IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 13, 2019, accepted February 25, 2019, date of publication February 28, 2019, date of current version March 18, 2019.
Digital Object Identifier 10.1109/ACCESS.2019.2902299

Fault Diagnosis of Power Transformers With Membership Degree

ENWEN LI, LINONG WANG, AND BIN SONG
School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China
Corresponding author: Enwen Li (lienwen@whu.edu.cn)

ABSTRACT Power transformers are important equipment for power systems, and a dissolved gas analysis (DGA) is widely used to detect incipient faults in oil-pregnant transformers. The conventional methods are prone to misinterpreting the gas data near the boundaries and the correct rate is low. Though a high correct rate is reported with intelligent methods as artificial neural network, support vector machine, and so on, these methods are usually too complicated to be implemented practically on a wide range. Based on clustering techniques, this paper proposes a new method for fault diagnosis of transformers with the DGA. A reference fault set is provided, and the fault diagnosis is implemented by calculating the membership of the DGA data to the reference fault set. Test with credible DGA dataset (201 field cases) shows that the correct rate of the new method is 89%, while the David triangle method is 79% and the IEC ratio method is 59%, which demonstrate the superiority of the proposed method to the conventional ones. The new method is simple and highly accurate, indicating a good application prospect in engineering practice.

INDEX TERMS Power transformer, fuzzy clustering, fault diagnosis, membership degree.

I. INTRODUCTION

The oil-paper insulation system in power transformers operates under the effects of high temperature and strong electromagnetic environment, and the insulation medium can slowly decompose into a number of small molecules. The decomposition gases dissolved in oil are H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, CO_2, CO and N_2. However, when a fault occurs, the insulation breaks down more quickly and the decomposition products will be different according to the type and severity of the fault [1], [2]. Dissolved gas analysis (DGA) is widely used to detect incipient faults in oil-pregnant transformers. This technique involves several steps, such as taking oil samples from a transformer, removing dissolved gases from oil, determining gas component content, and identifying fault types [3]. Fault identification is a decisive step in the internal fault state determination of the transformer in DGA analysis.

Various computational and graphical methods employing gas ratios and proportions of gases dissolved in oil determined by gas chromatography have been worked out for recognizing the characteristic patterns of the dissolved gases that are associated with the main types of faults [4], [5]. These methods available to interpreted DGA data include Key Gas Method, Doernenburg Ratio Method, Rogers Ratio Method, IEC Ratio Method and Duval Triangle Method, and they have been developed and validated using large sets of data for equipment in service. In these methods, the multiple numeric thresholds and gas boundaries are commonly set to classify features of the dissolved gas data. However, these thresholds and boundaries do not physically exist, and the gas data near the ratio boundaries are prone to misinterpretation. Moreover, these methods are often unable to diagnose complex faults when both thermal faults and discharge faults occur simultaneously. And at this case, if the overheat fault is more serious, it is easy to be diagnosed as an overheat fault, and vice versa. Therefore, the fault diagnosis accuracy rate of these methods is relatively low [4], [6]. Some intelligent approaches are also available for fault diagnosis of large oil-immersed power transformers [7], [8], such as fuzzy logic [9], [10], support vector machine [11], [12], artificial neural network [13], genetic programming [14], rough set theory [15], deep belief network [16], etc. and good results have been claimed for those approaches. However, these methods are usually complicated, and their results depend on training data, so the research work documented in these publications become difficult to reproduce for the lack of publicly available DGA data. Therefore, intelligent methods do not implemented practically on a wide range as

The associate editor coordinating the review of this manuscript and approving it for publication was Chuan Li.

conventional ones. Exploring new DGA data interpretation methods is still of practical significance.

Fuzzy clustering techniques are integral components of artificial intelligence, being unsupervised classification algorithms with a wide range of applications in areas such as data mining and pattern recognition. Based on clustering techniques, this article proposes a new method for the interpretation of DGA data. A reference fault set is provided, and then the interpretation is implemented by calculating the membership of the DGA data to the fault set. Examples demonstrate the effectiveness and advantages of the present method, and it has a good application prospect in engineering practice. This paper is organized as follows: clustering techniques and reference fault sets are provided in Section 2. Section 3 describes an application example of the method. Then, Section 4 reports a detailed verification of the effectiveness of the proposed method by comparison with other conventional ones. At last, Section 5 makes a conclusion of the full paper.

II. THE REFERENCE FAULT SET OBTAINED WITH CLUSTERING

Fuzzy c-means clustering (FCM) is a renowned machine learning algorithm of unsuper-vised classification. The primitive principle of the clustering algorithms is that "birds of a feather flock together" and it classifies data according to their similarity. Based on the fuzzy set theory, FCM transforms the classification problem into a mathematical optimization problem with constraints. It accomplishes the fuzzy partitioning and classifying of the dataset by seeking the minimum value of the objective function. However, the similarities of the samples are measured by the reciprocal of the squared vector norm in the conventional FCM algorithm, and many local minima exist in the membership function, resulting in the complexity of the clustering spatial structure. The iterative solution of the optimization problem is essentially the "mountain climbing" method of the local search. In that approach, it is easy to fall into the local extreme points owing to the sensitivity of the initial value. Different clustering results will be obtained for different iteration starting points of the same dataset, which seriously affects the clustering accuracy. It is difficult to obtain data classification that conforms to engineering practice with conventional FCM. In our previous work [17], the influence of monotonicity of the membership function on clustering analysis is examined. Accordingly, an improved membership function for FCM is constructed. This new membership function has an exponential format, eliminates the local extremes, and has excellent monotonicity, which optimizes the clustering spatial structure. Thus, the sensitivity of conventional FCM to initial value are relieved, which are essential factors for its practical application. The clustering method is as follows [17].

The Fuzzy clustering algorithm partitions a collection of N vectors $X = \{x_1, x_2, \cdots, x_N\}$ into c fuzzy groups, such that the weighted within-groups sum of the error objective function is minimized. N is the total number of data in X. The *i-th* sample in X is a p-dimensional vector with p features or attributes, and $x_i = \{x_{i1}, x_{i2}, \cdots, x_{ip}\}$. X contains a total of c classes. The cluster center of the *j-th* cluster is $v_j = \{v_{j1}, v_{j2}, \cdots, v_{jp}\}$. The membership of the *i-th* data in the *j-th* cluster is u_{ij}. The objective function for Fuzzy clustering is defined as:

$$F = J(U, V) = \sum_{i=1}^{N}\sum_{j=1}^{c} u_{ij}^{m} e^{-\gamma d_{ij}^{2}} \tag{1}$$

where $U = \{u_{ij}\}$ denotes the membership matrix, $V = \{v_j\}$ is the cluster center set, and d_{ij} denotes the distance between x_i and v_j.

Solve $\min J(U, V)$ by the Lagrange multiplier and differentiating $J(U, V)$ to v_j (for fixed u_{ij}) and u_{ij} (for fixed v_j), respectively. The necessary condition for $\min J(U, V)$ to reach its minimum is:

$$u_{ij} = \frac{p_{ij}}{\sum\limits_{k=1}^{c} p_{ik}} = \frac{\prod\limits_{h=1}^{p} e^{-\gamma (x_{ih}-v_{jh})^2}}{\sum\limits_{k=1}^{c}\prod\limits_{h=1}^{p} e^{-\gamma (x_{ih}-v_{kh})^2}} \tag{2}$$

$$v_j = \frac{\sum\limits_{i=1}^{N} u_{ij}^{m} x_i}{\sum\limits_{i=1}^{N} u_{ij}^{m}} \tag{3}$$

The iterative process is repeating (2) and (3) until a prespecified level of accuracy is obtained.

In this paper, the clustering method is used to classify a DGA dataset $X = \{x_1, \cdots, x_{60}\}$ with known fault types, and the reference (standard) fault set is obtained. The fault diagnosis is implemented by calculating the membership of the DGA data to the reference fault set.

More than 2,000 DGA data were collected from a large number of publicly published literatures, and screened out 60 of them as DGA dataset X. The screening process is as follows.

(1) Classify data by fault types, namely, low-, middle-, and high-temperature faults, and partial, spark, and arc discharges. Each data has five attributes: the H_2, CH_4, C_2H_6, C_2H_4, and C_2H_2 content (unit: μL/L).

(2) Initialize DGA dataset X by calculating the proportion of gases related to the various faults. The H_2 content in the DGA is usually large. During the initialization process, the percentage of H_2 to H_2 and hydrocarbons, and the percentage of each hydrocarbon to the total hydrocarbons, are calculated.

(3) The mean and standard deviation of each attribute of the DGA data for each fault type are obtained; Data that differ from the mean by more than two standard deviations are eliminated. This elimination applies to five attributes of the DGA data.

(4) After multiple rounds of screening, 60 DGA data are finally left as dataset X for cluster analysis. The dataset contains six classes of 10 instances each, where each class refers to a type of transformer fault, as is shown in APPENDIX.

Classify the DGA dataset $X = \{x_1, \cdots, x_{60}\}$ with fuzzy clustering algorithm, and the parameters in (2) and (3) are $\gamma = 10$, $m = 2$, where γ is a constructed shape parameter and set by exploring and testing, and m is the fuzzy weighting exponent for FCM algorithm and set to 2 as in most literatures. The convergence accuracy is set to 10^{-5}, where the convergence of each cluster is defined by the difference in the center of each cluster for two consecutive iterations.

A large number of clustering of these 60 DGA data were performed with different random initial values. Six fault classification patterns were obtained, and five of them are considered as invalid as their cluster centers coincide or two types of faults are statistically assigned to the same data set. One classification pattern is considered valid, with the subsets and fault types must form a bijective mapping for the classification process, and the cluster centers $V = \{v_1, v_2, v_3, v_4, v_5, v_6\}$ of this valid classification pattern is considered as the reference fault (standard fault) data set, as is shown in Table 1. In the table, corresponding to the initialization process, H_2 is the percentage of H_2 to H_2 and total hydrocarbons, and hydrocarbons are the percentage of each hydrocarbon to the total hydrocarbons.

TABLE 1. The reference fault set.

	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Low-temperature overheating	0.3361	0.5038	0.1781	0.3174	0.0007
Mid-temperature overheating	0.0961	0.3675	0.2081	0.4228	0.0016
High-temperature overheating	0.1517	0.2771	0.1235	0.5938	0.0056
Partial discharge	0.9079	0.5055	0.2884	0.1915	0.0147
Spark discharge	0.4362	0.2614	0.0568	0.1430	0.5388
Arc discharge	0.4697	0.2530	0.0450	0.3206	0.3814

As shown above, Table 1 is obtained from a large amount of DGA data by screening and clustering, and they are typical chromatographic data for these six faults, namely, the proportion of dissolved gases for each standard fault. Taking the arc discharge fault as an example, its typical fault chromatographic data is:

$$v_6 = (0.4697,\ 0.2530,\ 0.0450,\ 0.3206,\ 0.3814)$$

And the proportion of gases in the arc fault are: the percentage of H_2 to H_2 and total hydrocarbons is 46.97%, and the percentage of each hydrocarbon in total hydrocarbons is 25.30%, 4.50%, 32.06% and 38.14% respectively.

III. DIAGNOSTIC STEPS AND AN APPLICATION EXAMPLE OF THIS METHOD

Interpret the DGA data to be diagnosed by calculating its membership to the reference faults in Table 2 with (2), and the membership is the extent to which the data belongs to each faults. For clarity, the calculation is split into the following steps:

Step 1: Initialize DGA data by calculating the proportion of gases, the percentage of H_2 to H_2 and hydrocarbons, and the percentage of each hydrocarbon to the total hydrocarbons, are calculated;

Step 2: Calculate the distance of each gas between DGA data and the reference fault set in Table 1, $d_{ijh} = x_{ih} - v_{jh}$;

Step 3: Calculate the sub-similarity of each gas attribute, $p_{ijk} = e^{-10d_{ijk}^2}$;

Step 4: Calculate the comprehensive similarity by multiplying the five gas sub-similarities $p_{ij} = \prod_{h=1}^{5} p_{ijh}$;

Step 5: Normalize the comprehensive similarity and obtain membership $u_{ij} = \frac{p_{ij}}{\sum_{k=1}^{6} p_{ik}}$.

TABLE 2. The gas distance between DGA data and reference fault set.

	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Low-temperature overheating	0.0686	0.384	0.1386	0.1246	0.6472
Mid-temperature overheating	0.1714	0.2477	0.1686	0.23	0.6463
High-temperature overheating	0.1158	0.1573	0.084	0.401	0.6423
Partial discharge	0.6404	0.3857	0.2489	0.0013	0.6332
Spark discharge	0.1687	0.1416	0.0173	0.0498	0.1091
Arc discharge	0.2022	0.1332	0.0055	0.1278	0.2665

TABLE 3. The sub-similarity between DGA data and reference fault set.

	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Low-temperature overheating	0.9540	0.2289	0.8252	0.8562	0.0152
Mid-temperature overheating	0.7454	0.5414	0.7526	0.5892	0.0153
High-temperature overheating	0.8745	0.7808	0.9319	0.2003	0.0162
Partial discharge	0.0166	0.2259	0.5382	1.0000	0.0181
Spark discharge	0.7523	0.8183	0.9970	0.9755	0.8878
Arc discharge	0.6644	0.8374	0.9997	0.8493	0.4915

A spark discharge fault DGA data of IEC TC 10 databases in the literature [18] is as follows: $H_2 = 305$, $CH_4 = 100$, $C_2H_6 = 33$, $C_2H_4 = 161$ and $C_2H_2 = 541$ in ppm. And the fault is '*Sparking between HV braided connection and isolated copper tube*'. Interpret this DGA data by calculating its membership to the reference faults in Table 2 using the above five steps.

1) Calculate the percentage concentrations of gases.

$H_2 = 0.2675$, $CH_4 = 0.1198$, $C_2H_6 = 0.0395$, $C_2H_4 = 0.1928$ and $C_2H_2 = 0.6479$, namely $x_i = (0.2675,\ 0.1198,\ 0.0395,\ 0.1928,\ 0.6479)$.

2) Calculate the distance of each gas between DGA data x_i and the reference fault set: $d_{ijh} = |x_{ih} - v_{jh}|$, as is shown in Table 2. E.g., the difference of H_2 between this DGA data and the low temperature overheating reference fault is: $|0.2675\text{-}0.3361| = 0.0686$.

3) Calculate the sub-similarity of the k-th gas attribute: $p_{ijh} = e^{-10d_{ijh}^2}$, as is shown in Table 3. E.g., the sub-similarity of H_2 between this DGA data and the low temperature overheating reference fault is:

$$e^{-10\times 0.0686^2} = 0.9540.$$

4) Calculate the comprehensive similarity between sample DGA data x_i and cluster center v_j by multiply the

TABLE 4. The similarity and membership between DGA data and reference fault set.

	the similarity p_{ij}	the membership u_{ij}
Low-temperature overheating	0.0023	0.0030
Mid-temperature overheating	0.0027	0.0036
High-temperature overheating	0.0021	0.0027
Partial discharge	0.0000	0.0000
Spark discharge	0.5316	0.6895
Arc discharge	0.2322	0.3012

sub-similarity: $p_{ij} = \prod_{h=1}^{5} p_{ijh}$, as is shown in Table 4. E.g., the comprehensive similarity between this DGA data and the low temperature overheating reference fault is:

$$0.9540 \times 0.2289 \times 0.8252 \times 0.8562 \times 0.0152 = 0.0023.$$

Although there are high sub-similarities of H_2, C_2H_6 and C_2H_4 between this DGA data and the low temperature overheating reference fault, the similarity of C_2H_2 is very low, only 0.0152, so the comprehensive similarity obtained after multiplication is also low, only 0.0023. And all the sub-similarities between this DGA data and the spark discharge reference fault are fair, and the comprehensive similarity is also high, i.e. 0.5316.

5) Normalize the similarity and obtain membership degree of this DGA data to each fault: $u_{ij} = \frac{p_{ij}}{\sum_{k=1}^{6} p_{ik}}$, as is shown in Table 4.

The results show that the fault of this DGA data is 68.95% spark discharge and 30.12% arc discharge. In view of the relevance of the spark and arc discharge, (also known as low- and high-energy discharge, respectively), the diagnosis result is consistent with the actual fault '*Sparking between HV braided connection and isolated copper tube*'.

In the traditional method, such as IEC Ratio Method and Duval Triangle Method, there are strict boundaries between faults. But in reality, physical boundaries only exist between thermal fault, partial discharge fault, and discharge fault. Although the thermal faults were divided into low-, mid-, and high-temperature overheating with divisions of 300 °C and 700 °C, the differences between them are not physically clear, and neither is the difference between spark and arc discharge (also known as low- and high-energy discharge, respectively). Therefore, it is more reasonable to interpret DGA data with membership degree, rather than simply being interpreted as a single one. In this way, the problem of misjudgment near the boundaries is avoided, and the diagnosis of the composite fault is also expected.

IV. COMPARISON WITH OTHER CONVENTIONAL METHODS

A. SELECTION OF COMPETITOR FOR COMPARISON

The conventional methods available to interpreted DGA data include Key Gas Method (KGM), Doernenburg Ratio Method, Rogers Ratio Method, IEC Ratio Method, Duval Triangle Method and Mansour Pentagon Method, and they have been developed and validated using large sets of data for equipments in service. Among them, the Key Gas Method charts look simple, but they are not widely accepted as reliable diagnostic tools power transformers, and studies based on an IEC data bank inspected transformers shows that only 42% of KGM diagnoses are correct [4]. Inconsistencies and low success rate for correct faults type identification have also been reported for Rogers Ratio Method [19], [20]. Faiz and Soleimani [6] and Abu-Siada and Hmood [21] investigated the accuracy of conventional methods for transformer diagnosis based on DGA data obtained from oil samples of real transformers and the comparison results shows that the David Triangle Method yields out good results among these conventional techniques. Therefore, the David Triangle Method is selected as a comparison object to verify the effectiveness of the proposed method.

IEC Ratio Method has a wide range of applications worldwide [2], and Chinese standard regards it as the preferred method for transformer fault identification [22]. So this method is also incorporated for comparison.

Mansour Pentagon Method is firstly introduced in [23] and [24] in recent years. This method was built using a pentagon shape with its heads representing the percentage concentration of each individual gas to the total combustible gases. Compared to the David triangle method (only CH_4, C_2H_4, and C_2H_2 are considered), the Mansour Pentagon takes into consideration of another two important gases, namely ethane (C_2H_6) and Hydrogen (H_2), and good diagnosis accuracy has also been claimed. It is also incorporated for comparison.

B. THE DGA DATA SOURCES

The data set used in this article has four sources. One is the IEC TC 10 Databases, which had been used for the revision of Publication 60599 [18]. This data set contains a total of 117 filed data, with the faults reliably identified by visual inspection of the equipment after the fault occurred in service. The second source [25] of this paper including 39 cases of faults simulated in the laboratory, 22 of which are thermal faults and 17 are discharge faults. Another data source [22] is an appendix to the China's power industry standards DL/T722-2014, including 17 filed cases. The last dataset [17] includes 28 field DGA, and these data were from the book—Typical Cases: Application of Grid Equipment Status Detection Technology (2011-2013) [26], published by the Operation and Maintenance Department of the Chinese State Grid Corporation, which details actual faults found after transformer disassembly. These DGA data have a total of 201, covering six faults classes.

C. DIAGNOSTIC RESULT

Interpret the above 201 DGA data with IEC Ratio Method, Mansour Pentagon Method, David Triangle Method and the method proposed in this paper respectively, and the correct

TABLE 5. The total correct rate of each method.

	IEC Ratio Method	Mansour Pentagon	David Triangle Method	The Proposed method
Number of misinterpretation	82	75	42	22
Correct rate	59.2%	62.7%	79.1%	89%

TABLE 6. Correct rate of each fault.

	Cases number	IEC Ratio Method	Mansour Pentagon	David Triangle Method	Proposed method
PD	16	6.3%	75%	81.3%	87.5%
D1	49	46.9%	65.3%	79.6%	93.9%
D2	54	90.7%	77.8%	96.3%	94.4%
T3	37	73.0%	56.8%	91.2%	89.2%
T1 and T2	45	42%	51%	66.7%	77.8%

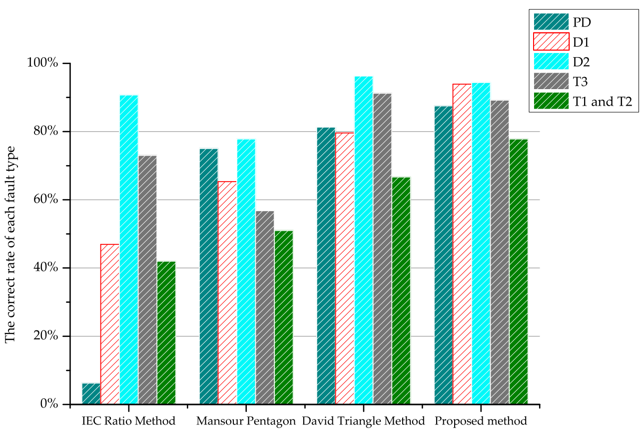

FIGURE 1. Correct rate of each fault type for each method.

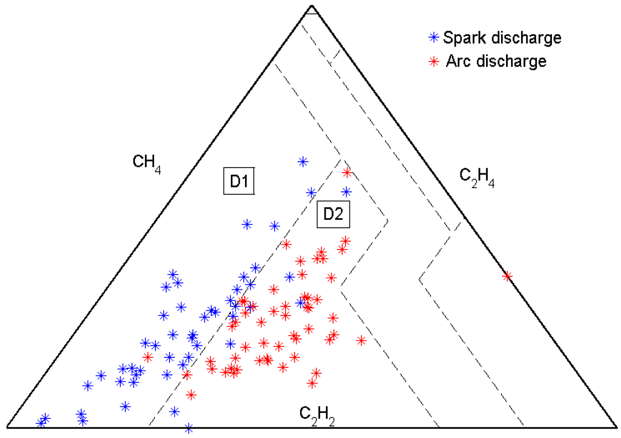

FIGURE 2. The diagnosis results for spark and arc discharge data with Duval triangle method.

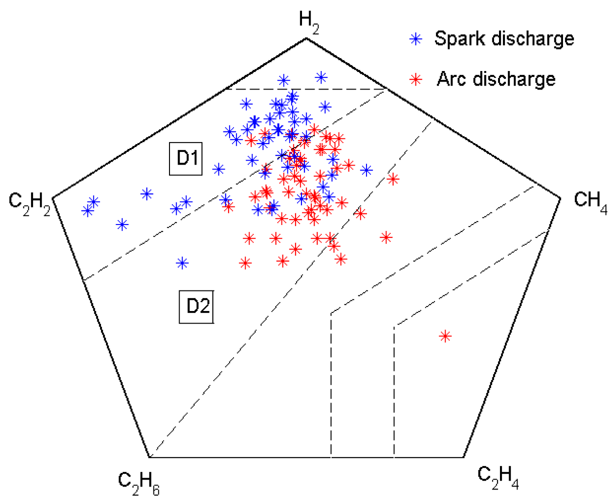

FIGURE 3. The diagnosis results for spark and arc discharge data with Mansour pentagon method.

rate of each method is shown in Table 5. The correct rate of each fault type for each method is shown in Table 6 and Figure 1. PD denotes partial discharge, D1denotes spark discharge, D2 denotes arc discharge, T1, T2, and T3 denote Low-, mid-, high-temperature overheating respectively. Because data for T1 and T2 are not separated in IEC TC 10 Databases [18], the two faults are not diagnosed separately, but denote as '*T1 and T2*'.

The test results show that the method in this paper has a high recognition accuracy rate for various faults, and outperformed the other methods.

D. DISCUSSION

The Duval Triangle Method also yields good results, but data near the boundaries are prone to misjudgment. Although spark and arc discharge, also known as low- and high-energy discharge respectively, are divided by a clear line in the Duval Triangle, the differences between them are not physically clear, and the data are overlap between the two fault type, as is shown in Figure 2. Those data that exceed the boundaries are misjudged. Strict boundaries also exist in the Mansour Pentagon, and overlap area is larger, so more data are misinterpreted, as is shown in Figure 3. This also holds for the thermal faults, whose differences are also not physically clear. There is a fuzzy transitional area between fault types, and it is more reasonable to interpret DGA data with membership degree, as is shown in the following example.

The method of this paper can identify the trend of fault development. Abnormal growth of oil chromatogram data occurred for a transformer in the appendix of [22], and the electrical test found that the grounding current of the transformer was abnormal. Returning to the factory for inspection, it was found that the core screw was in contact with the iron yoke caused by damaged insulation. This was a thermal fault caused by multi-point grounding of the core. The Oil DGA data before the transformer exits the operation are shown in Table 7. '*Delt gas*' denotes the increased gas between 2011.07.21 and 2011.07.27.

TABLE 7. DGA data for a thermal fault transformer.

Time	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
2011.07.21	362	533	91	553	10
2011.07.27	566	763	301	1299	13
Delt gas	204	230	210	746	3

TABLE 8. Membership to each reference fault.

	T1	T2	T3	PD	D1	D2
2011.07.21	0.3258	**0.3295**	0.2979	0.0002	0.005	0.0401
2011.07.27	0.1626	0.3472	**0.4576**	0.0004	0.003	0.0293
Delt gas	0.0681	0.3137	**0.5991**	0.0001	0.0013	0.0177

Interpret these two DGA data with the proposed method of this paper, and their membership to each reference fault is shown in Table 8.

The DGA data of July 21 shows that the transformer fault is 32.58% Low-temperature overheating, and 32.95% Mid-temperature overheating and 29.79% High-temperature overheating. After six days, more fault gases accumulated in the oil, and the fault characteristics became clear, and the fault is interpreted as 34.72% Mid-temperature overheating and 45.76% High-temperature overheating. Diagnostic result shifts from lower temperature fault toward a higher temperature fault. If the effect of residual gas is excluded, only the increased gas (Delt gas) is considered, then the interpretation is 31.37% Mid-temperature overheating and 59.91% High-temperature overheating. Therefore, the fault is between medium temperature overheating and high temperature overheating, and is more biased toward the high temperature side.

Mansour Pentagon Method and David Triangle Method also reveal the same fault trend in their graphs for the three DGA data, as is shown in Figure 4. The DGA data moving from a low temperature zone to a high temperature zone, and this consists with the results of the proposed method of this paper. In this way, membership describes the type and severity of the fault in more detail.

In addition, it is worth noting when using the method of this paper that if the degree of membership is almost equally divided into three types of faults, such as the case of July 21 above, this may be due to the fault feature being masked by residual gas, and it is recommended to further test the transformer and use incremental gas for analysis.

In normal operation, the insulating oil and organic insulating materials inside the oil-filled transformers generate gases under the actions of heat and electricity. Therefore, warning values are set in practice [1], including gas content and gas growth-rate warning values. Just like the David Triangle Method, IEC Ratio Method and other conventional DGA interpretation method, fault diagnostics with the proposed method is enabled only when the gas properties exceed the warning values and the gas content continues to increase, indicating that the transformer does have a fault.

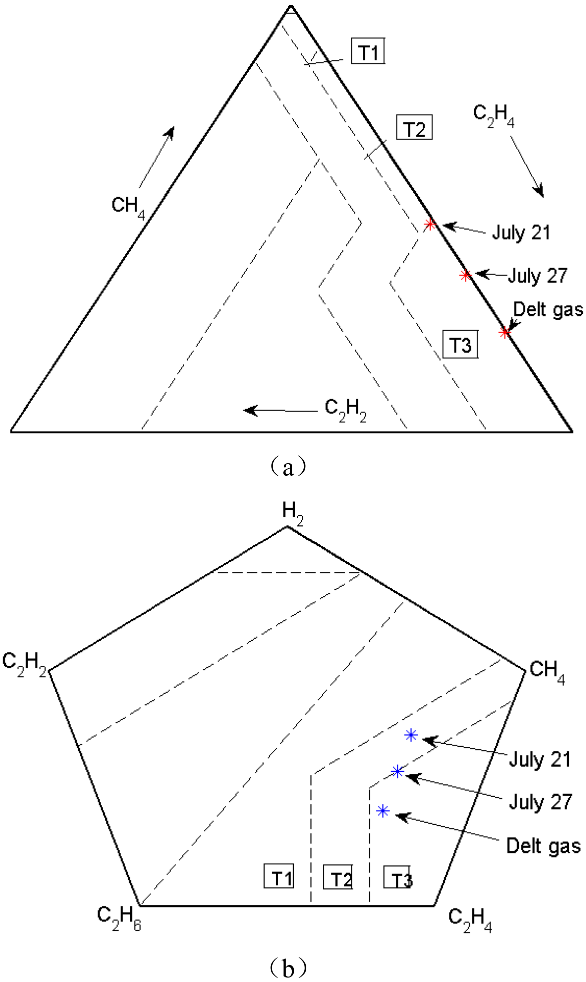

FIGURE 4. Distribution of three DGA data in (a) Duval triangle and (b) Mansour pentagon.

V. CONCLUSIONS

Interpretation of Oil chromatography data is the key step of dissolved gas analysis in transformer. The conventional DGA interpretation methods are prone to misinterpreting the gas data near the ratio boundaries, and the fault diagnosis accuracy rate of these methods is relatively low. Intelligent approaches are also available for transformer fault diagnosis, but they are usually too complicated to be implemented practically on a wide range. This paper proposes a new DGA interpretation method for oil impregnated transformers. A reference fault set is obtained from a large amount of DGA data by statistical screening and clustering, and the interpretation is implemented by calculating the membership of the DGA data to the fault set. Test with credible DGA dataset (201 field cases) shows that the correct rate of the new method is 89%, while David triangle method is 79% and IEC Ratio method is 59%, which demonstrate the superiority of the proposed method to the conventional ones. This method overcomes the defect that the conventional DGA analysis method is easy to misjudgment near the boundaries, and can provide detailed information on the type and severity of

TABLE 9. DGA data used to obtain the reference fault set of Table 1.

NUM	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2	Fault type
1	33	29	9	12	0	T1
2	46	98	26.3	41.3	0	T1
3	60	60	16	40	0.3	T1
4	87.2	73.18	27.14	56.88	0	T1
5	97	110	34	85	0	T1
6	110.4	112	32.5	80.8	0	T1
7	29.9	24.1	34.3	92.5	0.6	T1
8	120	120	33	84	0.55	T1
9	143.2	123	38	75	0	T1
10	181	162	70	132	0	T1
11	20	41.9	20.2	44.2	0.38	T2
12	20.37	59.79	45.24	80.49	0	T2
13	23.51	61.33	45.21	98.03	1.01	T2
14	46.9	161.6	94.1	193.6	0.56	T2
15	47	120	90	198	3	T2
16	72	442	221	461	0.7	T2
17	110.6	458.8	242.6	406.4	0	T2
18	128	419	269.5	614.1	0.35	T2
19	24	34.6	14.2	21.7	0	T2
20	613	3240	1432	2788	0	T2
21	35.1	50.6	16.1	93	1.1	T3
22	68	99.2	35.9	202.9	0	T3
23	156	240	54	399	0.98	T3
24	63	149.6	57.5	276	0	T3
25	30	25.5	31.5	93	1.8	T3
26	165.62	240.95	61.32	514.53	13.53	T3
27	164	244	103	497	8.3	T3
28	135.65	278.53	58.86	492	2.95	T3
29	135.88	362.42	125.22	826.65	3.74	T3
30	236	410.2	159	817.3	3.5	T3
31	1198	3.2	1.4	3.2	0.5	PD
32	1309	124	113	6	0	PD
33	2587	7.88	4.7	1.4	0	PD
34	85.87	7.01	4.49	2.64	0	PD
35	102	108	70	41	0	PD
36	625	49	9	7	0.6	PD
37	195.9	14.5	11.6	2.4	0	PD
38	420	37.3	14.9	30	0.2	PD
39	485	35	29	6	0	PD
40	83.26	45.32	18.1	36.45	0.26	PD
41	4.1	3.5	0.68	1.2	5.2	D1
42	9	3.9	0.8	4	13	D1
43	1198	3.2	1.4	3.2	0.5	D1
44	14.2	4	1.4	1.5	9.51	D1
45	30.1	17.1	2.2	5.5	30.1	D1
46	45	11	2.7	12.74	28.5	D1
47	65.2	20	3.9	8.13	25.1	D1
48	67.8	8.89	1.88	12.67	36.2	D1
49	101.72	27.65	7.13	16.92	53.87	D1
50	549	121.3	25.5	31.9	198.5	D1
51	56	10	1.3	13.5	17.6	D2
52	57	15	3.1	23	25.3	D2
53	65	26.1	10.1	41.6	57.8	D2
54	75.5	30.2	2.33	30.3	18.2	D2
55	145.88	40.65	9.37	34.02	59.71	D2
56	195.7	58	16.4	91.6	96.9	D2
57	1027	185	17	271	399	D2
58	475.3	195.8	32.6	187.3	221.2	D2
59	531	111.9	22.7	122.5	169	D2
60	755	229	32	404	460	D2

PD denotes partial discharge, D1 denotes spark discharge, D2 denotes arc discharge, T1, T2, and T3 denote Low-, mid-, high-temperature overheating respectively.

the fault. Besides, this method is simple and practical, not as complicated as SVM, ANN and other intelligent algorithms, which indicates a good application prospect in engineering practice.

APPENDIX

See Table 9.

REFERENCES

[1] *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*, IEEE Standard C57.104, New York, NY, USA, 2009.

[2] *British Standards Institute Staff. Mineral Oil-Impregnated Electrical Equipment in Service-Guide to the Interpretation of Dissolved and Free Gases Analysis*, IEC Standard 60599, Geneva, Switzerland, 2015.

[3] *Standard Test Method for Analysis of Gases Dissolved in Electrical Insulating Oil by Gas Chromatography*, ASTM Standard D3612-02, ASTM, Pennsylvania, PA, USA, 2017.

[4] N. Bakar, A. Abu-Siada, and S. Islam, "A review of dissolved gas analysis measurement and interpretation techniques," *IEEE Elect. Insul. Mag.*, vol. 30, no. 3, pp. 39–49, May/Jun. 2014.

[5] H. C. Sun, Y. C. Huang, and C. M. Huang, "A review of dissolved gas analysis in power transformers," *Energy Procedia*, vol. 14, pp. 1220–1225, Mar. 2012.

[6] J. Faiz and M. Soleimani, "Dissolved gas analysis evaluation in electric power transformers using conventional methods: A review," *IEEE Trans. dielectrics Electr. Insul.*, vol. 24, no. 2, pp. 1239–1248, Apr. 2017.

[7] L. F. Cheng and T. Yu, "Dissolved gas analysis principle-based intelligent approaches to fault diagnosis and decision making for large oil-immersed power transformers: A survey," *Energies*, vol. 11, no. 4, pp. 913–982, 2018.

[8] M. E. A. Senoussaoui, M. Brahami, and I. Fofana, "Combining and comparing various machine-learning algorithms to improve dissolved gas analysis interpretation," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 15, pp. 3673–3679, 2018.

[9] A. Velasquez, M. Ricardo, M. Lara, and V. Jennifer, "Principal components analysis and adaptive decision system based on fuzzy logic for power transformer," *Fuzzy Inf. Eng.*, vol. 9, no. 4, pp. 493–514, 2017.

[10] N. Masoud, E. Reza, and H. Payman, "Using dissolved gas analysis results to detect and isolate the internal faults of power transformers by applying a fuzzy logic method," *IET Gener. Transmiss., Distrib.*, vol. 11, no. 10, pp. 2721–2729, 2017.

[11] H. Wei, Y. Wang, L. Yang, C. Yan, Y. Zhang, and R. Liao, "A new support vector machine model based on improved imperialist competitive algorithm for fault diagnosis of oil-immersed transformers," *J. Elect. Eng. Technol.*, vol. 12, no. 2, pp. 830–839, 2017.

[12] J. Z. Li, Q. G. Zhang, K. Wang, and J. Y. Wang, "Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 23, no. 2, pp. 1198–1206, Apr. 2016.

[13] E. M. Danish, K. S. Ahmad, and I. Tarikul, " Transformer incipient fault diagnosis on the basis of energy-weighted DGA using an artificial neural network," *Turkish J. Elect. Eng. Comput. Sci.*, vol. 26, no. 1, pp. 77–88, 2018.

[14] T. Kari, W. Gao, and D. Zhao, "Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm," *IET Gener. Transmiss. Distrib.*, vol. 12, no. 21, pp. 5672–5680, 2018.

[15] J. Watada, S. Chen, and Y. Yabuuchi, "A rough set approach to data imputation and its application to a dissolved gas analysis dataset," in *Proc. IEEE Int. Conf. Comput. Meas. Control Sensor Netw.*, May 2017, pp. 24–27.

[16] D. Jiejie, S. Hui, and S. Gehao, "Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 5, pp. 2828–2835, Oct. 2017.

[17] E. Li and L. Wang, "Improved fuzzy C-means clustering for transformer fault diagnosis using dissolved gas analysis data," *Energies*, vols. 9–11, no. 9, pp. 2344–2361, 2018.

[18] M. Duval and A. de Pabla, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Elect. Insul. Mag.*, vol. 17, no. 2, pp. 31–41, Mar. 2001.

[19] F. Jawad and S. Milad, "Assessment of computational intelligence and conventional dissolved gas analysis methods for transformer fault diagnosis," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 25, no. 5, pp. 1798–1806, Oct. 2018.

[20] A. Abu-Siada and S. Islam, "A new approach to identify power transformer criticality and asset management decision based on dissolved gas-in-oil analysis," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 19, no. 3, pp. 1007–1012, Jun. 2012.

[21] A. Abu-Siada and S. Hmood, "A new fuzzy logic approach for consistent interpretation of dissolved gas-in-oil analysis," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 20, no. 6, pp. 2343–2349, Dec. 2013.

[22] *National Energy Administration of PRC, Guide to the Analysis and the Diagnosis of Gases Dissolved in Transformer Oil*, Standard DL/T722-2014, Beijing, China, 2014.

[23] D. A. Mansour, "A new graphical technique for the interpretation of dissolved gas analysis in power transformers," in *Proc. IEEE Conf. Electr. Insul. Dielectr. Phenomena*, Oct. 2012, pp. 195–198.

[24] D. A. Mansour, "Development of a new graphical technique for dissolved gas analysis in power transformers based on the five combustible gases," *IEEE Trans. Dielectr. Electr. Insul*, vol. 22, no. 5, pp. 2507–2512, Oct. 2015.

[25] M. Duval, "A review of faults detectable by gas-in-oil analysis in transformers," *IEEE Elect. Insul. Mag.*, vol. 18, no. 3, pp. 8–17, May 2002.

[26] State Grid Corporation Operation and Maintenance Department, *Typical Cases: Application of Grid Equipment Status Detection Technology*. Beijing, China: China Electric Power Press, 2014, pp. 76–191.

ENWEN LI received the bachelor's degree in packaging engineering from Wuhan University, in 2013, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include transformer fault diagnosis and maintenance strategies.

LINONG WANG received the B.E. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1998, and the M.E. and Ph.D. degrees in high voltage and insulation technology from Wuhan University, Wuhan, China, in 2006 and 2011, respectively, where he is currently a Professor with the School of Electrical Engineering. His research interests include transmission lines operation and maintenance technology, live working technology, and electrical engineering material.

BIN SONG was born in Tianmen, China, in 1970. He received the bachelor's and master's degrees from the Thermal Energy and Power Engineering School, Wuhan University, in 1995 and 2000, respectively, and the Ph.D. degree from the Electric Engineering School, Wuhan University, in 2004. He is currently an Associate Professor of electrical engineering with Wuhan University. His research interests include condition assessment, maintenance strategies, fault diagnosis, and information data process for electrical equipment.