Electric Power Systems Research 155 (2018) 196–205

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

A novel model based on wavelet LS-SVM integrated improved PSO algorithm for forecasting of dissolved gas contents in power transformers

Hanbo Zheng[a,b,1], Yiyi Zhang[a,f,*,1], Jiefeng Liu[a,c,*,1], Hua Wei[a], Junhui Zhao[d], Ruijin Liao[e]

[a] Guangxi Key Laboratory of Power System Optimization and Energy Technology, Guangxi University, Nanning, Guangxi 530004, China
[b] State Grid Henan Electric Power Research Institute, Zhengzhou, Henan 450052, China
[c] Shijiazhuang Power Supply Branch of State Grid Electric Power Company, Shijiazhuang 050093, China
[d] Department of Electrical and Computer Engineering & Computer Science, University of New Haven, West Haven, CT 06516, USA
[e] State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University, Chongqing 400044, China
[f] National Demonstration Center for Experimental Electrical Engineering Education, Guangxi University, Nanning, Guangxi 530004, China

ARTICLE INFO

Article history:
Received 11 July 2017
Received in revised form 7 September 2017
Accepted 10 October 2017

Keywords:
Wavelet technique
Least squares support vector machine
Forecasting
Dissolved gases
Oil-immersed power transformers
Particle swarm optimization

ABSTRACT

Finding out the transformer incipient faults and their development trend has always been a central issue for electric power companies. In this paper, a novel approach combing wavelet technique with least squares support vector machine (LS-SVM) for forecasting of dissolved gases in oil-immersed power transformers has been proposed. The algorithm of particle swarm optimization (PSO) with mutation is developed to optimize the parameters of constructed wavelet LS-SVM regression (W-LSSVR). The existence of admissible wavelet kernels is proven by theoretic analysis. Evaluation of forecasting performance is based upon the measures of mean absolute percentage error (MAPE) and squared correlation coefficient (r^2). On the basis of the proposed approach, a procedure is put forward to serve as an effective tool and experimental results show that this approach is capable of forecasting the dissolved gas contents accurately. Comparing with the back propagation neural network (BPNN), the radial basis function neural network (RBFNN), the generalized regression neural network (GRNN), and the SVM regression (SVR) in two practical cases (taken hydrogen as an example here), the MAPEs of the proposed approach are significantly better than that of the four methods (5.4238% vs 19.1458%, 11.7361%, 7.7395%, 8.3248%; 2.1567% vs 18.9453%, 10.2451%, 7.8636%, 2.4628%) respectively.

1. Introduction

As a major component of the power system, a malfunction of an oil-immersed power transformer is among the more frequent causes of interruptions in power supplies with serious repercussions on the system stability and reliability [1]. In order to promote the operation reliability of power transformers, electric power companies perform online monitoring to obtain the condition of oil-immersed power transformers.

Dissolved gases in oil always result from the decomposition of electrical insulation materials (oil or paper), as a result of faults or chemical reactions in the power transformers. Therefore, with the advantages of non-destructive monitoring and sensitively detecting, dissolved gas analysis (DGA) is considered to be a widely used online monitoring method for detecting the early incipient faults in oil-filled power transformers [2]. Various computational and graphical methods employing gas ratios and proportions of gases dissolved in oil include the key gas method, Doernenburg ratio method, Rogers ratio method, IEC ratio method, and Duval triangle method [3]. In the IEC 60599 three-ratio method, ratios of certain gases are used for diagnostic analysis. These techniques were standardized in 1978 by IEC and later revised in 2008 [4]. In the last two decades, artificial intelligent (AI) techniques have been proposed for transformer fault diagnosis, such as fuzzy logic inference system [5–17], artificial neural network (ANN) method [18–26], expert systems [27–29], grey clustering analysis method [30,31], and rough

* Corresponding authors at: Guangxi Key Laboratory of Power System Optimization and Energy Technology, Guangxi University, Nanning, Guangxi 530004, China.
E-mail addresses: hanbozheng@163.com (H. Zheng), yiyizhang@gxu.edu.cn (Y. Zhang), liujiefeng9999@163.com (J. Liu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.epsr.2017.10.010
0378-7796/

set theory method [32,33], among others. Additional AI techniques have been used by researchers, such as self-organizing polynomial networks [34], organizing-map algorithm [35], data mining approach [36], extension theory [37], Bayesian network [38], and kernel-based possibilistic theory [39], etc.

From the above brief review, the fault diagnosis approaches mentioned above can only detect the present-time happening or already happened faults in a transformer rather than forecast the faults. Due to the fact that the forecasting of the gas content in oil can effectively predict the faults of oil-immersed transformers, some data-centric machine-learning techniques have been introduced for the prediction of transformer failures from DGA data [40–47]. In the past, the main drawback of the traditional forecasting methods was that they were established on the principle of empirical risk minimization such as the back propagation neural network (BPNN) [48], radial basis function neural network (RBFNN) [49] and generalized regression neural network (GRNN) [50]. As discussed in Refs. [48–50], these approaches related to artificial neural network can fully approach to arbitrary complex nonlinear relationship and show the good performance in a forecasting problem. However, a large amount of historical data is needed for model training to overcome the over fitting problem, which may result in unacceptable performance in the application because of the limitation of the key gas content data in practice [51–53]. A support vector machine (SVM) relies on the structural risk minimization principle, in considering both empirical risk minimization and the complexity of the learning machine, and therefore it is good for solving small samples and optimal problems, and has good generalization ability [54–59], so it is an effective approach for solving the forecasting problems [56]. The least squares support vector machine (LS-SVM) is introduced by Ref. [60] as reformulations to the standard SVM [54,56] which simplifies the model of standard SVM in a great extent by applying linear least squares criteria to the loss function instead of traditional quadratic programming method. The simplicity and inherited advantages of SVM such as basing principle of structural risk minimization and kernel mapping promote the applications of LS-SVM in many pattern recognition and regression problems [61–64].

However, the common used kernels for SVM such as Gaussian and polynomial kernels are not orthonormal bases, whereas the wavelet function is orthonormal in $L_2(R^N)$ space [65–67]. That is to say the wavelet function can approximate arbitrary curves in $L_2(R^N)$ space. So it is not surprising that wavelet kernel gets better approximation than Gaussian kernel [68–71], which is shown by computer simulations. In addition, most of past studies do not consider optimizing parameters of the kernel function for SVM, leading to get unsatisfactory classification or prediction accuracies in practice. In order to overcome these limitations, the main contribution of this paper can be summarized as follows:

(1) Proving the existence of three admissible wavelet kernels and building the LS-SVM regression algorithm based on wavelet kernels.
(2) A mutation operation employing certain probability is applied to the algorithm of traditional particle swarm optimization (PSO) [72–76] to optimize parameters of the kernel function for LS-SVM, which can overcome the drawback of premature convergence of traditional PSO.
(3) Introducing the wavelet LS-SVM regression (W-LSSVR) to forecast dissolved gases in oil-immersed power transformers.

The novelty of this paper is to integrate the advantages of wavelet kernels, LSSVM, and PSO with mutation to forecast dissolved gases in oil-immersed power transformers. And the performance evaluation of proposed W-LSSVR is guided by the two measure criteria; namely, mean absolute percentage error (MAPE) and squared correlation coefficient (r^2). Research results and comparisons are performed to emphasize the potential of the provided method with satisfactory forecasting accuracy and valuable information.

2. Methodology

2.1. Establishment of wavelet least squares support vector machine

2.1.1. Conditions for support vector's kernel function

Generally, the formation of an SVM kernel is a kernel of dot-product type in some feature space $K(x, x') = \varphi(x)^T\varphi(x')$. According to Hibert–Schmidt theory, $K(x, x')$ can be any symmetric function satisfying the following Mercer's condition [77].

Theorem 1. *To guarantee that the symmetric function $K(x, x')$ from $L_2(R^N \times R^N)$ space has an expansion*

$$K(x, x') = \sum_{k=1}^{\infty} a_k \psi_k(x)\psi_k(x') \quad (1)$$

with positive coefficients $a_k > 0$, (i.e., $K(x, x')$ describes an inner product in some feature space), it is necessary and sufficient that the condition.

$$\iint K(x, x')g(x)g(x')dxdx' \geq 0 \quad (2)$$

be valid for all $g \neq 0$ for which

$$\int g^2(x)dx < \infty \quad (3)$$

For the translation invariant kernels, *i.e.*, $K(x, x') = K(x - x')$ are also admissive SVM kernels if they satisfy Mercer's condition. However, it is difficult to decompose the translation invariant kernels into the dot-product of two functions and then to prove them as SVM kernels according to Theorem 1. Consequently, we state a necessary and sufficient condition for translation invariant kernels as follows [78]:

Theorem 2. *A translation invariant kernel $K(x, x') = K(x - x')$ is an admissible SVM kernel if and only if the Fourier transform need satisfy the following condition:*

$$F[K](\omega) = (2\pi)^{-N/2} \int_{R^N} \exp(-j(\omega \cdot x))K(x)dx \geq 0 \quad (4)$$

2.1.2. Wavelet kernel functions as support vector's kernels

Suppose that a wavelet function $\psi(x)$ satisfies the admissibility condition

$$\int_0^{\infty} \frac{|\psi'(x)|^2}{\omega} d\omega < \infty \quad (5)$$

where $\psi'(x)$ is the Fourier transform of $\psi(x)$, then $\psi(x)$ is called the mother wavelet. So the wavelet function group can be defined as:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right) \quad (6)$$

where $a, b \in R$, a is a dilation factor, $a > 0$ and b is a translation factor.

For a common multidimensional wavelet function, we can construct it as the product of one dimensional wavelet functions:

$$\Psi(x) = \prod_{i=1}^{N} \psi(x_i) \quad (7)$$