9.84% (±0.1456) |
| GS_LDM | 91.17% (±0.0081) | 11.31% (±0.1059) | 8.05% (±0.0159) |
| STA_LDM | 96.49% (±0.0142) | 5.07% (±0.1091) | 2.97% (±0.0143) |

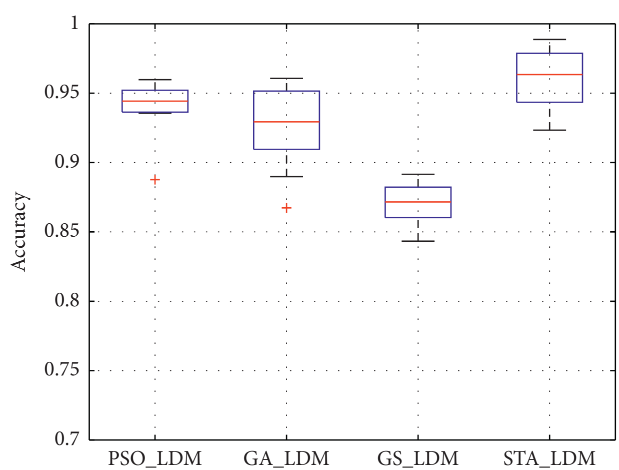


FIGURE 4: Boxplot of variable pitch blade server drive temperature over-limit fault detection accuracy.

TABLE 4: Performance comparison of variable pitch blade server drive temperature over-limit fault detection.

| Fault detection model | F1-score | FPR | FNR |
|---|---|---|---|
| PSO_LDM | 96.24% (±0.0092) | 5.44% (±0.0166) | 2.73% (±0.0371) |
| GA_LDM | 95.68% (±0.0457) | 6.71% (±0.1048) | 3.04% (±0.0018) |
| GS_LDM | 87.37% (±0.0135) | 8.75% (±0.0191) | 13.64% (±0.0713) |
| STA_LDM | 98.72% (±0.0241) | 1.10% (±0.0049) | 1.16% (±0.0118) |

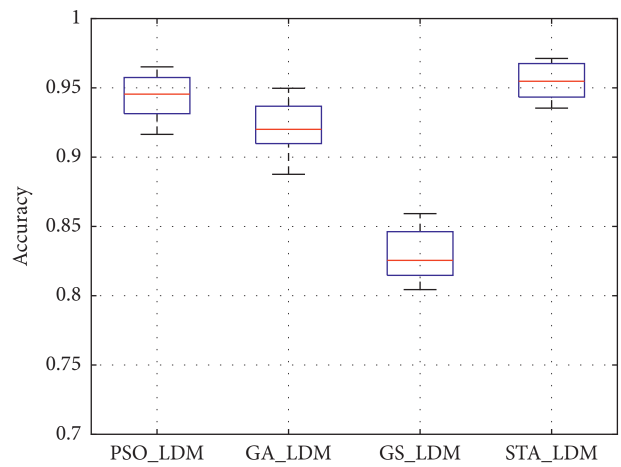


Figure 5: Boxplot of variable pitch system emergency stop fault detection accuracy.

Table 5: Performance comparison of variable pitch system emergency stop fault detection.

| Fault detection model | F1-score | FPR | FNR |
|---|---|---|---|
| PSO_LDM | 96.55% (±0.0107 ) | 2.73% (±0.1102 ) | 4.23% (±0.0544) |
| GA_LDM | 95.50% (±0.0105 ) | 6.69% (±0.0174) | 3.98% (±0.0137) |
| GS_LDM | 81.85% (±0.0134 ) | 11.72% (±0.0753) | 15.16% (±0.0096) |
| STA_LDM | 97.19% (±0.0275 ) | 2.21% (±0.0024 ) | 3.65% (±0.0168 ) |

higher than the values in terms of the methods of parameter adjustment through PSO, GA, and GS. FNR and FPR were the lowest among the four parameter adjustment methods.

## 6. Conclusion

Concerning the problem of dependent parameter selection of the fault detection model, this paper introduces the STA to improve LDM in terms of the parameter optimization of the classification algorithm. First, in order to meet the structure need of the optimization problem, the three parameters in LDM were regarded as a three-dimensional vector form, a state in the STA. In addition, a new state candidate assembly was generated by alternately using the four transformation operators. Second, the accuracy of the fault detection model output is used as a fitness function to support parameter updating and optimization. Finally, for verifying the effectiveness of the wind turbine variable pitch system fault detection method based on the improved LDM, the paper introduced the GS method, PSO, and GA for comparison on parameter optimization. The evaluation indexes were accuracy, F1-score, FPR, and FNR. The experimental data were variable pitch main power supply fault data, variable pitch blade server drive temperature over-limit fault data, and variable pitch system emergency stop fault data.

Experimental results showed that the fault detection model which used the STA for parameter optimization had higher accuracy and lower FPR and FNR than the other three optimization algorithms, which proved that the improved LDM has stronger capability of detecting wind turbine variable pitch system fault.

On account of the vulnerability of the wind turbine to be affected by the environment and load while running, it is incomprehensive to use a single detection method in the process of fault detection. As a result, it is indispensable to study a hybrid fault detection method based on various fault detection methods and technologies in the future.

## Data Availability

The data used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, [6/12 months] after publication of this article, will be considered by the corresponding author.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Authors' Contributions

All authors contributed equally to this work.

## Acknowledgments

This research was funded by the National Natural Science Foundation of China (grant no. 61403046), the Natural Science Foundation of Hunan Province, China (grant no. 2019JJ40304), the Changsha University of Science and Technology "The Double First Class University Plan" International Cooperation and Development Project in

Scientific Research in 2018 (grant no. 2018IC14), the Research Foundation of Education Bureau of Hunan Province (grant No.19K007), the Science and Technology Progress and Innovation Plan Project of Hunan Provincial Department of Transportation in 2018 (grant no. 201843), the Key Laboratory of Renewable Energy Electric-Technology of Hunan Province, the Key Laboratory of Efficient and Clean Energy Utilization of Hunan Province, the Innovative Team of Key Technologies of Energy Conservation, the Emission Reduction and Intelligent Control for Power-Generating Equipment and System, the CSUST, Hubei Superior and Distinctive Discipline Group of Mechatronics and Automobiles (grant no. XKQ2020009), the National Training Program of Innovation and Entrepreneurship for Undergraduates (grant no. 202010536016), and the Major Fund Project of Technical Innovation in Hubei (grant no. 2017AAA133), Hubei Natural Science Foundation Youth Project (2020CFB320).

## References

[1] X. Xu, D. Niu, B. Xiao, X. Guo, L. Zhang, and K. Wang, "Policy analysis for grid parity of wind power generation in China," *Energy Policy*, vol. 138, Article ID 111225, 2020.

[2] M. Nazir, A. Q. Khan, G. Mustafa, and M. Abid, "Robust fault detection for wind turbines using reference model-based approach," *Journal of King Saud University–Engineering Sciences*, vol. 29, no. 3, pp. 244–252, 2017.

[3] S. Cho, Z. Gao, and T. Moan, "Model-based fault detection, fault isolation and fault-tolerant control of a blade pitch system in floating wind turbines," *Renewable Energy*, vol. 120, pp. 306–321, 2018.

[4] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, "A review of process fault detection and diagnosis," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 293–311, 2003.

[5] Y.-Y. Ko, "A simplified structural model for monopile-supported offshore wind turbines with tapered towers," *Renewable Energy*, vol. 156, pp. 777–790, 2020.

[6] R. Isermann, "Model-based fault-detection and diagnosis -status and applications," *Annual Reviews in Control*, vol. 29, no. 1, pp. 71–85, 2005.

[7] S. Ibaraki, S. Suryanarayanan, and M. Tomizuka, "Design of luenberger state observers using fixed-Structure<tex>$ cal H_ infty $</tex>Optimization and its application to fault detection in lane-keeping control of automated vehicles," *IEEE/ASME Transactions on Mechatronics*, vol. 10, no. 1, pp. 34–42, 2005.

[8] M. Zhong, S. X. Ding, Q.-L. Han, and Q. Ding, "Parity space-based fault estimation for linear discrete time-varying systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 7, pp. 1726–1731, 2010.

[9] B. Qiu, Y. Lu, L. Sun, X. Qu, Y. Xue, and F. Tong, "Research on the damage prediction method of offshore wind turbine tower structure based on improved neural network," *Measurement*, vol. 151, Article ID 107141, 2020.

[10] H. Chen, S. Jing, X. Wang, and Z. Wang, "Fault diagnosis of wind turbine gearbox based on wavelet neural network," *Journal of Low Frequency Noise, Vibration and Active Control*, vol. 37, no. 4, pp. 977–986, 2018.

[11] X. Zhu and Z. Gao, "An efficient gradient-based model selection algorithm for multi-output least-squares support vector regression machines," in *Pattern Recognition Letters*, vol. 111, pp. 16–22, 2018.

[12] H. Chih-Wei and L. Chih-Jen, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.

[13] J. Hang, J. Zhang, and M. Cheng, "Application of multi-class fuzzy support vector machine classifier for fault diagnosis of wind turbine," *Fuzzy Sets and Systems*, vol. 297, pp. 128–140, 2016.

[14] L. Wenyi, W. Zhenfeng, H. Jiguang, and W. Guangfeng, "Wind turbine fault diagnosis method based on diagonal spectrum and clustering binary tree SVM," *Renewable Energy*, vol. 50, pp. 1–6, 2013.

[15] M. Tang, "Cost-sensitive large margin distribution machine for fault detection of wind turbines," *Cluster Computing*, vol. 22, no. 3, pp. 7525–7537, 2019.

[16] M. E. Barrios Aguilar, D. V. Coury, R. Reginatto, and R. M. Monaro, "Multi-objective PSO applied to PI control of DFIG wind turbine under electrical fault conditions," *Electric Power Systems Research*, vol. 180, Article ID 106081, 2020.

[17] Y. Zhang, H. Zheng, J. Liu, J. Zhao, and P. Sun, "An anomaly identification model for wind turbine state parameters," *Journal of Cleaner Production*, vol. 195, pp. 1214–1227, 2018.

[18] H. Yan, H. Mu, X. Yi, Y. Yang, and G. Chen, "Fault diagnosis of wind turbine based on PCA and GSA-SVM," in *Proceedings of the Prognostics and System Health Management Conference*, (PHM-Paris), Beijing, China, pp. 13–17, 2019.

[19] J.-T. Yin, Y.-F. Xie, Z.-W. Chen, T. Peng, and C.-H. Yang, "Weak-fault diagnosis using state-transition-algorithm-based adaptive stochastic-resonance method," *Journal of Central South University*, vol. 26, no. 7, pp. 1910–1920, 2019.

[20] R. Murugesan, J. Solaimalai, and K. Chandran, "Computer-aided controller design for a nonlinear process using a Lagrangian-based state transition algorithm," *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, pp. 977–996, 2020.

[21] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[22] T. Zhang and Z. Zhou, "Large margin distribution machine," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 313–322, New York, NY, USA, 2014.

[23] Z. Zhou, "Large margin distribution learning," in *Artificial Neural Networks in Pattern Recognition,* , pp. 1–11, Springer Verlag, Montreal, Canada, 2014.

[24] X. Zhou, C. Yang, and W. Gui, "State transition algorithm," *Journal of Industrial Management Optimization*, vol. 8, no. 4, pp. 1039–1056, 2013.