An improved CS-LSSVM algorithm-based fault pattern recognition of ship power equipments

Yifei Yang\textsuperscript{1,2}, Minjia Tan\textsuperscript{2}, Yuewei Dai\textsuperscript{1,2}

\textsuperscript{1} School of Automation, Nanjing University of Science and Technology, Nanjing, Jiangsu, China, \textsuperscript{2} School of Electronics and Information, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, China

* seayyf@163.com


citation: Yang Y, Tan M, Dai Y (2017) An improved CS-LSSVM algorithm-based fault pattern recognition of ship power equipments. PLoS ONE 12(2): e0171246. doi:10.1371/journal.pone.0171246

1. Introduction

Ship power equipments often work under terrible working conditions. Thus, the occurrence of faults is unavoidable. It should be mentioned that serious faults can lead to economic losses and threaten the safety of the crew’s life. However, the structure of ship power equipments is complex, and it is difficult to figure out the fault pattern. Then, it is necessary to analyze the fault pattern of ship power equipments and guarantee the reliability and safety of ship power equipments accordingly.

Ship power equipments include diesel, gearbox, air compressor, pump, etc. The fault pattern recognition for these equipments falls into a multi-classification issue which focuses on the fault samples. As observed from [1], neural network demonstrates the characteristics of nonlinear processing, adaption, robustness, etc. Then it can be widely used to deal with nonlinear equipment test signal, complex fault modes, and fault recognition. However, this method does show some shortcomings such as the complex network structure, the occurrence of local extremum, and slow converging speed [2].

Some new algorithms, especially the SSVM in [3] and the CBoPS in [4], are proposed in [3–8] to enhance the searching accuracy and searching speed. Moreover, the proposed algorithms find applications in mobile visual search [4], mobile visual location recognition (MVLR) [5–7], and mobile landmark recognition [8].

Abstract

A ship power equipments’ fault monitoring signal usually provides few samples and the data’s feature is non-linear in practical situation. This paper adopts the method of the least squares support vector machine (LSSVM) to deal with the problem of fault pattern identification in the case of small sample data. Meanwhile, in order to avoid involving a local extremum and poor convergence precision which are induced by optimizing the kernel function parameter and penalty factor of LSSVM, an improved Cuckoo Search (CS) algorithm is proposed for the purpose of parameter optimization. Based on the dynamic adaptive strategy, the newly proposed algorithm improves the recognition probability and the searching step length, which can effectively solve the problems of slow searching speed and low calculation accuracy of the CS algorithm. A benchmark example demonstrates that the CS-LSSVM algorithm can accurately and effectively identify the fault pattern types of ship power equipments.
The least squares support vector machine (LSSVM), which is a fusion of neural network and support vector machine (SVM), is an extension of SVM [9]. By using equality constraints to substitute inequality constraints and solving linear equations to obtain the support vector, one can simplify the calculation and enhance the training speed [10].

The parameter optimization of LSSVM can affect greatly the prediction model performance of LSSVM. Considering that the optimization of LSSVM parameters is necessary, this paper adopts the Cuckoo search algorithm (CS).

2. Description of the LSSVM method

The LSSVM method introduces the idea of square sum of errors to the objective function of the standard SVM, where a training data with \( n \) samples is taken into account.

\[
\text{Train} = \{(x_i,y_i)|i=1,2,\cdots,n\}, \quad \text{where the input data is } x_i \in \mathbb{R}^d, \text{ and the output characteristic is } y_i \in \mathbb{R}.
\]

Utilize a nonlinear mapping \( \phi(\cdot) \) to project the sample input space to the feature space \( \phi(x) = \{\phi(x_1), \cdots, \phi(x_n)\} \), and construct the following optimal linear decision function [11]

\[
y(x) = \omega^T \cdot \phi(x) + \beta
\]

By using the structural risk formula

\[
R = \frac{1}{2} \|\omega\|^2 + C \cdot R_{\text{emp}}
\]

one can calculate the weight vector \( \omega \) and the offset \( \beta \) in Eq (1). In Eq (2), \( C \) and \( R_{\text{emp}} \) denote the penalty factor and the loss function, respectively. In LSSVM, the loss function is always equal to the quadratic loss function \( R_{\text{emp}} = \sum_i \xi_i^2 \), where \( \xi_i \) denotes the degree that the misclassification sample deviates from the ideal sample. Then, getting the solutions of \( \omega \) and \( \beta \) is equivalent to solving the following optimization problem

\[
\text{min} R = \frac{1}{2} \|\omega\|^2 + C \cdot \sum_{i=1}^{n} \xi_i^2 \quad \text{s.t. } y_i = \omega^T \cdot \phi(x_i) + \beta + \xi_i, \quad i = 1, \cdots, n
\]

Introduce the Lagrangian multiplier \( \alpha = \{\alpha_1,\alpha_2,\cdots,\alpha_n\} \) to formula (3) and construct an equation as

\[
L(\omega, \beta, \xi, \alpha) = \frac{1}{2} \|\omega\|^2 + C \cdot \sum_{i=1}^{n} \xi_i^2 - \sum_{i=1}^{n} (\alpha_i \cdot (\omega^T \cdot \phi(x_i) + \beta + \xi_i - y_i))
\]

Calculate the derivative of every factor in the formula (4). For the purpose of eliminating \( \omega \) and \( \xi \), introduce the equations \( \omega = \sum_{i=1}^{n} \alpha_i \phi(x_i) \) and \( 2C \cdot \xi_i = \alpha_i \) to (4). Then one can obtain a new formula described as

\[
y_i = \sum_{j=1}^{n} \left( \alpha_j \cdot \phi(x_j)^T \cdot \phi(x_i) \right) + \beta + \frac{\alpha_i}{2C}
\]
Suppose that there exists a kernel function $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ which satisfies the Mercer condition [12]. Then the formula (5) can be described as

$$y_i = \sum_{j=1}^{n} \left( x_j \cdot K(x_i, x_j) \right) + \beta + \frac{x_i}{2C} \quad (6)$$

Solve the formula (6) to get the model parameters $\beta$ and $[\alpha_1, \cdots, \alpha_n]^T$. Then the decision function of LSSVM is shown as (see also [13])

$$y(x) = \text{sgn} \left[ \sum_{i=1}^{n} (x_j \cdot K(x, x_i)) + \beta \right] \quad (7)$$

It is well-known that different types of kernel function $K(x_i, x_j)$ will lead to different performance of LSSVM, while the polynomial kernel function, the sigmoid kernel function, the linear kernel function and the radial basis function (RBF) are widely used now. The experiment in [14] illustrates that the RBF is more general than other functions mentioned above. The RBF is usually described as

$$K(x, x_i) = \exp \left( -\frac{|x - x_i|^2}{2\sigma^2} \right) \quad (8)$$

Thus we choose the formula (8) as the kernel function for constructing the LSSVM classifier of ship power equipments’ fault. However, the factor $\sigma$ in the kernel function and the penalty factor $C$ should be optimized. The improved Cuckoo Search (CS) algorithm is proposed in this paper to optimize the system performance.

3. The improved CS algorithm

3.1 Description of the standard CS algorithm

The main mechanism of the CS algorithm is based on the strategy of cuckoo’s population with parasitical multiplying. This algorithm is to find out the optimal egg through its Lévy flight, which approach can achieve a high efficiency optimization [15] with one egg representing one optimal solution. The nest host will abandon its own nest once an outside egg is discovered, and then rebuild the nest in another place, where the new nest represents the new optimal solution. Therefore, the main idea of the CS algorithm is to adopt new and good solutions to substitute bad solutions [16]. There are two methods to signify the process of producing offspring.

3.1.1 The Lévy flight. If the Lévy flight [15] is introduced, the path of seeking nest and the expression of location updating can be described as

$$x_i^{k+1} = x_i^k + \alpha \cdot L(\lambda), \quad i = 1, 2, \cdots, n \quad (9)$$

where $x_i^k$ represents the location of the $i$th nest in the $k$th generation, $\circ$ denotes Hadamard product, $\alpha$ is the degree of step control, $L(\lambda)$ is a random searching path which satisfies Lévy flight distribution, while $s$ is the random step length satisfying

$$L(s, \lambda) \sim s^{-\lambda}, \quad (1 < \lambda \leq 3) \quad (10)$$

Adopt a formula about the path of Lévy flight which is proposed by Mantegna to describe this distribution (see also [17])

$$L(\lambda) = \mu/|\nu|^{1/\beta}, \quad 0 < \beta \leq 2 \quad (11)$$

where $\beta = \lambda - 1$ with $\beta$ being usually chosen as $\beta = 1.5$. The parameters $\mu$, $\nu$ satisfy a normal distribution.
3.1.2 Deviation of random walking. Under some cases, the nest host will probably discover the eggs of cuckoos. Define the probability of the nest host discovering the cuckoo egg as \( P_a \). Then one has

\[
x_{i+1} = x_i + r \cdot (x_j - x_i) \otimes H(P_a - r)
\]

(12)

where \( x, x_j, x_t \) are randomly chosen different solutions, \( H(\cdot) \) is a Heaviside function and \( r \) is a random number which satisfies the (0,1) uniform distribution.

3.2 The improved CS algorithm

3.2.1 The improvement of the recognition probability. When the nest location is updated, if \( r > P_a \), the nest location \( x_{i+1} \) will be updated, otherwise, it will keep unchanged. Then, the smaller the \( P_a \), the larger the probability of discovering outside eggs, and the number of updated nests will increase. One should choose an appropriate \( P_a \) to achieve a balance between the influences of the global extremum and the local extremum.

As observed from [18], if the probability falls into the interval \([0.1, 0.75]\), the global search capability will increase with the increase of the number of iterations. Introduce a dynamic adaptive mechanism to improve the probability within an appropriate range

\[
P_{a,i} = P_\text{amin} + (P_\text{amax} - P_\text{amin}) \left(1 - \frac{f_{\text{best}}^k}{f_i^k}\right)
\]

(13)

where \( P_{a,i} \) represents the probability of the \( i \)th nest host finding outside eggs in the \( k \)th generation. Similarly, \( P_\text{amax} \) and \( P_\text{amin} \) represent the bounds of maximum and minimum probabilities of recognition. At the \( k \)th generation population, \( f_{\text{best}}^k \) represents the best individual and \( f_i^k \) represents the \( i \)th individual fitness.

Note that if \( f_i^k = f_{\text{best}}^k \), the computation process will terminate. Thus, the value of \( P_a \) in formula (13) will increase with nests fitness values getting close to each other. Since the adaptive strategy of the algorithm is fully realized, the CS algorithm can achieve a balance between the global random searching ability and the local searching ability.

3.2.2 The improvement of adaptive step. In the Standard CS algorithm, the Lévy flight is adopted to generate a randomly variable step length. For the searching process, a large step length will lead to the increase of the probability of getting the global optimum. However, the searching precision will decrease and the phenomenon of oscillation is unavoidable. If a small step length is adopted, the searching precision is increased at the cost of reduced searching speed.

To reduce the influence of the Lévy flight randomly determined searching step length, and deal with the relationship between the ability of global optimization and the accuracy of optimization, we introduce the following equation (see also [19])

\[
d_i = \frac{\|n_i - n_{\text{best}}\|}{d_{\text{max}}}
\]

(14)

Based on formula (14), one can propose an adaptive strategy to renew the step length which is described as

\[
\text{step}_i = \text{step}_{\text{min}} + (\text{step}_{\text{max}} - \text{step}_{\text{min}})d_i
\]

(15)

In the above given formula, \( n_i \) represents the \( i \)th position of the nest, \( n_{\text{best}} \) represents the best nest corresponding to the best position, \( d_{\text{max}} \) represents a maximum value of the distance
between the optimal nest location and the others, while \( \text{step}_{\text{max}} \) and \( \text{step}_{\text{min}} \) represent the maximum and the minimum searching step lengths.

The smaller the distance between the \( i \)th nest position and the best nest position, the smaller the step length, otherwise, the larger the mentioned distance, the larger the step length. Then, the step length of the current iteration can be updated dynamically via the result of the last iteration, which demonstrates a good adaptivity. Then one can choose \( \alpha = \text{step}_{\text{a}} \) and adopt formula (9) to calculate the path of seeking nest and the nest location update.

3.3 Performance analysis of improved CS algorithm

To verify the effectiveness of the improved CS algorithm, we compare the performance between the improved CS (ICS) algorithm and the existing particle swarm optimization (PSO), genetic algorithm (GA), and standard CS algorithm. The parameter optimization results of the PSO method, the GA method, the standard CS method, and the ICS method are presented in Fig 1(A), 1(B), 1(C) and 1(D), respectively.

![Fig 1. Optimization parameters of the PSO method, the GA method, the CS method and the improved CS method.](image-url)

(a) Curve of fitness (PSO method) with \( c_1 = 1.5, c_2 = 1.7 \), and number of the populations = 20. (b) Curve of fitness (GA method) with number of the populations = 20. (c) Curve of fitness (CS method) with number of the nests = 20. (d) Curve of fitness (ICS method) with number of the nests = 20.

doi:10.1371/journal.pone.0171246.g001
As one can see from Fig 1, the improved CS method provides better fitness and higher optimization speed than other algorithms. Moreover, the average fitness of ICS is relatively stable. Choosing the mean square error of test sample’s predictive value as fitness, one can obtain the convergence speeds of the standard CS method and the improved CS method, see Fig 2. As one can see from Fig 2, the improved CS method provides a higher speed of optimization convergence than the standard CS method, which illustrates the merits of the newly proposed improved algorithm.

4. The improved CS-LSSVM algorithm-based fault pattern recognition

4.1 The process of fault pattern recognition

Step 1: In order to reduce the dimensionality, the equipment vibration signal data is collected and extracted from the kernel principal component by using the KPCA method for reducing the correlation between the characteristics of the data. The LSSVM training samples and testing samples are generated later.

Step 2: Initialize parameters which include the number of nests $n$, searching accuracy, the maximum number of iterations, the maximum and the minimum step length $\text{step}_{\text{max}}$ and $\text{step}_{\text{min}}$ of the CS algorithm. Generate initial locations of $n$ nests randomly, which are defined as $p_0 = [x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}]^T$ here. Assume that the location of a nest can be defined as $(C, \sigma)$. Then we can find out the best nest location $x_{\text{best}}^{(t)}$ and the best fitness $f_{\text{best}}^{(t)}$, where $t = 0$.

Step 3: Reserve the location of last generation’s nest as $x_{\text{best}}^{(t-1)}$. Calculate the step length of Lévy flight by using formula (14) and (15). Update the locations of $n$ nests. For $n$ nests, one can get the new locations from formula (12) and calculate the fitness $f_n^{(t)}$ by using LSSVM.

Step 4: Compare the newest locations with the last nests locations $p_{t-1} = [x_1^{(t-1)}, x_2^{(t-1)}, \ldots, x_n^{(t-1)}]^T$. Update them for getting better locations $p_t = [x_1^{(t)}, x_2^{(t)}, \ldots, x_n^{(t)}]^T$.

Step 5: Determine to reserve or change the locations of nests according to the adaptive recognition probability of formula (13). Generate a random number $r$, which is uniformly distributed, as the probability of nest host recognizing the outside eggs, and compare $r$ with $P_a$. Reserve the nest which is unidentified, in the meantime, change the nest location $x_n^{(t+1)}$.
which can recognize the outside eggs according to formula (12). Get new locations of the nests \( p_{t+1} = [x_1^{(t+1)}, x_2^{(t+1)}, \ldots, x_n^{(t+1)}]^T \).

**Step 6:** Compare \( p_{t+1} \) in **Step 5** with \( p_t \) in **Step 4**. Substitute the location in \( p_t \) with the location which provides better fitness in \( p_{t+1} \). Judge whether the best fitness \( f_{best}^{(k)} \) meets the searching accuracy or not. If this condition is satisfied, go to **Step 7**, otherwise, go to **Step 3**.

**Step 7:** Adopt the best nest location \( x_{best}^{(k)} \) and the corresponding value \((C, \sigma)\) as the optimal parameters for LSSVM to recognize the fault pattern. Judge whether the recognition accuracy
meets the requirement or not. If this condition is satisfied, output the result, otherwise, go to Step 3.

The process of the improved CS-LSSVM-based fault pattern recognition is presented in Fig 3.

4.2 The case analysis of fault pattern recognition

The analysis of fault pattern recognition is based on a practical MTU diesel engine. The feasible fault patterns include that fuel delivery valve is worn (FW), fuel feeding is abnormal (FA), fuel supply advance is late (FL), fuel supply advance is early (FE), needle valve is stuck (NS), needle valve is worn (NW), and so on. The definition of different fault patterns is shown in Table 1.

The diesel engine vibration signal is collected under the sampling frequency 6000Hz. Then the KPCA method is adopted to reduce the dimensionality of itself for getting 8 characteristic values including peak-to-peak value $x_{p-p}$, mean of absolute value $x_{av}$, root-mean-square $x_{rm}$, kurtosis $\beta$, wave index $s_f$, peak index $C_f$, impulse index $I_f$, and margin index $C_{L_f}$ [20]. There are 7 different vibration signals. Every vibration signal includes 8 characteristic values. Every characteristic value includes 10 samples. Then the total number of sample data is 560. Choose 5 samples among them as training samples with the total number being equal to 35. The other 35 samples are test samples. Use the improved CS-LSSVM algorithm to deal with these samples. Because of the randomness of the meta-heuristic algorithm, we take 200 iterations here. By using the improved CS-LSSVM, we can get $C$ and $\sigma$. The corresponding recognition result is shown in Fig 4.

![Fig 4. The improved CS-LSSVM-based recognition result.](https://example.com/fig4)

doi:10.1371/journal.pone.0171246.g004
As one can see in Fig 4, the recognition rate is close to 100% under the improved CS-LSSVM algorithm, which illustrates the effectiveness of the newly proposed algorithm. Compared with other intelligent algorithms, the proposed algorithm in this paper can improve the recognition accuracy greatly, refer to Table 2 for more details.

5. Conclusions

For LSSVM, precise selection of the parameter in RBF kernel function and the penalty factor is the key to recognize the fault pattern. Since the artificial parameter setting may lead to inaccurate recognition, an improved CS algorithm has been proposed in this paper to optimize parameters. The simulation results have illustrated that the improved CS algorithm can converge to the global optimum quickly, and lead to perfect fault pattern recognition performance of ship power equipments. The improved CS-LSSVM can provide a technical support for maintaining ship power equipments timely and accurately.

Acknowledgments

This work was supported in part by the Open Foundation for Joint Laboratory of Ocean-based Flight Vehicle Measurement and Control under Grant FOM2016OF001. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceptualization: YFY.
Data curation: YFY MJT.
Formal analysis: YFY.
Funding acquisition: YFY.
Investigation: YFY YWD.
Methodology: YFY MJT.
Project administration: YFY.
Resources: YFY.
Software: MJT.
Supervision: YFY MJT.
Validation: YYF TMJ.
Visualization: YFY MJT.
Writing – original draft: YFY MJT.

| Algorithm     | Number of Samples | C     | σ     | Precision |
|---------------|-------------------|-------|-------|-----------|
| PSO-LSSVM     | 35                | 6.705 | 18.675| 85.71%    |
| GA-LSSVM      | 35                | 13.067| 9.7103| 91.43%    |
| CS-LSSVM      | 35                | 17.717| 6.5610| 94.29%    |
| ICS-LSSVM     | 35                | 12.373| 8.6546| 99.96%    |

doi:10.1371/journal.pone.0171246.t002
Writing – review & editing: YFY.

References
1. Aminian F, Aminian M, Collins HW. Analog fault diagnosis of actual circuits using neural networks. IEEE Transaction Instrumentation and Measurement. 2002; 51(3): 544–550.
2. Shen MH, Xiao L, Wang FX. Application of support vector machine (SVM) in pattern recognition. Telecommunication Engineering. 2006; 04: 9–12. (In Chinese)
3. Ji RR, Cao LJ, Wang Y. Joint depth and semantic inference from a single image via elastic conditional random field. Pattern Recognition. 2016; 59: 268–281.
4. Ji RR, Gao Y, Liu W, Xie X, Tian Q, Li XL. When location meets social multimedia: A survey on vision-based recognition and mining for geo-social multimedia analytics. ACM Transactions on Intelligent Systems & Technology. 2015; 6(1): 1–18.
5. Ji RR, Duan LY, Chen J, Huang TJ, Gao W. Mining compact bag-of-patterns for low bit rate mobile visual search. IEEE Transactions on Image Processing. 2014; 23(7): 3099–3113. doi: 10.1109/TIP.2014.2324291 PMID: 24835227
6. Zhang Y, Guan T, Duan LY, Wei BC, Gao J, Mao T. Inertial sensors supported visual descriptors encoding and geometric verification for mobile visual location recognition applications. Signal Processing. 2015; 112: 17–26.
7. Guan T, Wang YS, Duan LY, Ji RR. On-device mobile landmark recognition using binarized descriptor with multi-feature fusion. ACM Transactions on Intelligent Systems & Technology. 2015; 7(1): 1–29.
8. Pan HL, Guan T, Luo YW, Duan LY, Tian Y, Yi L, et al. Dense 3D reconstruction combining depth and RGB information. Neurocomputing. 2016; 175: 644–651.
9. Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995; 20(3): 273–297.
10. Cao SG, Liu YB, Wang YP. A forecasting and forewarning model for methane hazard in working face of coal mine based on LS-SVM. Journal of China University of Mining and Technology. 2008; 18(2): 172–176.
11. Gencoglu MT, Uyar M. Prediction of flashover voltage of insulators using least squares support vector machines. Expert Systems with Applications. 2009; 36(7): 10789–10796.
12. Gu YP, Zhao WJ, Wu ZS. Least squares support vector machine algorithm. Journal of Tsinghua University (Science and Technology). 2010; 07: 1063–1066+1071. (In Chinese)
13. Van Gestel T, Suykens JAK, Baesens B, Vlaene S, Vanthienen J, Dedene G, et al. Benchmarking least squares support vector machine classifiers. Machine Learning. 2004; 54(1): 5–32.
14. Hsu CW, Lin CJ. A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Networks. 2002; 13(2): 415–425. doi: 10.1109/72.991427 PMID: 18244442
15. Yang XS, Deb S. Cuckoo search via levy flights. Nature & Biologically Inspired Computing (NaBiC). World Congress on IEEE. 2009; 210–214.
16. Liu XN, Ma M. Application of cuckoo search algorithm in multi-threshold image segmentation. Computer Engineering. 2013; 39(7): 274–278. (In Chinese)
17. Mantegna RN. Fast, accurate algorithm for numerical simulation of Lévy stable stochastic processes. Physical Review E. 1994; 49(5): 4677–4683.
18. He M, Ruan Q, Zheng XG, Huang YL, Zhang L. Adaptive cuckoo search algorithm. Computers and Applied Chemistry. 2014; 31(8): 961–968. (In Chinese) Available from: http://d.wanfangdata.com.cn/Periodical/jsjgcyy201310019
19. Zheng HQ, Zhou YQ. Self-adaptive step cuckoo search algorithm. Computer Engineering and Applications. 2013; 49 (10): 68–71. (In Chinese) Available from: http://d.wanfangdata.com.cn/Periodical/jsjgcyy201310019
20. Li HB. Study on ensemble learning diagnosis method for diesel engine fault. M.Sc. Thesis, Xi’an Shiyou University. 2015. (In Chinese) Available from: http://d.wanfangdata.com.cn/Thesis/Y2796558