Outlier detection for on-line monitoring data of transformer based on wavelet transform and weighted LOF

J F Qin¹, Y Yang¹, H Y Du²⁻³ and Z J Hong²

¹State Grid Shandong Electric Power Research Institute, Jinan, China
²School of Mathematical Sciences, Fudan University, Shanghai, China

E-mail: hydu15@fudan.edu.cn

Abstract. Power transformers are one of the core equipment in the power grid, so it is of great significance to guarantee transformers’ normal operation. By analyzing the dissolved gas content in transformer oil, we can monitor the operation of the power transformer. However, usually there are outliers in the data generated using the on-line monitoring system. In this paper, we propose a new outlier detection method based on wavelet transform and local outlier factor (LOF) algorithm. Using wavelet transform, we get the high dimensional representation of the original data in frequency domain, and by adding the weighted LOF (WLOF), we can identify outliers in high dimensional data set. Furthermore, we use the sliding window method to improve the efficiency of the algorithm, and achieve transformer oil on-line outlier detection efficiently. The experimental results on transformer data from several power transformers indicate that this algorithm can identify the outliers that exceed the threshold value, as well as the oscillations due to fluctuations in gas content. This can help achieve initial diagnosis of transformer oil on-line monitoring system rapidly.

1. Introduction

With the continuous development and advancement of smart grid construction, the amount of data generated by grid operation and monitoring has grown exponentially, which puts higher requirements for data processing. The power transformer is one of the most important equipment in the power system, and its stable operation affects the safety and reliability of the entire power system directly [1]. The on-line monitoring system of the transformer can simultaneously monitor the status and detect the anomaly of the transformer. This system generates a large amount of data, which can be used to detect the malfunction of the power system and facilitate troubleshooting. Therefore, on-line monitoring of the transformer has a significant impact on the safety and stability of the entire power system [2].

Based on current techniques, dissolved gas analysis (DGA) is one of the most effective ways to detect the transformer faults. The DGA method detects various characteristic gases dissolved in transformer oil, mainly including hydrogen (H₂), methane (CH₄), acetylene (C₂H₂), ethylene (C₂H₄), ethane (C₂H₆), carbon monoxide (CO) and carbon dioxide (CO₂) [3]. When an abnormal activity or malfunction occurs in the power transformer, the transformer oil will crack and generates a variety of characteristic gases, which can reflect the operating status of the transformer. And many researches and experiments have proved that on-line fault detection of transformer can be realized by monitoring the characteristic gas dissolved in transformer oil [4-6]. Jia [7] has summarized the advantages of on-line monitoring of dissolved gases in transformer oil. It is proposed that on-line monitoring is the initial diagnosis of equipment, and it can provide preliminary analysis of equipment...
failures. After finding the anomaly in the transformer, further use of chromatographic analysis can be implemented for follow-up diagnosis.

There have already been various methods using the DGA method to monitor the operating status of the transformer, such as Rogers codes [8] and IEC codes [9]. However, these fault detection methods are based on existing experiences, and the anomaly identification criteria are too absolute. Thus these methods do not perform well in most cases in practice. Recently, machine learning algorithms, neural networks are considered in smart grids, new algorithms such as clustering based method (CBT) [10], support vector machine (SVM) [11] and artificial neural network (ANN) [12] are widely used in the on-line anomaly monitoring of the transformer, and have achieved good experimental results. Although these methods can establish a non-linear relationship between DGA gas content and transformer fault types, they also have some problems in practice, such as the slow convergence speed, the inordinately long training time, and model overfitting due to lacking of training data with abnormal labels.

In [13], the dissolved gas in transformer oil is regarded as a kind of signal data, and most of these signals are non-stationary signals. Based on that, we propose an anomaly detection algorithm by analyzing the dissolved gas content in transformer oil based on wavelet decomposition coefficient and weighted local outlier factor. Wavelet transform is a fundamental method in signal processing. By multi-scale analysis of signals, the information in frequency domain can be extracted easily, and through sliding window analysis, we can obtain the high-dimensional representation of the original signal quickly. Breunig et al. proposed local outlier factor (LOF) algorithm [14] to measure the degrees of the outliers based on density. This method gives each data a factor indicating its degree of outlier-ness, but LOF algorithm will be inaccurate when there are many duplicate data. Based on the characteristics of the dissolved gas in transformer oil, we propose a new degree weighted LOF (WLOF) to identify the outliers in the original signal, and solve the problem of inaccurate calculation of LOF algorithm due to duplicate data in the signal.

The paper is organized as follows. In section 2, we introduce a frequency domain representation of the data based on wavelet transform and sliding window analysis. In section 3 we introduce some definitions of LOF and define weighted LOF. In section 4 we discuss in detail about the model construction. In section 5 we perform some experiments using our methods.

2. Frequency domain representation of the data based on wavelet transform

2.1. Introduction of wavelet transform

Wavelet transform is a digital signal processing technology developed from Fourier transform. It uses the transformation and scaling of wavelet base to perform multi-scale decomposition in time and frequency domains of signals. The process of decomposing a signal into wavelet components of different frequencies mainly depends on two functions: scale function \( \phi \) and wavelet function \( \psi \), which are both easy to describe. And according to the multi-scale analysis theory proposed by Mallet [15], the use of continuous \( \phi \) and \( \psi \) can improve the effect of wavelet decomposition. Let \( \{V_j, j \in \mathbb{Z}\} \) be the approximate space based on scale function \( \phi \) in \( L^2(\mathbb{R}) \), and they satisfy the following properties: (1) \( V_j \subset V_{j+1} \), (2) \( \bigcup_{j \in \mathbb{Z}} V_j = L^2(\mathbb{R}) \), (3) \( \bigcap_{j \in \mathbb{Z}} V_j = \{0\} \), (4) \( f(x) \in V_j \Leftrightarrow f(2^{-j}x) \in V_0 \). At this time, \( \phi \in V_0 \) and \( \{\phi(x-k), k \in \mathbb{Z}\} \) is the standard orthogonal basis of \( V_0 \). The following scale relationship can be established to this set of approximate space:

\[
\phi_{j-1,l} = 2^{-1/2} \sum_k p_{k-2l} \phi_{jk}
\]  

(1)

where \( \phi_{jk}(x) = 2^{j/2} \phi(2^j x - k) \), \( p_k = 2 \int_{-\infty}^{\infty} \phi(x) \phi(2x-k) dx \).

Since \( V_j \) is the subspace of \( V_{j+1} \), we can represent \( V_{j+1} \) as direct sum of \( V_j \) and its ortho
complement $W_j$. Similarly, we can find a function $\psi$ that generates the full space $W_j$ by transformation, and that is to find the standard orthonormal basis $\{\psi_{j,k}\}_{k \in \mathbb{Z}}$ of the space $W_j$. Then for any subspace $V_j$, we can obtain the orthogonal decomposition successively:

$$V_j = W_{j-1} \oplus V_{j-1}$$
$$= W_{j-1} \oplus W_{j-2} \oplus V_{j-2}$$
$$\vdots$$
$$= W_{j-1} \oplus W_{j-2} \oplus \ldots \oplus W_0 \oplus V_0$$

(2)

Similarly, we can obtain the orthogonal decomposition of $L^2(\mathbb{R})$ successively. Specifically, for any $f \in L^2(\mathbb{R})$, it can be uniquely represented as a sum form

$$\sum_{k=0}^{\infty} \pi_k \quad \text{where} \quad \pi_k \in W_k \quad \text{and they are mutually orthogonal},$$

which means wavelet function basis $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ is a standard orthonormal basis of $L^2(\mathbb{R})$. So when $j$ is big enough, it is possible to achieve approximation of any $f \in L^2(\mathbb{R})$ using wavelet function basis $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$, and $\sum_{k=0}^{\infty} \pi_k$ are called wavelet decomposition coefficients.

2.2. High dimensional representation of signal based on wavelet transform

We assume that the original data is a one-dimensional signal containing some noise. According to [16], we can select proper scale function $\{\phi_{j,k}\}_{j,k \in \mathbb{Z}}$ and wavelet function $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ and for a given positive integer $j_0$, it turns out that $f \approx f_h \in V_{j_0}$ in the approximate space $V_j$. Then we can begin with $f_h$, and decompose the original signal into the approximate part $f_{h-1} \in V_{h-1}$ and the wavelet part $w_{h-1} \in W_{h-1}$.

After initialization, $f$ can be decomposed into the following form based on discrete wavelet transform theory:

$$f(t) \approx f_j(t) = \sum_{k} c_h(k)2^{j \cdot 2} \phi(2^j t - k) + \sum_{k} d_h(k)2^{j \cdot 2} \psi(2^j t - k)$$

(3)

where $\phi$ is the scale function and $\psi$ is the wavelet function, $j_0 = 0$, and $j_1$ is a given positive integer representing the levels of wavelet decomposition. $c_h(k)$ are called approximate coefficients and $d_h(k)$ are called detail coefficients. Given the wavelet base and the levels of decomposition, we can obtain the wavelet coefficients step by step by the Mallat algorithm [15]:

Decomposition:

$$c_j(l) = 2^{-1} \sum_{k \in \mathbb{Z}} p_{k-2l} c_j(k)$$
$$d_j(l) = 2^{-1} \sum_{k \in \mathbb{Z}} (-1)^k p_{l-k+2} c_j(k)$$

(4)

Reconstruction:

$$c_j(k) = \sum_{l \in \mathbb{Z}} p_{k-2l} c_{j-1}(l) + \sum_{l \in \mathbb{Z}} (-1)^l p_{l-k+2} d_{j-1}(l)$$

(5)

And we can perform equation (4) by discrete filter form:

$$c_{j+1} = D(l * a_j)$$
$$d_{j+1} = D(h * a_j)$$

(6)
Where \( D \) is the downsampling operator, which takes samples every other time point. \( l_k = \frac{1}{2} p_{k+1} \) and 
\( h_k = \frac{1}{2} (-1)^k p_{k+1} \) are convolution operators corresponding to the two discrete filters, and \(*\) is the convolution operation.

The basis function \( \{\phi_{j,k}\}_{k \in \mathbb{Z}} \) describes the overall characteristics of the signal in the level \( j \), and \( \{\phi_{j,k}\}_{k \in \mathbb{Z}} \bigcup \{\psi_{j,k}\}_{k \in \mathbb{Z}} \) describes the smooth and oscillating features. Therefore, the corresponding wavelet coefficients can reflect characteristics of the original signal. The following theorem states that the approximate coefficients of signal can be approximated by orthogonalization of the original signal [17]:

**Theorem 2.1** [17]: Let \( \{V_j, j \in \mathbb{Z}\} \) be the approximate space based on scale function \( \phi \) in \( L^2(\mathbb{R}) \), if \( f \in L^2(\mathbb{R}) \) is continuous, then for \( j \) big enough:

\[
c_j(k) = 2^j \int_{-\infty}^{\infty} f(x) \phi(2^j x - k) dx \approx mf(k / 2^j)
\]

(7)

Where \( m = \int \phi(x) dx \).

According to the Theorem 2.1, for those \( j \) big enough, we can obtain the approximate coefficients of signal by orthogonalization of the original signal, so the coefficients represent the characteristic information of the signal. Given the wavelet base \( \phi \) and \( \psi \), and the levels of decomposition \( j \), then for the original signal \( f \in L^2(\mathbb{R}) \), we construct the following map \( F \) to acquire the high-dimensional representation of signal based on wavelet decomposition:

\[
F : N_{j_0} \rightarrow [j_{-k}]^{j-1} \backslash 1
\]

\[
F(f(k)) \rightarrow (c_{j_0}(k), d_{j_0}(k), d_{j_0+1}(k), \ldots, d_j(k))
\]

(8)

Where \( N_{j_0} \) is the sample space sampled at frequency \( f_s \).

Based on the Theorem 2.1 and equation (8), the original signal can be mapped into high-dimensional space \( [j_{-k}]^{j-1} \backslash 1 \) by wavelet decomposition. And we can decompose features of different frequencies effectively by selecting different \( j \) and \( j_0 \). And the anomaly in the original signal can be effectively identified by analyzing the characteristics of wavelet coefficients in high-dimensional space.

### 2.3 Sliding window analysis

We have already proposed a method to obtain the high-dimensional representation of the original signal in the above section, but in practice, the signal of power transform is updated very fast and contains a large amount of information. Meanwhile, the wavelet transform speed cannot be greater than that of FFT, which means the amount of computation of the wavelet transform is \( O(n \log n) \). We propose to improve the efficiency of the algorithm by adding the sliding window. The wavelet coefficients of the original signal in the previous time can be pre-stored by introducing the window function of length \( m \) \( (m \leq n) \). And when the new data is updated, the signal can be decomposed in a new window by simply shifting the window. This method avoids repeated coefficient calculations and the amount of computation each time is reduced to \( O(m \log m) \), thus it can improve the efficiency of the algorithm.

Assume that the original signal is \( f(t) \), and we use a window function \( \omega(t) \) of length \( m \) to intercept the signal, obtaining the truncated signal in the \( k \)-th window \( S(t) \):
\[ S(t) = f(t)w(t), \; km \leq t \leq (k + 1)m \] (9)

Assume that the time interval between two slides is \( \Delta T \), and the sampling frequency is \( fs' \), then the number of samples of the signal obtained in each window is \( \Delta N = \Delta T / fs' \). Then the truncated signal \( S \) intercepted by the sliding window is:

\[ S(km + i \cdot fs') = f(km + i \cdot fs')w(km + i \cdot fs'), \; i = 0, 1, \ldots, \Delta N, \; k \in \mathbb{Z} \] (10)

Or

\[ S(j) = f(j)w(j), \; j = km, km + fs', \ldots, km + \Delta Nfs', \; k \in \mathbb{Z} \] (11)

3. Weighted local outliers factor (WLOF)

After using the wavelet transform to obtain the high-dimensional representation of the original signal, we will perform anomaly monitoring by the outlier recognition algorithm. Breunig et al. proposed local outlier factor (LOF) algorithm based on density [14]. By calculating the local reachable distance of each data point, then comparing the local reachable density to the surrounding data points, LOF can give each data point a quantified indicator as the degree of outlier-ness. Although the traditional LOF algorithm can quantitatively identify the outliers, it performs unsatisfactory when there are many repeated data points. We propose weighted local outlier factor (WLOF), with the weights defined based on the repeatability of data points. WLOF can deal with the impact of data duplication better, so that it can efficiently identify the anomalies in transform signals.

The raw data point set is \( D_0 \) with partial repetition, and we introduce a weight for each point \( p \in D_0: w(p) = |R_p| \), where \( R_p = \{ q \in D_0 \mid q = p \} \) is the number of duplicate points in \( D_0 \). If \( p \) is a single point, then \( w(p) = 1 \). Then all the duplicated data points are taken as one point, and the processed set of non-repeated points is denoted as \( D \).

In the following part, we introduce some basic definitions of LOF in [14], and get the calculation formula of WLOF. We use \( d(p, q) \) to denote the Euclidean distance between \( p \) and \( q \), and \( k\text{-neighbor}_p \) denotes the set of \( k \) points that are closest to the point \( p \).

Definition 3.1 [14]: (k-distance of \( p \))

\[ k\text{-dist}(p) = \max\{d(p, q) \mid q \in k\text{-neighbor}_p \} \] (12)

Definition 3.2 [14]: (k-distance neighborhood of \( p \))

\[ N_k(p) = \{ q \in D \setminus \{ p \} \mid d(p, q) \leq k\text{-dist}(p) \} \] (13)

Definition 3.3 [14]: (k-reachability distance of an \( p \) w.r.t. \( q \))

\[ \text{reach-dist}_k(p, q) = \max\{k\text{-dist}(q), d(p, q)\} \] (14)

Definition 3.4: (weighted \( k\)-local reachability density of \( p \))

\[ \text{wlrd}_k(p) = \left( \frac{\sum_{q \in N_k(p)} w(o) \ast \text{reach-dist}_k(p, o)}{\sum_{q \in N_k(p)} w(o)} \right)^{-1} \] (15)

Definition 3.5: (weighted \( k\)-LOF of \( p \))
The value of the WLOF denotes the degree of outliers-ness. For most normal data point \( p \), it has similar \( k \)-reachability distance with the data points in its \( k \)-distance neighborhood, thus \( WLOF_i(p) \) is approximately equal to 1. On the contrary, the WLOF of an outliers \( \hat{p} \) will be significantly larger than 1. So the larger the WLOF value is, the greater the possibility of being outliers is.

Compared with the traditional LOF algorithm, if there exists a duplicated set \( \{ p_0, p_1, \ldots, p_i \} \) in \( D_0 \), where \( i > k \), then for any point in \( D_0 \), the \( k \)-reachability distance will reach 0, and \( \text{lrdr}_k(p_i) \to \infty \), which will result in a wide difference between the LOF value and the actual data distribution. By introducing the weight based on the repeatability of data points \( w(p) \), we can both avoid calculation abnormality caused by the repeated points in the local reachable density, and also measure the repeated distribution of data points.

4. Model construction

The construction of the on-line anomaly monitoring model proposed in this paper is mainly divided into five steps. The first step is to select the wavelet base \( \phi \) and \( \psi \), and the levels of decomposition \( j \), and truncate the historical observation data using the window function \( o(n) \). The second step is to perform wavelet decomposition with the specified base \( \phi \) and \( \psi \) on the truncated signal \( S(t) \), and store the corresponding wavelet coefficients. The third step is to map the wavelet coefficients into high-dimensional space using \( F : F(S(k)) \to (c_0(k), d_{j_0}(k), d_{j_1}(k), \ldots, d_j(k)) \), acquiring the high-dimensional representation of signal \( D \). The fourth step is to perform WLOF algorithm in \( D \) to obtain \( WLOF(k) \), and last is to set the anomaly recognition criteria and identify anomalous data based on the value of \( WLOF(k) \).

Specifically, when the model is constructed completely and the new observation \( f_i(t) \) is updated, where \( 1 \leq t \leq n_i \), the on-line monitoring of transformer data can be achieved as follows:

Step1: Assume that the number of sampling in each window is \( \Delta N \), and form a signal set with a certain repetition \( \hat{f}(t), 1 \leq t \leq n_i + l \), by intercepting a signal with length of \( l = \left[ \frac{n_i}{\Delta N} \right] \cdot \Delta N - n_i \) from the historical signal.

Step2: Obtain the truncated signal \( \hat{S}(t) \) according to equation (9).

Step3: Perform wavelet decomposition with the specified base \( \phi \) and \( \psi \), and the levels of decomposition \( j \) on the truncated signal \( \hat{S}(t) \) according to equation (6), and obtain the wavelet coefficients. Then map the wavelet coefficients into high-dimensional space according to equation (8), and acquire the high-dimensional representation of the truncated signal \( \hat{S}(t) \).

Step4: Denote \( D_1 = D \cup \hat{D} \), for \( \forall p \in \hat{D} \), compute \( k \)-dist \((p) \) and \( N_i(p) \) in \( D_1 \) according to equations (12) and (13).

Step5: If there exists \( q \in D \bigcap \bigcup_{p \in \hat{D}} N_i(p) \), then update \( k \)-dist \((q) \) and \( N_i(q) \) according to equations (12) and (13).

Step6: Compute the \( WLOF_i(p) \) for each data point \( p \) in \( D_1 \) according to equations (14)-(16), and identify anomalous data according to the criteria.
Step7: Store high-dimensional representations of all normal data into \( D \) to update the historical data set.

In this way, the anomalous information in the newly collected data can be identified, and the initial anomaly monitoring of the transformer oil chromatographic data can be completed. For the anomalous point detected, the second diagnosis can be further carried out by means of chromatographic analysis.

5. Simulations

We first demonstrate our fault diagnosis method on a transformer data set produced by a 220 kv main transformer from Jan. to Sep. in 2016. The oil chromatography on-line monitoring device used in this transformer has a sampling period of 12 hours, so the measured gas content is updated every 12 hours, and the total amount of data is around 500. Fix the missing point of the original data using the cubic spline method, and then we take total hydrocarbon value as an example to test our fault diagnosis methods. The total hydrocarbon value is the sum of four gas contents of \( CH_4 \), \( C_2H_4 \), \( C_2H_6 \) and \( C_3H_2 \), and figure 1 shows the original data of total hydrocarbon value.

![Figure 1. Original data of total hydrocarbon value.](image)

We consider using a window of length \( m = 24 \) to perform fault diagnosis. First, we obtain the truncated signal with a length of \( m \) following the steps 1 and 2. Then we use rectangular window and Hamming window as shown in equations (17) and (18), respectively:

\[
\text{Rectangular window: } \omega(n) = R_{\Delta N}(n) = \begin{cases} 
1, & 0 \leq n \leq \Delta N \\
0, & \text{otherwise}
\end{cases} 
\]

\[
\text{Hamming window: } \omega(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{\Delta N - 1}\right), 1 \leq n \leq \Delta N
\]

Use the ‘db2’ wavelet base, and chose the levels of decomposition to be 3, then we can obtain the high-dimensional representation of the original data and the corresponding WLOF value following the steps 3 - 6. The results are shown in figures 2 and 3.

When using WLOF value to perform anomaly monitoring, the criteria or the thresholds are generally chosen empirically. According to the test results on the total hydrocarbon value data of multiple different main transformers, the data points with WLOF values greater than 2 are generally
identified as ‘potential anomalous data’ or ‘potential outliers’, and the data points with WLOF values greater than 2.5 are identified as ‘anomalous data’ or ‘outliers’. Following the above identification criteria, we can obtain the fault diagnosis results using two different window functions, and the results are shown in the figures 4 and 5, where the black curve represents the original data, and star points are suspected outliers with WLOF values between 2 and 2.5, while the solid circle points are the anomalies with WLOF values greater than 2.5.

Figure 2. WLOF using rectangular window.
Figure 3. WLOF using Hamming window.

Figure 4. Outlier detection results using rectangular window.
Figure 5. Outlier detection results using Hamming window.

The results shown above indicate that the fault diagnosis using our methods can not only obtain the anomalous data where the value exceeds the normal threshold, but also consider the fluctuation frequency of the original data. Moreover, it can be clearly seen that, compared with the Hamming window, the rectangular window is more sensitive for the signal frequency, so it is more likely to identify the data that is inconsistent with the fluctuation frequency of the surrounding data as an anomaly.

Furthermore, in order to test the fault diagnosis method for transformer on-line anomaly monitoring, we take the historical data of a 220 kv main transformer in 2016 and 2017 as an example, and apply our algorithm to achieve on-line anomaly monitoring. Similarly, we take total hydrocarbon value as an example. First, we process 4000 data to build the historical data set. Here we also select ‘db2’ as the wavelet base and the levels of decomposition is chosen to be 3. We use a rectangular
window with a length of $m=96$ to perform sliding time window wavelet transform and store the corresponding wavelet coefficients and WLOF value. The results are shown in figures 6 and 7.

Figure 6. Total Hydrocarbon value historical data.

Figure 7. WLOF value of the data in figure 6.

In order to simulate the situation of on-line monitoring, we assume that the device receives 96 new data points every day. The observation data of the next 10 days are sequentially input, and we perform on-line outlier detection following steps 1 – 7. According to the identification criteria, we mark data points with WLOF values greater than 2.5 as anomalies and keep the wavelet coefficients of other normal data, then update the historical data. By repeating the above procedures, we analyze the next 10 days’ data in turn to obtain the anomaly detection results as shown in figure 8, where the black curve represents the normal data, and blue solid circle points are anomalies in real historical data, and red star points are anomalies recognized by our methods. It can be seen that the fault diagnosis method we propose can perform on-line anomaly detection in transform data well, which can identify both anomaly data with values exceeding the normal thresholds, and abnormalities caused by fluctuations and rapid oscillation in gas content.

Figure 8. The result of on-line anomaly detection on total hydrocarbon value.

6. Conclusions
In this paper, we propose an outlier detection method of on-line monitoring data of power transformer based on wavelet transform and weighted LOF (WLOF). This method takes both the advantages of the
wavelet transform to extract the overall characteristics and local features of digital signals, and WLOF algorithm to identify the outliers based on density quickly and accurately. It can analyze the fluctuation characteristics of the data according to the historical data, and obtain the high-dimensional representation of the data using wavelet transform, and then identify the anomaly using WLOF. The experimental results show that this method is able to identify anomalies in transformer data accurately, and realize the initial diagnosis of the transformer quickly and efficiently.

Acknowledgments
The research of this work is supported by Science and Technology Commission of Shanghai Municipality 16JC1402600.

References
[1] Liang Y L, Li K J, Zhao J G 2014 Research on the dynamic monitoring cycle adjustment strategy of transformer chromatography on-line monitoring devices Proceedings of the CSEE 34 1446-53
[2] Zou J M 2007 Application of on-line monitoring technology on power grid High Voltage Engineering 8 203-6
[3] Hung C P and Wang M H 2004 Diagnosis of incipient faults in power transformers using CMAC neural network approach Electr. Power Syst. Res. 71 235-44
[4] Wang Y Y, Liao R J, Chen W G and Du L 2005 Online fault predicting method based on dissolved gas analysis for transformer J. Chongqing Univ. (Natural Science Edition) 7 34-7
[5] Chen W and Sun C 2000 Study on fault diagnosis methods of transformer DGA with fuzzy model hierarchy classification ICMEP 9 167-71
[6] Aburaghieta E, Farrag M E, Hepburn D M and Garcia B 2016 Enhanced condition monitoring of power transformers through improvement in accuracy of DGA interpretation 2016 51st International Universities Power Engineering Conference (UPEC) (Coimbra, Portugal) 1-6
[7] Jia R 1998 Review about gas dissolved in transformer oil on-line detection Power Syst. Technol. 22 49-55
[8] Rogers R R 1978 IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis IEEE Trans. Elect. Insul. 5 349-54
[9] Duval M 2001 Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases IEEE Electr. Insul. Mag. 17 31-41
[10] Xiong H, Sun C X, Liao R J, Li J and Du L 2005 Study on kernel-based possibilistic clustering and dissolved gas analysis for fault diagnosis of power transformer Zhongguo Dianji Gongcheng Xuebao (Proc. Chinese Soc. Electr. Engin.) 25 162-6
[11] Zheng H B, Liao R J, Grzybowski S and Yang L J 2011 Fault diagnosis of power transformers using multi-class least square support vector machines classifiers with particle swarm optimisation IET Electr. Power. App. 5 691-6
[12] Gao J and He J 2010 Application of quantum genetic ANNs in transformer dissolved gas-in-oil analysis Proc. CSEE 30 121-7
[13] Chen W, Pan C, Yun, Y and Liu Y 2009 Wavelet networks in power transformers diagnosis using dissolved gas analysis IEEE Trans. Power Deliv. 24 187-94
[14] Breunig M M, Kriegel H P, Ng R T and Sander J 2000 LOF: Identifying density-based local outliers ACM Sigmod Record 29 93-104
[15] Mallat S 1989 A theory for multi-resolution approximation: the wavelet approximation IEEE Trans. PAMI 11 674-93
[16] Vetterli M and Herley C 1992 Wavelets and filter banks: Theory and design IEEE Trans. Signal Process. 40 2207-32
[17] Boggess A and Narcowich F J 2015 A First Course in Wavelets with Fourier Analysis John Wiley and Sons (USA) 165-7