Power transformer fault diagnosis using FCM and improved PCA

Tusongjiang Kari, Wensheng Gao

Department of Electrical Engineering, Tsinghua University, Beijing 100084, People’s Republic of China
E-mail: tsjkr14@mails.tsinghua.edu.cn

Published in The Journal of Engineering; Received on 17th November 2017; Accepted on 24th November 2017

Abstract: In order to improve fault diagnosis accuracy of power transformer, a new fault diagnosis model based on fuzzy C-means (FCM) clustering algorithm and improved principal component analysis (IPCA) is presented. First, dissolve gas analysis samples are clustered with FCM and cluster centre for each fault type is regarded as reference sequence. Then, the IPCA approach is implemented to obtain main principal components containing 95% of original information. Finally, Euclidean distances between principal components of reference sequence and testing sample are calculated to identify final fault type. Cases studies and test results show that the proposed approach achieves recognition of transformer fault effectively and has a higher diagnostic accuracy than the international electrotechnical commission (IEC) ratio method and the improved three ratio method.

1 Introduction

Power transformer is one of the most expensive and vital assets in power system, which plays an important role in converting voltage level, connecting transmission and distribution networks [1]. Power transformer operating condition has direct impact on safety, reliability and stability of the power system. However, faults occurrence of power transformer seems inevitably due to aging and kinds of stresses, including electrical, mechanical, thermal and environmental stresses. Therefore, it is imperative to detect and diagnose incipient faults of power transformer early and correctly for a reliable and steady operation of power system.

Dissolved gas analysis (DGA) is an extensively used interpretation approach for power transformer fault diagnosis. The main advantage of the DGA method is that it diagnoses faults simply and effectively regardless of magnetic field around. Besides, DGA can be carried out without derogation of power transformer [2]. Plenty of fault interpretation techniques have been developed based on DGA, such as the key gas method [3, 4], IEC ratio method [5], improved three ratio (ITR) method [6], Roger ration method [7], Dornenburg ratio method [8] and Duval triangle method [9]. DGA methods mentioned above have made progresses in practical application, but there are still some drawbacks, such as code absence, sharp boundary of codes, diagnosis inconsistence and undesired performance [10]. Recently, since the development of artificial intelligent (AI) and machine learning, various AI techniques and theories have been utilised in fields of power transformer fault diagnosis. Support vector machine (SVM) [11], artificial neuro-networks (ANNs) [12], fuzzy cluster method (FCM) [13], grey theory [14], matter element analysis [15] and evidential theory [16] have been applied to improve fault diagnosis performance. There are limitations of AI techniques mentioned above, i.e. ANN has good self-learning and recognition ability but with slow convergence and local optimal. For SVM, suitable kernel function and parameters are hard to determine and new structures need to be designed for multi-class problems. Determination of distinguish coefficient and standard sequence is a complex task for grey theory. Effective fusion of conflict evidences is still a challenge for evidence theory. Therefore, there are rooms for traditional DGA techniques and AI-based models to further improve the fault diagnosis capability.

Many variables can be used to interpret incipient faults of power transformer. Nevertheless, these variables may be related and redundant, which makes the fault diagnosis model more complex and diagnosis performance poor. So, it is necessary to reduce dimension and remove redundancy in variables. The PCA method is a simple yet effective way to compress data space, extract key features and eliminate redundancies between variables. A DGA-based fault diagnosis method using fuzzy C-means (FCM) clustering algorithm and improved principal component analysis (IPCA) is put forward in this paper. Cluster centres of original DGA dataset obtained by FCM are used as reference sequences. Main principal components of sample matrix are computed using IPCA. Final diagnosis results are based on the comparison of Euclidean distance between reference sequences and testing samples. Results of case studies and experimental test verify the validity and accuracy of the proposed method.

2 Establishment of reference sequence

Power transformer reference sequence is usually employed to determine fault types according to the distance between testing samples and typical sequences [17]. Fault types corresponding to the minimum distance can be treated as diagnosis result. So, creation of typical sequence, known as reference sequence, has a significant influence on the determination of final result. Till now, most of reference sequence is based on the mean value of input variables, which is subjected to limited samples size, imprecise and inconsistency. Besides, the conventional normalisation process to acquire reference sequence puts too much emphasis on samples’ uniformity and pays less attention on sample difference or distribution characteristic. As a result, the reference sequence may not be typical enough and lead to inaccurate results. Therefore, more reliable and precise reference sequences are needed in fault diagnosis of power transformer.

2.1 Data normalisation

Different range and magnitude of raw samples result in long computational time, low accuracy and poor generalisation of the learning model. Hence, normalisation is a critical preprocessing of AI and machine learning. In general, the original data matrix $X$ is

$$J. Eng., 2017, Vol. 2017, Iss. 14, pp. 2605–2608$$

This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0)
normalised with (1) shown as

$$X^* = (X - X_m)/S$$  \hspace{1cm} (1)$$

where $X^*$ is the standardised matrix, $X_m$ stands for the mean value of each variable and $S$ is standard deviation.

In order to take sample difference and distribution characteristic into consideration, relative proportion of samples is used to standardise the sample matrix, described as follows:

$$x_{ij}^* = x_{ij} / \sum_{j=1}^{m} x_{ij} (i = 1, 2, \ldots; j = 1, 2, \ldots m)$$  \hspace{1cm} (2)$$

where $x_{ij}$ and $x_{ij}^*$ are the variables before and after normalisation, respectively, $n$ stands for samples size and $m$ for dimension number.

2.2 Fuzzy c-mean cluster algorithm

FCM clustering algorithm is a popular unsupervised clustering approach, which is introduced by Dunn and has been intensively used in many different fields, such as quantisation, segmentation and pattern recognition [18]. FCM implements fuzzy clustering by searching sets of clusters and determined cluster centres that are intend to mark the average location of each cluster. FCM assigns each data a membership grade for each cluster and updates cluster centres and membership grade iteratively to make cluster centre moving towards the ‘right’ location of dataset [19]. Determination of final cluster centre can be regarded as an optimisation issue that minimises the constrained objective function presented as follows:

$$\min J = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2$$

subject to $\sum_{j=1}^{c} u_{ij} = 1$  \hspace{1cm} (3)$$

where $c$ and $n$ are the number of clusters and samples in dataset, respectively. $\mu_{ij} \in [0, 1]$ presents the membership degree of the data $x_{ij}$ in the $i$th cluster. $d_{ij}$ is the Euclidean distance between the $i$th cluster centre $v_i$ and $j$th data $x_{ij}$, $m \in [1, + \infty]$ is a weighting exponent and generally sets to 2.

With the application of the Lagrange multiplier method, an iterative optimisation is carried out and updated membership $u_{ij}^*$ and cluster centres $v_i$ are shown, respectively, as (4) and (5):

$$u_{ij} = \frac{\sum_{k=1}^{c} (d_{ij}/d_{ik})^{-(m-1)}}{\sum_{k=1}^{c} (d_{ij}/d_{ik})^{-(m-1)}} (1 \leq i \leq c, 1 \leq k \leq n)$$  \hspace{1cm} (4)$$

$$v_i = \sum_{j=1}^{n} u_{ij}^m x_{ij} / \sum_{j=1}^{n} u_{ij}^m (1 \leq i \leq c)$$  \hspace{1cm} (5)$$

The iteration will stop when (6) is satisfied:

$$\|u_{ij}^{t+1} - u_{ij}^{t}\| < \delta$$  \hspace{1cm} (6)$$

where $\delta$ is a termination criterion and $t$ is the iteration number.

To sum up, the establishment of reference sequence consists of the following steps:

(i) normalise raw dataset with (2);
(ii) initialise FCM algorithm, including initial cluster centres, cluster number, iteration number, weighting exponent and termination criterion;
(iii) calculate objective function and update membership and cluster centres with (4) and (5); and
(iv) terminate calculation process if (6) is satisfied or return step (iii).

The final cluster centres are viewed as reference sequence of each fault types and used to determine incipient fault of new DGA samples [20].

3. Improved principle component analysis

PCA is one of the most classical feature extraction methods in statistical analysis. This approach extracts significant factors via linear transformation and gets minimal dimensions to represent original information. Therefore, PCA is used to reduce variables dimensions, eliminate redundant information, simplify structure of classifier with acceptable information lost and promote classification performance [21].

Generally speaking, information contained in the raw dataset can be divided into two parts. One is the variables relationship information that is presented by correlation coefficient, and another one is the variables’ difference information that is revealed through variance. The common method for normalisation is shown as (1), with whom the mean value is ‘0’ and variance is ‘1’. As a result, relationship information is kept only and difference information is lost. Furthermore, the PCA algorithm using the aforementioned normalisation method may not be able to represent original information entirely and correctly [22]. To overcome these shortcomings, an averaging approach described as (7) is employed to normalise raw data:

$$x_{ij}^* = x_{ij} / \left(\frac{1}{n} \sum_{j=1}^{m} x_{ij}\right) (i = 1, 2, \ldots; j = 1, 2, \ldots m)$$  \hspace{1cm} (7)$$

where $x_{ij}^*$ is the standardisation value of $x_{ij}$. The main advantage of averaging value normalisation method is that there is no different in correlation matrix and no information lost after normalisation [23].

Given data matrix $X_{n \times p}$, then the implementation steps of PCA are listed as below:

(i) Normalisation of data matrix $X$: In order to keep all information, equation is applied to standardise the original samples.

(ii) Build correlation coefficient matrix $R$ and calculation of eigenvalues $\lambda$ and eigenvector $\gamma$: The correlation coefficient can be computed with equation (8) after normalization and set up the correlation coefficient matrix $R$. After solving characteristic equation $|\lambda I - R| = 0$, eigenvalues $\lambda$ and corresponding eigenvector $\gamma$ are obtained:

$$r_{ij} = \frac{\sum_{k=1}^{n} (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{\sqrt{\sum_{k=1}^{n} (x_{ik} - \bar{x}_i)^2 \sum_{k=1}^{n} (x_{jk} - \bar{x}_j)^2}}$$  \hspace{1cm} (8)$$

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\
 r_{21} & r_{22} & \cdots & r_{2m} \\
 \vdots & \vdots & \ddots & \vdots \\
 r_{m1} & r_{m2} & \cdots & r_{mm} \end{bmatrix}$$

where $r_{ij}$ is the correlation coefficient and $\bar{x}_i$ is the mean value.

(iii) Determination of the principal component: Contribution rate $\eta_j$ and cumulative contribution rate $\eta(h)$ can be obtained with (9) and (10), respectively. $\eta(h)$ is used to evaluate the information proportion contained by the first $h$ principal components. Threshold value for $\eta(h)$ is often set to 85% to provide...
enough information for further application:

\[
\eta_i = \left( \frac{\lambda_i}{\sum_{k=1}^{m} \lambda_k} \right) \times 100\% \quad (i = 1, 2, \ldots, m) \tag{9}
\]

\[
\eta(h) = \sum_{i=1}^{h} \lambda_i / \sum_{i=1}^{m} \lambda_i > 85\%
\]

\[(i = 1, 2, \ldots, h, \ldots; m; h < m) \tag{10}\]

(iv) Calculate the first \( h \) principal component of sample: After determining the exact number of \( h \), then principal components \( Y_i \) of samples are obtained using (11):

\[
Y_i = \begin{bmatrix}
Y_{i1} \\
Y_{i2} \\
\vdots \\
Y_{ik}
\end{bmatrix} = \begin{bmatrix}
\gamma_{11} & \gamma_{12} & \cdots & \gamma_{1p} \\
\gamma_{21} & \gamma_{22} & \cdots & \gamma_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
\gamma_{ih} & \gamma_{i2} & \cdots & \gamma_{ip}
\end{bmatrix} x_i \tag{11}
\]

\[(i = 1, 2, \ldots, p)\]

4. Power transformer fault diagnosis based on IPCA

4.1 Preprocessing of fault diagnosis

A total of 524 DGA samples with labels are selected from literatures and records of China Southern Power Grid. Five types of key gases, including hydrogen (H\(_2\)), methane (CH\(_4\)), ethane (C\(_2\)H\(_6\)), ethylene (C\(_2\)H\(_4\)) and acetylene (C\(_2\)H\(_2\)), are used as input variables of diagnose model. According to IEC 60599 and GBT 7275 [6], fault types are classified into six categories: low energy discharge (D1), high energy discharge (D2), partial discharge (PD), thermal fault of low temperature (T1), thermal fault of medium temperature (T2) and thermal fault of high temperature (T3). The sample distribution of each fault types is shown as Table 1.

To obtain the reference sequence, a matrix with all 524 samples is first normalised with (2), and then the FCM method is implemented to obtained six cluster centres for all fault types, and the result is shown in Table 2.

Table 1 Sample distribution of fault types

| Fault type | D1 | D2 | PD | T1 | T2 | T3 |
|------------|----|----|----|----|----|----|
| Number     | 70 | 157| 60 | 42 | 55 | 140|

Table 2 Cluster centres of DGA samples

| Fault types | Relative percentage of gases, % |
|-------------|---------------------------------|
|             | H\(_2\) | CH\(_4\) | C\(_2\)H\(_6\) | C\(_2\)H\(_4\) | C\(_2\)H\(_2\) |
| D1          | 52.00  | 13.33   | 7.33          | 8.67       | 18.67        |
| D2          | 40.00  | 15.95   | 1.69          | 19.62      | 22.74        |
| PD          | 83.59  | 7.42    | 2.97          | 5.97       | 0.03         |
| T1          | 35.91  | 32.84   | 7.89          | 23.02      | 0.34         |
| T2          | 10.19  | 34.58   | 18.78         | 36.26      | 0.21         |
| T3          | 12.26  | 27.04   | 9.64          | 50.63      | 0.42         |

It can be concluded from Table 2 that, for discharging faults (such as D1, D2 and PD), the major proportion of gases are hydrogen (H\(_2\)). While the proportion of C\(_2\)H\(_2\) and C\(_2\)H\(_4\) increases as discharging intensity increases. Meanwhile, the main proportion for T1 is H\(_2\) and CH\(_4\), but the proportion of CH\(_4\) and C\(_2\)H\(_2\) increases as temperature rises. The major proportion of key gas and changing tendency suggested by Table 2 is agreed with [4–6].

After reference sequences are established, a sample matrix \( X \) combined with reference sequence and testing samples is built and IPCA is implemented to extract main principal components from \( X \) and determine final diagnosis result of testing samples. The main steps of IPCA are described in Section 3. When the first \( h \) principal components are gained by following step (iv), the Euclidean distance \( D \) between principal components of testing sample \( C_t \) and reference sequence \( C_i \) can be calculated with (12) [24]:

\[
D_t = \left[ \sum_{k=1}^{h} \sqrt{(C_{tk} - C_{ik})^2} \right]^{1/2} \tag{12}
\]

Fault type of the reference sequence corresponding to the minimum distance stand for the diagnose result. For example, if the minimum distance is the third one \( d_3 \), then the final result is high energy discharge.

4.2 Case study and experimental test

In order to reveal the fault diagnosis process concretely and demonstrate the validity and effectiveness of the proposed approach, a practical case and another 130 samples from China Southern Power Grid are studied.

Case 1: Dissolve gases of a power transformer (type: SF7-31500/110) are tested after alarm. Concentration of H\(_2\), CH\(_4\), C\(_2\)H\(_6\), C\(_2\)H\(_4\) and C\(_2\)H\(_2\) is 1582, 95.5, 35, 48 and 0 (units: ppm). Interpretation codes based on IEC ration methods are 0, 1 and 1, which is out of code and unable to determine exact fault types [24].

According to the proposed methods, the testing sample is normalised and then the sample matrix \( X \) is created. After that IPCA is applied to calculate correlation coefficient matrix \( R \), eigenvalues \( \lambda \) and the corresponding eigenvectors \( \gamma \), which are shown as below:

\[
X = \begin{bmatrix}
52.00 & 13.33 & 7.33 & 8.67 & 18.67 \\
40.00 & 15.95 & 1.69 & 19.62 & 22.74 \\
83.59 & 07.42 & 2.97 & 5.97 & 0.03 \\
35.91 & 32.84 & 7.89 & 23.02 & 0.34 \\
10.19 & 34.58 & 18.78 & 36.26 & 0.21 \\
12.26 & 27.04 & 9.64 & 50.63 & 0.42 \\
89.86 & 5.42 & 1.98 & 2.72 & 0.00
\end{bmatrix}
\]

\[
R = \begin{bmatrix}
0.206 & -0.5923 & 0.1524 & 0.1360 & 0.7517 \\
-0.2199 & 0.2844 & 0.0645 & -0.830 & 0.4215 \\
-0.3196 & 0.3683 & 0.7843 & 0.3506 & 0.1552 \\
-0.2784 & 0.4342 & -0.5924 & 0.4114 & 0.4633 \\
0.8541 & 0.4950 & 0.0803 & -0.0188 & 0.1363
\end{bmatrix}
\]

Table 3 Contribution rate and cumulative contribution rate of principal component

| Component | Eigenvalue | Contribution, % | Cumulative contribution, % |
|-----------|------------|-----------------|---------------------------|
| PC1       | 2.7505     | 67.38           | 67.38                     |
| PC2       | 1.0421     | 25.53           | 92.91                     |
| PC3       | 0.2260     | 5.54            | 98.45                     |
| PC4       | 0.0631     | 1.55            | 100                       |
| PC5       | 0          | 0               | 100                       |
Comparison results of fault diagnosis are presented in Table 4. The proposed approach is capable to diagnose incipient faults of power transformer effectively and correctly. Moreover, the proposed method has higher accuracy than that of the IEC ratio method and ITR method.

5. Conclusion
A fault diagnosis method based on FCM clustering and IPCA algorithm is proposed in this paper. This method normalises raw dataset with relatively percentage of gas concentration and regards cluster centres of sample matrix as reference sequence. IPCA is then applied to reduce dimension and eliminate redundancy in input variables. The distance between principal components of testing sample and reference sequence is used to determine final diagnostic results. Case study and experimental test verify the effectiveness and accuracy of the proposed approach. Comparison results show that the proposed methods have a better fault diagnosis performance than the IEC ratio method and ITR method.

Table 4 Comparison of fault diagnosis results

| Fault type | Number | IEC | ITR | Proposed |
|------------|--------|-----|-----|----------|
| LD         | 20     | 70  | 80  | 85       |
| HD         | 20     | 75  | 80  | 90       |
| PD         | 25     | 68  | 68  | 88       |
| LT         | 20     | 80  | 80  | 90       |
| MT         | 20     | 75  | 85  | 90       |
| HT         | 25     | 68  | 76  | 88       |
| Total      | 130    | 72.7 | 78.2 | 88.5    |

The distance $D$ between principal component of testing sample and other reference sequence is shown as below:

$$D = [3.3022, 4.0009, 0.2649, 2.1098, 3.6085, 3.1898]$$

The minimum distance is $d_1 = 0.2649$, which indicates that final diagnosis result of the testing sample is partial discharge. The fault diagnosis result accords with actual situation, which manifests the fault diagnosis capability of the proposed approach.

In order to evaluate effectiveness and accuracy of the proposed approach, 130 extra samples provided by a power supply bureau are diagnosed by three different interpretation techniques, including the IEC ratio method, the ITR method and the proposed methods. Comparisons results of fault diagnosis are presented in Table 4. It is illustrated from Table 4 that the proposed approach is capable to diagnose incipient faults of power transformer effectively and correctly. Moreover, the proposed method has higher accuracy than that of the IEC ratio method and ITR method.

6. Acknowledgments
This project is supported by the Nation High-tech R&D program of China (No. 2015AA050201). Also, the authors thank all the reviewers for their useful comments.

7 References
[1] Faiz J., Soleimani M.: ‘Dissolved gas analysis evaluation in electrical power transformers using conventional methods a review’, IEEE Trans. Dielectr. Electr. Insul., 2017, 24, (2), pp. 1239–1248
[2] Sun H.C., Huang Y.C., Huang C.M.: ‘A review of dissolved gas analysis in power transformer’, Energy Proc., 2012, 14, pp. 1220–1225
[3] IEE Standard C57.104: ‘IEE guide for the interpretation of gases generated in oil-immersed transformers’, 2008
[4] IEC Standard 60596: ‘International Electro-Technical Commission Guide for the sampling of gases and of oil-filled electrical equipment and for the analysis of free and dissolved gases’, 2005
[5] IEC Standard 60599: ‘Mineral oil-impregnated electrical equipment in service guide to the interpretation of dissolved and free gases analysis’, 2007
[6] GB/T 7725: ‘Guide to the analysis and the diagnosis of gas dissolved in transformer oil’, 2001
[7] Rogers R.R.: ‘IEEE and IEC codes to interpret incipient faults in transformers using gas in oil analysis’, IEEE Trans. Electr. Insul., 1978, 13, pp. 349–354
[8] Dornenburg E., Strittmatter W.: ‘Monitoring oil-cooled transformers by gas analysis’, Brown Boveri Rev., 1974, 61, pp. 238–247
[9] Duvall M.: ‘A review of fault detectable by gas-in-oil analysis in transformers’, IEEE Electr. Insul. Mag., 2002, 18, pp. 8–17
[10] K. Fard, R.H., Costa J.G.S., Oliwas J.L.M.: ‘A review of monitoring methods for predictive maintenance of electric power transformers based on dissolved gas analysis’, Renew. Sustain. Energy Rev., 2015, 46, pp. 201–209
[11] Yin Z., Hou J.: ‘Recent advances on SVM based fault diagnosis and process monitoring a complicated industrial processes’, Neurocomputing, 2016, 174, pp. 643–650
[12] Jian Li Y.U., Shuai B.: ‘Fault diagnosis model of transformer based on BP neural network’, J. Syst. Simul., 2014, 26, (6), pp. 1343–1349
[13] Weihe L., Rujin L., Lijun Y.: ‘Power transformer fault diagnosis based on dot density weighted fuzzy kernel clustering’, Electr. Power Autom. Equip., 2012, 32, (6), pp. 66–69
[14] Lei Y., Lin D., Junwei W.: ‘Transformer Multi-parameter integrated fault diagnosis method based on fuzzy membership function’, High Volt. Appar., 2011, 47, (5), pp. 35–42
[15] Wu Y., Zhu H.B., Zhou Z.C., et al.: ‘Transformer condition assessment based on entropy fuzzy matter-element and principal component analysis’, Power Syst. Prot. Control, 2015, 43, (17), pp. 1–7
[16] Tianfang Y., Haifang L., Xiangjun Z., et al.: ‘Application of a combined decision model based on optimal weights in incipient faults diagnosis for power transformer’, IEEE Trans. Electr. Electron. Eng., 2017, 12, (2), pp. 169–175
[17] Zhao F., Li S.: ‘Fault diagnosis for traction transformer based on DGA and improved grey correlation analysis model’, High Volt. Appar., 2015, 47, (15), pp. 41–45
[18] Chatoopadhyay S.: ‘A comparative study of fuzzy C-means algorithm and entropy-based fuzzy clustering algorithm’, Comput. Inf., 2011, 30, pp. 701–720
[19] Sopharak A., Uyyanonvara B., Barman S.: ‘Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy C-means clustering’, Sensor, 2009, 9, pp. 2148–2161
[20] Xin D.U., Gao R.G.: ‘Transformer fault diagnosis and analysis based on FCM’, J. Electr. Power, 2011, 26, (5), pp. 376–379
[21] Tang Y., Gui W., Fang T., et al.: ‘Transformer fault diagnosis model based on PCA and KPCA feature extraction’, High Volt. Eng., 2014, 40, (2), pp. 557–563
[22] Jie W., Hesheng Z., Liming J.: ‘State monitoring approach based on improved principal component analysis’, J. Electron. Meas. Instrum., 2009, 25, (7), pp. 51–55
[23] Yajing X., Yuanzhong W.: ‘The improvement of the application method of principle component analysis’, Math. Pract. Theory, 2006, 36, (6), pp. 68–75
[24] Tingfang Y., Hang Z., Libin H.: ‘Incipient fault diagnosis based on improved principal component analysis for power transformer’, Electr. Power Autom. Equip., 2015, 35, (6), pp. 149–153