Electrical Power and Energy Systems 105 (2019) 846–855

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

A novel approach to transformer fault diagnosis using IDM and naive credal classifier

B. Zhao[a], M. Yang[a,*], H.R. Diao[b], B. An[c], Y.C. Zhao[a], Y.M. Zhang[a]

[a] Key Laboratory of Power System Intelligent Dispatch and Control (Shandong University), Ministry of Education, Jinan 250061, China
[b] CSG Power Dispatching Control Center, Guangzhou 510663, China
[c] State Grid Shandong Electric Power Maintenance Company, Jinan 250118, China

ARTICLE INFO

Keywords:
Fault diagnosis
Imprecise Dirichlet model
Imprecise probability
Naive credal classifier
Transformer fault

ABSTRACT

Transformer fault diagnosis is important for improving the operational reliability of power systems. Despite great efforts to enhance the accuracy of fault diagnosis, the precise detection of transformer faults remains difficult. The major barriers are attributed to the lack of transformer fault records and related data, which may produce unconvincing diagnostic results, and to the fact that multiple latent faults in a transformer are difficult to detect distinctly. In this paper, an imprecise probability-based approach to transformer fault diagnosis by using the imprecise Dirichlet model and naive credal classifier is proposed. Instead of providing a single-valued probability, the approach calculates the interval probability for each possible type of transformer faults to explicitly indicate the uncertainty of the diagnosis caused by the insufficiency of historical fault records. Meanwhile, according to whether an overlap exists among the probability intervals of different types of faults, the proposed approach can output either one explicit fault or multiple probable latent faults that occurred in the transformer. In the proposed approach, the imprecise Dirichlet model (IDM) is used to evaluate the imprecise probabilistic relationships between each type of transformer fault and its corresponding symptom attributes. A naive credal classifier (NCC) is established to integrate the IDM estimation results and calculate the interval-valued probabilities of all types of possible faults. Then, the transformer fault diagnosis results can be obtained by the classification criteria of the NCC. The approach can compensate for the defects of conventional methods, including their inability to perform reliable diagnosis with insufficient fault samples or distinctly indicate the multiple latent transformer faults. The accuracy of diagnosis and the efficiency of transformer maintenance can be significantly improved by implementing the proposed approach. The effectiveness of the approach is verified through case studies.

1. Introduction

Power transformers are important components of power systems, and damage to or failure of transformers may cause significant economic and social losses. Therefore, performing efficient and reasonable transformer fault diagnosis is vital for the safe and reliable operation of power systems.

An in-service transformer may be subjected to various harmful operating conditions that can break down the insulating materials and release gaseous decomposition products dissolved in the oil. In recent decades, dissolved gas analysis (DGA) and its improved approaches, e.g., the Rogers Ratios, modified Rogers Ratios, and IEC Ratios, have been widely recognized as convenient diagnostic techniques for transformer faults [1]. The ratio analysis of specific dissolved gas concentrations in the insulation oil of a transformer provides knowledge regarding a fault state, and thus allows performing necessary maintenance to keep the normal operation of the transformer [2]. However, conventional DGA interpretation methods continue to have disadvantages despite their extensive applications. For instance, the potential for misdiagnosis is high when the measured and calculated gas ratios are near the specified boundaries. In some cases, the DGA-based methods are not always efficient because the physical interpretations of some ratio coding combinations and the explicit diagnosis result under a specific code value cannot be reasonably obtained [3].

To address these restrictions, artificial intelligence (AI) techniques, such as fuzzy logic inference systems [4,5], nuero fuzzy inference system [6], artificial neural networks (ANNs) [7,8] and support vector machines (SVMs) [9,10], have been developed to detect different types of faults combined with DGA. AI-based methods can establish the complex and nonlinear relationships between DGA data and transformer faults [10].

* Corresponding author.
E-mail address: myang@sdu.edu.cn (M. Yang).

https://doi.org/10.1016/j.ijepes.2018.09.029
Received 22 May 2018; Received in revised form 8 August 2018; Accepted 16 September 2018
0142-0615/

However, the majority of the previously mentioned methods have inherent drawbacks. For instance, the accurate rate of fuzzy logic-based diagnosis methods is poor, as their inference rules and fuzzy membership functions are determined by subjective experience [11]. ANNs often encounter problems of over-fitting and local minimum [12]. Parameters in SVMs, e.g., penalty factors and kernel parameters, should be continuously optimized during the training process to ensure the classification performances. Notably, all DGA-based methods, either conventional methods or AI-based methods, encounter a common and inherent defect: DGA is not suitable for detecting precise electrical or mechanical faults because these types of faults affect the dissolved oil in an indirect manner [13]. Therefore, recent researches have proposed comprehensive diagnosis methods using chromatography data and electrical test data [14,15]. Compared with the fault diagnosis methods that are only based on DGA, techniques that combine DGA with chromatography or electrical test data can yield significantly better diagnostic effects [13].

However, existing methods often fail to diagnose real faults. The major barriers are attributed to the following two aspects. One is the lack of fault samples and transformer test data. With the development of smart grid and the energy internet, a large number of new components have been employed. As their outages or failures are small probability events, their fault documents and sample records are scarce; therefore, obtaining a convincing fault diagnosis result according to the limited data is difficult. On the other hand, the precise detection of multiple latent faults that occurred in a transformer is usually difficult. Although the gases detected for diagnosis may be generated by a single fault or multiple faults, conventional diagnosis methods usually assume a single major fault and do not consider the possibility of multiple simultaneous faults [16]. Based on these obstacles, an efficient diagnosis approach to detecting transformer faults considering limited fault samples and multiple latent faults has not been reported, which leaves a large margin for improvements for transformer fault diagnosis techniques.

In this paper, a novel approach to transformer fault diagnosis based on imprecise probability theory, imprecise Dirichlet model and naive credal classifier is proposed. Instead of providing the single-valued probability of a certain type of transformer fault, the proposed approach provides the possible interval probability of each type of fault considering the lack of available historical samples and can objectively reflect the uncertainty of transformer fault diagnosis. The imprecise Dirichlet model (IDM) is adopted to estimate the imprecise probabilistic relationships between the transformer faults and their corresponding symptom attributes. Then, a naive credal classifier (NCC) that can perform imprecise probabilistic classification is established to obtain the interval-valued probabilities of all possible transformer faults. Depending on whether an overlap exists among the interval probabilities of the probable faults, the NCC can identify a single fault or multiple latent faults as the diagnosis result. The proposed approach can objectively estimate the probability of a transformer fault and can compensate for the disadvantage that the conventional diagnosis methods usually disregard the occurrence of multiple latent faults. Thus, the proposed approach can provide more reliable diagnosis results, especially for the cases with limited historical samples.

The remainder of this paper is organized as follows. In Section 2, some basic concepts concerning the proposed approach, including imprecise probability, IDM, credal classifier and its classification criteria, are introduced. Section 3 compares the effectiveness and accuracy of NCC with those of other classifiers through several experiments. Details of the transformer fault diagnosis model are provided in Section 4. Section 5 presents the case studies and corresponding results. Finally, the conclusions are drawn in Section 6.

2. Basic concepts

2.1. Imprecise probability

Imprecise probability theory is a generalization of classical probability theory that enables partial probability specifications when available information is scarce. The theory bloomed in the 1990s owning to the comprehensive foundations proposed by Walley [17].

In imprecise probability theory, the occurrence probabilities of a random event B and its complementary event B^c are represented by interval-valued probabilities, which can be expressed as

$$P_{\text{im}}(B) = [\underline{P}(B), \overline{P}(B)], \tag{1a}$$

$$P_{\text{im}}(B^c) = [\underline{P}(B^c), \overline{P}(B^c)]. \tag{1b}$$

where $\underline{P}(B)$, $\overline{P}(B)$, $\underline{P}(B^c)$ and $\overline{P}(B^c)$ are the lower and upper bounds of the imprecise probabilities $P_{\text{im}}(B)$ and $P_{\text{im}}(B^c)$, respectively, which satisfy $0 \leqslant \underline{P}(B) \leqslant \overline{P}(B) \leqslant 1$ and $0 \leqslant \underline{P}(B^c) \leqslant \overline{P}(B^c) \leqslant 1$. Meanwhile, according to the property of complementary events, i.e., $P(B) + P(B^c) = 1$, the constraints $\underline{P}(B) = 1-\overline{P}(B^c)$ and $\overline{P}(B) = 1-\underline{P}(B^c)$ should also be satisfied.

When no statistical information is available, the occurrence possibilities of the events will have maximal probability intervals, i.e., $\underline{P}(B) = \underline{P}(B^c) = 0$ and $\overline{P}(B) = \overline{P}(B^c) = 1$. Conversely, if sufficient statistical information is available, the probability interval may shrink to a single point and a precise probability will be obtained [18].

2.2. Imprecise Dirichlet model (IDM)

The imprecise Dirichlet model is an effective method for estimating the imprecise probability. Essentially, the IDM is an extension of the deterministic Dirichlet model [19]. Assume a multinomial distribution that has N types of outcomes. To estimate the occurrence probability of each outcome, the deterministic Dirichlet model employs a single Dirichlet distribution as the prior distribution, and its prior probability density function (PDF) can be represented as

$$f(\theta) = \Gamma\left(\sum_{n=1}^{N} \alpha_n\right)\left[\prod_{n=1}^{N} \Gamma(\alpha_n)\right]^{-1} \prod_{n=1}^{N} \theta_n^{\alpha_n - 1}, \tag{2}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_N)$ refers to the occurrence probabilities of the outcomes, satisfying $\theta_n \geq 0$, $n = 1, 2, \ldots, N$ and $\sum_{n=1}^{N} \theta_n = 1$; $\alpha_1, \alpha_2, \ldots, \alpha_N$ represent the positive parameters of the Dirichlet distribution and Γ represents the gamma function.

According to Bayesian theory, since the Dirichlet distribution is a conjugate prior of the multinomial distribution [20], the posterior of $\boldsymbol{\theta}$ with respect to the newly obtaining observations $\boldsymbol{M} = (m_1, m_2, \ldots, m_N)$ also belongs to a Dirichlet distribution, which is called the updating process, and the corresponding posterior PDF can be represented as

$$f(\theta \mid \boldsymbol{M}) = \Gamma\left(\sum_{n=1}^{N} (m_n + \alpha_n)\right)\left[\prod_{n=1}^{N} \Gamma(m_n + \alpha_n)\right]^{-1} \prod_{n=1}^{N} \theta_n^{m_n + \alpha_n - 1}, \tag{3}$$

where m_n, $n = 1, 2, \ldots, N$, represents the number of times that the nth outcome is observed.

Therefore, the parameter θ_n of the multinomial distribution can be estimated by calculating the mathematical expectation of the posterior distribution, as

$$\theta_n = E(\theta_n \mid \boldsymbol{M}) = (m_n + \alpha_n)\Big/\left(\sum_{n=1}^{N} (m_n + \alpha_n)\right), \quad n = 1, 2, \ldots, N. \tag{4}$$

When analysing the estimation results of the deterministic Dirichlet model, if no observations are available, the probability θ_n of the nth outcome is determined by the parameter α_n, i.e., $\theta_n = \alpha_n / \sum_{n=1}^{N} \alpha_n$, where the parameter α_n is called the prior weight of the outcome, and $\sum_{n=1}^{N} \alpha_n$ is often denoted as the parameter s, which is known as the equivalent sample size in the Dirichlet distribution. In the probabilistic estimation process, s implies the influence of the prior distribution on posterior probabilities [21], i.e., a larger value of s indicates that more observations are needed to tune the parameters assigned by the prior [22].

As shown in Eq. (4), a disadvantage of using the deterministic Dirichlet model is that the estimation result will be significantly