IOP Conference Series: Earth and Environmental Science

PAPER • **OPEN ACCESS**

Oil immersed transformer fault diagnosis based on cross entropy algorithm optimized support vector machine

To cite this article: L Bian and H He 2018 *IOP Conf. Ser.: Earth Environ. Sci.* **188** 012053

IOP Conf. Series: Earth and Environmental Science **188** (2018) 012053 doi:10.1088/1755-1315/188/1/012053

Oil immersed transformer fault diagnosis based on cross entropy algorithm optimized support vector machine

L Bian[1,2] and H He[1]

[1]Heilongjiang Institute Of Science and Technology, China

E-mail: branran@163.com

Abstract. With the increase of the voltage level of the power grid and the increase of its capacity, the probability of transformer failure is getting higher and higher. In order to discover the early latency faults of transformers, a cross-entropy algorithm was proposed to optimize the support vector machines. This method established a support vector machine classification model and used the cross-entropy algorithm to optimize the penalty factor and kernel function parameters. The transformer fault data is used to verify the classification model and compared with the test results of other algorithms. The results show that the accuracy of this algorithm for oil-immersed transformer fault diagnosis has reached 86.7%, which is higher than that of genetic algorithm and particle swarm algorithm, and iterating over 6 times of the fitness curve tends to be smooth and takes less time. After many tests, the forecast results are stable.

Nomenclature

C : Penalty factor
g : Parameters of the Kernel Function in Support Vector Machine
x^* : Optimal solution to the problem
ξ : Given real number
ξ^* :The maximum value of the function S
I : Indicator function
f : Probability density function
l : The probability of $S(X)$ larger than ξ
X :Random samples produced by f
E_v :Expected value
t : Number of iterations
M :Random sample number
β : Smoothness coefficient
ρ : Quantile
n : Dimension
$S_{(t)}$: Sequence of fitness function values
$\bar{S}$:The matrix after the adjustment of the fitness function value
μ : Mean value
σ^2 : Variance

List of abbreviations
CE: Cross entropy algorithm
GA: Genetic algorithm
PSO: Particle swarm optimization
SVM: Support Vector Machines

1. Introduction

The power transformer is one of the most expensive and most important equipment in the power system transmission, transformation and distribution equipment. Its normal operation can ensure the safe, reliable and stable operation of the power system. The analytical method of dissolved gas in oil is one of the most convenient and effective methods for fault diagnosis of oil-immersed transformers. On this basis, traditional methods such as three-ratio method and Rogers method as well as artificial neural network, Bayesian network, support vector machine and other artificial intelligence methods have been formed.

The traditional algorithm has many shortcomings such as lack of coding, too strict coding boundaries, etc. Artificial neural networks have disadvantages such as overfitting and falling into local optimum, which limits their accuracy. The Bayesian network requires a large amount of sample data. Among many algorithms, the support vector machine classifier has been successfully applied to fault diagnosis of oil-immersed transformers. However, there is no theoretical basis or effective methods for the selection of parameters that play a key role in its performance. The current application of a wide range of intelligent optimization algorithms includes genetic algorithms, particle swarm optimization and so on. Genetic algorithm has a global search, so it has better fault tolerance. However, this method contains selection, crossover, and mutation links, and it is easy to generate redundant iterations, resulting in a large amount of calculation and low solution efficiency [1]. Particle swarm optimization algorithm is simpler and easier than genetic algorithm, without crossover and mutation links, and the convergence speed is fast. However, in the face of complex problems, the algorithm is easy to fall into a local optimum, and there is a phenomenon of premature convergence.

Cross-entropy algorithm is a global stochastic optimization algorithm. It uses parameterized probability density distribution to generate random samples, so that the candidate samples used in each iteration change, so the optimization process is not easy to fall into the local optimal solution [2,3]. At the same time, the algorithm makes full use of the feedback information has the advantages of fast convergence, accurate diagnosis, and good stability.

As oil-immersed transformers have many types of faults, their properties are complex. Therefore, timely detection of early failure of the transformer, and its diagnosis and maintenance, can not only ensure the normal operation of the system, but also extend the life of the transformer. If the transformer early fault can not be found in time, the fault will deteriorate further, not only the transformer itself will be destroyed, but the normal and stable operation of the whole system will also be affected, and it will also endanger the safety of the person, cause irreparable economic loss and extremely bad social influence. This requires a more rapid and accurate diagnosis. Fastness and accuracy are the most important indicators of the diagnosis process. Therefore, the author uses the cross entropy algorithm to optimize the parameters of the SVM to diagnose faults, and compare it with genetic algorithm and particle swarm optimization.

2. Introduction to support vector machines

Support vector machine is based on the principle of VC theory and structural risk minimization. It is a classic machine learning algorithm for sample classification, especially on small data [4]. Rimjhim Agrawal *et al* applied support vector machines to the fault diagnosis of distribution networks [5]. Ravikumar *et al* applied support vector machines to the fault diagnosis of transmission systems [6]. Wang used support vector machines for the fault diagnosis of CNC machine tool spindles [7]. This shows that the support vector machine has good applicability in the field of classification.

Selecting different kernel functions can constitute different SVM classifiers. Common kernel functions include polynomial kernel functions, Sigmoid kernel functions, and radial basis functions (RBF). The SVM classifier constructed by RBF has very good nonlinear classification performance. Moreover, the kernel function has only one parameter and the parameter optimization is simple. Therefore, RBF is used in this paper. The RBF kernel function is as follows:

$$K(x,y)=\exp(-\frac{\|x-y\|^2}{2g^2}) \tag{1}$$

In formula (1) [8], g is the radial base width.

Since the kernel function uses RBF, the SVM classifier parameters C and g are artificially selected parameters, and the selection of C and g is an important factor influencing the performance of the SVM classifier.

3. The basic principle of cross-entropy algorithm

Each iteration of the cross entropy algorithm can be split into two steps [9]: 1) Based on a certain probability density function, a set of random sample data is generated; 2) Update the parameters of the probability density function according to these data samples, so as to contribute more optimized samples for the next iteration.

For optimization issues:

$$S(x^*) = \xi^* = \max_{x\in\chi} S(x) \tag{2}$$

In equation (2) [10], S is a real value function of X , and the maximum value of the function S is found. The CE algorithm converts the above optimization problem into a probabilistic problem, that is, it is converted into a probability problem that the S(x) is larger than the known parameter γ. This problem can be expressed by equation (3), where the random vector $X = (x_1, x_2......x_n)$.

$$l(\xi) = P_v(S(X) \ge \xi) = \sum_{x\in\chi} I_{\{S(X)\ge\xi\}} f(x;v) = E_v I_{\{S(X)\ge\xi\}} \tag{3}$$

When ξ is close to ξ^*, the value of l will be smaller and smaller, so in order to make sense, the value of l cannot be too small, and the choice of ξ and v is crucial. In order to solve this problem, a multi-level algorithm $\{v_t, t > 0\}$ is used to construct the distribution parameter sequence and the level sequence $\{\xi_t, t > 0\}$ as the number of iterations. Then, iteratively updates v_t and ξ_t until the maximum value of the corresponding element change in the distributed parameter sequence is less than a specified parameter b_{tol} after an iteration, and the iteration ends.

4. Cross entropy optimization SVM algorithm steps

Continuous cross-entropy algorithm is adopted, and SVM (c, g) is used as the optimization goal. The cross-validation probability of SVM is the fitness function. The specific optimization steps are as follows [11]:

Step (1): The initial values ($\mu_{(0)}$) of the parameters c and g (the dimension of $\mu_{(0)}$ is n), $\sigma^2{}_{(0)}$, the random sample number M, the smoothing coefficient β, and the quantile $\rho_{(0)}$ are respectively assigned to the initial values, so that the number of iterations is t=0.

Step (2): t=t+1, with $N(\mu_{(t-1)}, \sigma^2_{(t-1)})$ distribution, generates M candidate sample matrices $X^T = [X_1,......, X_M]^T$, where each X_M is an n-dimensional vector, $X_m = (x_{m1},......, x_{mn})$.

Step (3): Sort the sequence $S_{(t)} = [S_{1(t)},..., S_{K(t)}]^T$ of the fitness function values from the smallest to the largest, obtain a new matrix $\bar{S}_{(t)}$, and then use the formula (4) to calculate the $(1-\rho_0)$ quantile of the $\tilde{S}_{(t)}$ sequence.

$$\gamma_t = \tilde{S}_{((1-\rho_0)M)} \tag{4}$$

Step (4): Using the generated M random sample substitutions, update parameters $\mu_L = (\mu_1,..., \mu_n)$

and $\sigma_L^2=(\sigma_1^2,...,\sigma_n^2)$, where L=1, 2,..., n.

$$\mu_L=\frac{\sum_{m=1}^{M} I_{\{S(X_m)\le\gamma_t\}} x_{mL}}{\sum_{m=1}^{M} I_{\{S(X_m)\le\gamma_t\}}} \tag{5}$$

$$\sigma_L^2=\frac{\sum_{m=1}^{M} I_{\{S(X_m)\le\gamma_t\}} (x_{mL}-\mu_L)^2}{\sum_{m=1}^{M} I_{\{S(X_m)\le\gamma_t\}}} \tag{6}$$

Step (5): Calculate for m=1,...,M

$$\mu_{L(t)}=\beta\mu_{L(t)}+(1-\beta)\mu_{L(t-1)} \tag{7}$$

$$\sigma_{L(t)}^2=\beta\sigma_{L(t)}+(1-\beta)\sigma_{L(t-1)} \tag{8}$$

Where 0.5≤β≤0.9, and $\mu_{L(t)},\sigma_{L(t)}^2$ are the Lth elements in the sequence after the tth iteration.

Step (6): If the termination condition $\max(\sigma_{(t)}^2)<\theta(\theta=|\ \sigma_{L(t-1)}^2-\sigma_{L(t)}^2|)$ is satisfied after the tth iteration in the iterative process, the iteration is terminated; otherwise, the process returns to step (1) to execute again.

Finally, the radial base width $g=\sigma_{L(t)}^2$ of the optimal solution $X^*=\mu_{(t)}$ and the penalty factor $c=\mu_{L(t)}$ is obtained.

5. Cross entropy optimization support vector machine for transformer fault diagnosis

5.1. *Establishment process of fault diagnosis model based on CE-SVM*

The flowchart of the diagnostic model of Oil-immersed transformer based on CE-SVM is shown in Figure 1. The specific steps are as follows:

- Analyze and summarize oil-immersed transformer fault data, extract feature attributes and decision attributes, and normalize them.
- Continuous cross-entropy algorithm is used to optimize SVM (C, g) parameters to obtain the best parameters of SVM classification model.
- Acquisition of training samples. In this paper, 43 sets of sample data are selected from 58 sets of sample data as training samples for training based on CE-SVM fault diagnosis model.
- Test sample selection and evaluation of SVM model. The remaining 15 sets of sample data were used as test samples, and the test sample was used to output and evaluate the established SVM classification model, and the accuracy of the classification model for failure classification was obtained.

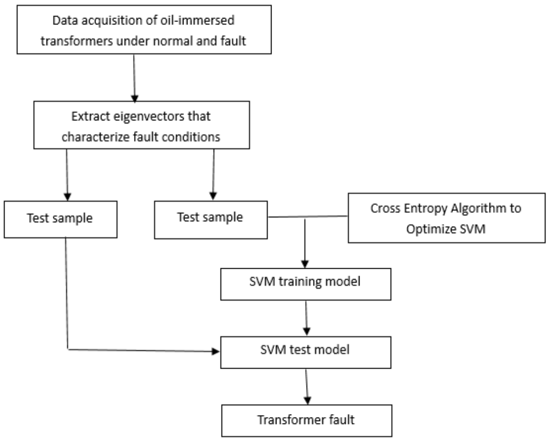

Figure 1. Flowchart of fault diagnosis based on cross-entropy algorithm to optimize support vector machine.

5.2. Preliminary processing of data

This article obtains 58 sets of data through the literature [12], Divide oil-immersed transformer fault data into two groups, one of which contains 43 sample data and use it as a training sample to build a test model; The other group contains 15 sample data, which is used as a test sample to test the model.

Before building the model, the raw data needs to be normalized. Data that has not been normalized will have some influence on the recognition and convergence of SVM. In order to minimize the training time of the predictive model, prevent the numerical change, and make the parameter easier in the calculation process, generally the sample set is uniformly normalized before the predictive model is constructed. Table 1 shows a sample of transformer faults. Table 2 shows the sample to be normalized.

Table 1. Partial fault samples.

Sample number	1	2	3	4	5	6
C3H8	46.13	22.43	1.88	3.56	10.07	47.71
C3H6	18.16	6.00	43.02	68.00	44.75	9.86
C3H4	7.00	67.83	52.37	28.73	67.49	71.34
CO2	5828.09	11827.09	9162.90	10796.01	6242.65	8204.98
C2H4	173.46	164.16	97.66	164.70	47.77	171.42
H2	1182.10	989.96	228.43	79.19	56.28	565.68
CO	732.99	214.49	698.84	537.46	273.68	445.76
C2H2	216.04	15.13	95.73	57.53	57.72	65.57
C2H6	120.77	159.65	52.71	21.04	156.80	152.82
CH4	633.98	154.26	945.30	158.23	191.24	315.30
C2H6/CH4	0.19	1.03	0.06	0.13	0.82	0.48
C2H4/C2H6	1.44	1.03	1.85	7.83	0.30	2.33
C2H2/CH4	0.34	0.10	0.10	0.36	0.30	0.21
C2H2/C2H4	1.25	0.09	0.98	0.35	1.21	0.38
CH4/H2	0.54	0.16	4.14	2.00	0.34	0.56
C2H6/C2H2	0.56	10.55	0.55	0.37	2.72	2.33

Table 2. Partial sample normalization.

Sample number	1	2	3	4	5	6
C3H8	0.9333	0.4335	0	0.0354	0.1727	0.9667
C3H6	0.2401	0.0642	0.5996	0.9608	0.6246	0.1200
C3H4	0.0511	0.8352	0.6359	0.3312	0.8309	0.8805
CO2	0.4429	1	0.7526	0.9043	0.4814	0.6637
C2H4	0.6133	0.5796	0.3382	0.5815	0.1571	0.6059
H2	0.9975	0.8320	0.1764	0.0480	0.0282	0.4668
CO	0.7594	0	0.7094	0.4730	0.0867	0.3387
C2H2	1	0.0651	0.4401	0.2624	0.2633	0.2998
C2H6	0.6021	0.8004	0.2548	0.0933	0.7859	0.7656
CH4	0.6679	0.1601	0.9975	0.1643	0.1993	0.3306
C2H6/CH4	0.0210	0.1192	0.0058	0.0140	0.0946	0.0546
C2H4/C2H6	0.0776	0.0546	0.1007	0.4370	0.0135	0.1277
C2H2/CH4	0.0231	0.0068	0.0068	0.0244	0.0204	0.0142
C2H2/C2H4	0.0913	0.0066	0.0716	0.0256	0.0884	0.0278
CH4/H2	0.0596	0.0152	0.4801	0.2301	0.0362	0.0619
C2H6/C2H2	0.0034	0.0680	0.0034	0.0022	0.0174	0.0149

5.3. *Transformer fault diagnosis analysis*

Figure 2 shows the fitness curve. It can be seen from the figure that when the evolution algebra is 6, the curve has stabilized and reaches an ideal effect. This shows that this algorithm has the advantage of fast convergence. Figure 3 compares the test sample with the prediction result. 15 test samples can be found in the figure. Only the second sample and the fifteenth sample are misclassified. The actual test set classification and the test set classification mark do not overlap. After many tests, the result is stable, and the accuracy rate is about 86.7%. This shows that the cross entropy algorithm has good reliability for the optimization of the support vector machine.

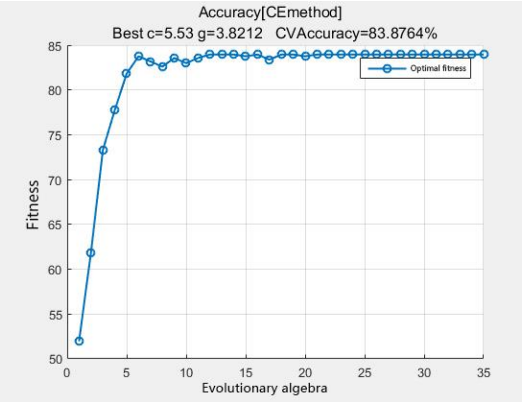

Figure 2. Fitness curve.

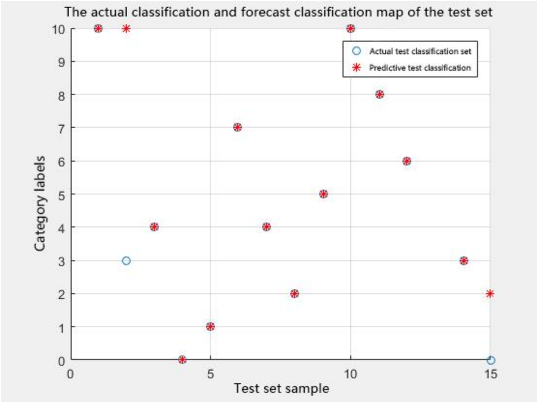

Figure 3. Cross-entropy algorithm optimization forecast results.

Figure 4 shows the results of genetic algorithm optimization support vector machine (SVM) optimization. Although the error classification is mostly maintained at 2, there are occasionally three error cases. Figure 5 shows that the particle swarm optimization algorithm optimizes the support vector machine (SVM) prediction results. The number of error classifications is maintained at about 2-3, and 4 errors occur with a small probability.

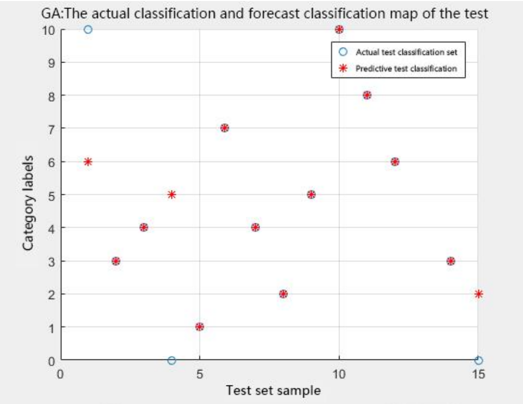

Figure 4. Genetic algorithm optimization forecasting result.

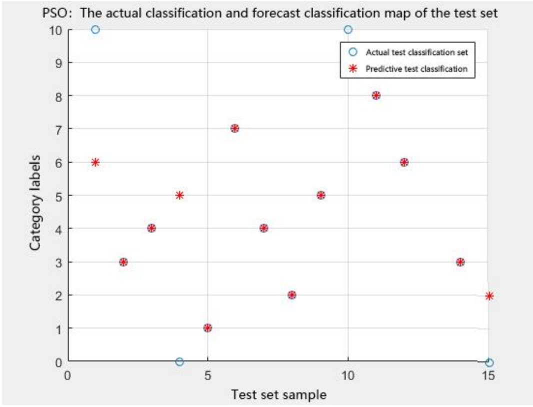

Figure 5. Particle Swarm Optimization algorithm forecasting result.

Table 3. Performance comparison of three intelligent algorithms.

	CE	PSO	GA
Accuracy	86.7%	79.4%	82.9%
Training time (s)	10.7086	17.2150	21.2253

The cross-entropy algorithm, genetic algorithm and particle swarm optimization algorithm's performance after optimizing the parameters of the SVM is shown in table 3. The cross-entropy algorithm takes less time than the other two algorithms, and the accuracy rate is higher than the other two intelligent algorithms. It can be analyzed that cross-entropy algorithm optimized support vector machine does have better fault diagnosis capability.

6. Conclusion

This paper applies cross-entropy algorithm and support vector machine to transformer fault diagnosis. Through modeling, simulation, and comparison with other intelligent algorithms, The following conclusions are obtained.

- The algorithm has higher accuracy for fault diagnosis and proves that the algorithm has the advantages of fast convergence speed and short running time.
- Genetic algorithm and particle swarm optimization have their own shortcomings. Cross entropy algorithm is more suitable for transformer fault diagnosis.
- This algorithm has a good effect on the optimization of support vector machines to solve transformer fault diagnosis. This method can be applied to other fault diagnosis and can solve problems more quickly and accurately.

Transformer fault diagnosis is a complex technology. It is almost impossible to classify the faults accurately and quickly only by relying on a single technology or a subject alone. Therefore, it is necessary to integrate various theories and intelligent methods in order to effectively and accurately diagnose faults, and will inevitably become the future development trend of the field of fault diagnosis.

Acknowledgments

foundation item：The National Natural Science Foundation of China (51504085);
Natural Science Foundation for Returness of Heilongjiang Province of China (LC2017026).
Thanks to Professor Bian Li for her guidance and Liu Wenjing's help to me.

References

[1] Bian L, Zhang X X and Li T T 2016 Method for predicting the permeability of coalbed methane

reservoir *J. Heilongjiang Univ. Sci. Technol.* **5** 480-4
[2] Reuven Y R 2004 *The Cross-entropy Method: A Unified Approach to Combinatorial Optimization* (Australia: Springer-Verlag New York Inc)
[3] Li Y H, Wu Y Q and Ao L 2010 Reliability evaluation technology based on cross entropy *Aeronaut. Standard. Qual.* **21** 37-9
[4] Deng N and Tian Y J 2004 *A New Method in Data Mining--Support Vector Machines* (Beijing: Science Press)
[5] Rimjhim A and Thukaram D 2012 Identification of fault location in distribution networks using multi class support vector machines *Int. J. Emerg. Elec. Power. Syst.* **13** 1-17
[6] Ravikumar B, Thukaram D and Khincha H P 2011 Intelligent approach for fault diagnosis in power transmission systems using support vector machines *Int. J. Emerg. Elec. Power. Syst.* **8** 1-32
[7] Wang Z X 2017 Research on fault diagnosis of CNC spindle based on support vector machine (Anhui, China: Hefei University of Technology)
[8] Chen C W, Du Y G and Zhou C 2018 Fish feeding behavior recognition technology based on support vector machine *Jiangsu Agricultural Sciences* **7** 226-9
[9] Pieter Tjerk De B and Reuven Y R 2015 A tutorial on the cross-entropy method *Ann. Oper. Res.* **134** 19-67
[10] Bian L and Che X Q 2016 *The Application of Cross Entropy Algorithm in the Field of Electronic Engineering* (Shaanxi, China: Xidian University Press)
[11] Bian L and Bian C Y 2016 Grid fault diagnosis using cross-entropy support vector machine and fuzzy integration *J. Elec. Mach. Contr.* **20** 112-20
[12] Wang G P 2016 Research on reliability evaluation and fault diagnosis model of oil-immersed power transformer (Guangdong, China: South China University of Technology)