

基于 DGA 的粗糙集与决策信息融合变压器故障诊断

李春茂, 周妹末, 袁海满, 高 波, 吴广宁

(西南交通大学电气工程学院, 四川 成都 610031)

摘要: 针对变压器故障特征信息不确定性、冗余性及传统故障诊断手段的单一性问题, 本文构建了一种粗糙集与多决策信息融合的变压器故障诊断模型。该方法首先考虑将 16 种特征气体比值作为故障特征参量, 并利用离散化规则与粗糙集知识约简对其进行知识提取, 以有效降低特征信息冗余度。其次, 将降维后属性集作为 BP 神经网络、支持向量机以及贝叶斯网络 3 种单一诊断方法的特征输入, 进行故障类型初步判定。最后, 利用 DS 信息融合规则对 3 种初步判定结果进行决策融合, 以获得更为高效的故障判断结论。实例分析表明, 该方法有效削弱了冗余特征信息对诊断结果的影响, 能够合理解决证据融合冲突, 并切实提高了故障识别准确率, 其性能明显优于单一诊断方法。

关键词: 变压器; DS 理论; 属性约简; 离散化; 粗糙集理论; 信息融合; 故障诊断

DOI: 10.12067/ATEEE1611080

文章编号: 1003-3076(2018)01-0084-07

中图分类号: TM411

1 引言

随着电力网络电压等级的不断提升, 变压器作为电力系统的枢纽设备, 其安全可靠的运行是电网稳定供电的前提。然而, 变压器在长期运行中, 难免会因各种内外因素致使内绝缘结构破坏, 从而导致异常情况乃至事故的发生。因此, 随着状态维修^[1,2]策略重要性的不断提升, 及时准确判定变压器异常状况, 并制定相应状态检修计划, 对于整个系统而言意义重大。

近年来, 研究者们相继提出人工神经网络(ANN)^[3,4]、支持向量机(SVM)^[5,6]、贝叶斯(Bayesian)网络^[7]、关联分析^[8-10]、粗糙集理论^[11-13]以及证据理论^[14-16]等多种基于数据挖掘与智能信息处理的变压器故障诊断方法, 在实际应用中取得了较好成效。文献[3-5, 7]以油中溶解气体分析(DGA)作为特征输入, 分别利用 BP 神经网络(BPNN)、SVM 及多 Bayesian 网络相结合建立故障诊断模型; 文献[8, 9]运用关联分析确立特征信息与故障类型的关系, 并结合智能方法进行故障诊断。但变压器作为

一个复杂的综合体, 其特性信息繁多而复杂, 含有较多不确定因素, 与故障类别间存在较大模糊性^[17], 仅仅依靠单一诊断难以获得更准确的诊断结果。为此, 文献[6, 11]利用粒子群、人工鱼群算法对单一智能方法的参数进行优化, 以改进单一算法不足之处; 文献[14, 15]采用信息融合的方式将多种智能方法的初步诊断结果进行决策融合, 使得诊断准确度得到较大提升。虽然这些研究有效解决了单一方法结论的片面性, 但冗余特征量依旧会使算法训练时间加长从而导致复杂度增加, 影响诊断结果的准确性。

鉴于此, 本文基于智能互补思想, 将粗糙集理论与证据理论相结合, 构建了一种粗糙集与故障决策信息融合的故障诊断方法, 以解决现有诊断方法的不足。粗糙集理论能克服特征参量的不完备性并降低故障特征参量维度, 以实现特征信息的有效提取与充分挖掘。同时, 本文利用 BPNN、SVM 以及 Bayesian 网络 3 种不同诊断方法进行故障状态初步判定, 并根据证据理论中的证据合成规则对初步结论进行决策融合, 以获得更为准确

收稿日期: 2016-11-30

作者简介: 李春茂 (1963-), 男, 四川籍, 教授, 博士, 研究方向为高压电气设备在线监测与状态评估;
周妹末 (1993-), 女, 四川籍, 硕士研究生, 研究方向为电力变压器状态监测与故障诊断。

可靠的判定结果。

2 粗糙集理论

粗糙集理论是一种处理模糊性和不精确性问题的新型数学工具,能够在保留关键信息的前提下对知识进行处理,并求得知识的最小表达^[18]。

四元组 $S = (U, A, V, f)$ 是一个信息系统,或数据库系统,其中 $U = \{x_1, x_2, \dots, x_n\}$, 为非空的对象集,也称为论域; $A = C \cup D$ 表示属性的非空有限集合, $C = \{c_1, c_2, \dots, c_m\}$ 为条件属性集合, $D = \{d_1, d_2, \dots, d_n\}$ 表示决策属性集合,且 $C \cap D$ 为空集; V 为属性值域。 $f: U \times A \rightarrow V$ 是 U 和 A 的关系集,也称信息函数集,它指定了 U 中每一个对象的属性值。在故障诊断中,故障特征信息为条件属性,故障类型为决策属性,两者间所建立的联系关系为信息函数,即决策表。

2.1 属性离散化

粗糙集理论都是针对离散化数据,而原始决策表通常表现为连续系统,因此,需对其进行离散化。数据离散化的最优准则是:离散化后属性值的信息丢失要尽量少;离散化后属性的空间维数要小,即离散化后的种类要少^[12]。

为此,本文一部分属性利用已有先验知识直接离散化,另一部分则采用等频划分法 (Equal Frequency Intervals)。等频划分法将属性值从小到大进行排列后,根据给定参数 m 将其平均划分为 m 段,并得到各属性断点集,要求相邻断点间属性个数相同。最后,依据断点对连续属性进行离散编码。

2.2 知识约简

定义 1: 对数据表 $S = (U, A, V, f)$, $B \subseteq C$, 若 $U/R_B \neq U/R_{B \setminus \{b\}}$, 称 $b \in B$ 在 B 中是必要的; 否则称 $b \in B$ 在 B 中是不必要的。若 $B \subseteq C$ 中每个属性在 B 中都是必要的, 称 B 是独立的。

定义 2: 对数据表 $S = (U, A, V, f)$, 称 C 中所有必要属性构成的集合为 C 的核, 记为 $\text{Core}(C)$ 。若 $U/R_B = U/R_C$, 且 B 是独立的, 称 $B \subseteq C$ 是一个约简, 记为 $\text{red}(C)$ 。约简与核的关系如下:

$$\text{Core}(C) = \bigcap \text{red}(C) \quad (1)$$

知识约简是粗糙集理论核心内容之一,要求在不丢失信息的前提下,获得与原信息系统表达同样知识的最小属性集^[11]。由此,针对特征信息降维,知识约简能取得较好成效。其约简步骤如下:

(1) 删除决策表中某条条件属性列,若剩余规

则发生矛盾,则此条属性不能删除。

(2) 计算条件属性核值,只保留必要属性。

(3) 简化规则,获得最小决策表与降维后特征量。

3 证据理论

3.1 DS 证据理论合成规则

DS 证据理论作为多源信息融合方法,不需要先验概率和条件概率,能表示、测量和组合证据中不确定性并对冲突信息进行有效融合^[19],因此,被广泛应用于目标识别、故障诊断等领域。其中,相关定义^[20]如下:

定义 1: 设定 U 是一辨识框架,若集函数 $m: 2^U \rightarrow [0, 1]$ (2^U 为 U 的幂集) 满足:

$$\begin{cases} m(\phi) = 0 \\ \sum_{A \subseteq U} m(A) = 1 \end{cases} \quad (2)$$

则称 m 为 U 上的基本概率赋值 (Basic Probability Assignment, BPA), 也即对命题的信任度。 $m(A)$ 称作 A 的基本信任函数 (或 mass 值)。

定义 2: 设定 U 是一辨识框架, $\forall A \subseteq U, m(A) > 0$, 则将 A 称作焦元。

定义 3: 设定同一辨识框架 U 上的 n 组 BPA 分别为 m_1, m_2, \dots, m_n , 相应焦元分别为 $A_i (i = 1, 2, \dots, N)$, 则 Dempster 组合规则为:

$$m(A) = \begin{cases} 0 & A = \phi \\ \frac{\sum_{\bigcap A_i = A} \prod_{1 \leq i \leq N} m_i(A_i)}{1 - k} & A \neq \phi \end{cases} \quad (3)$$

式中

$$k = \sum_{\bigcap A_i = \phi} \prod_{1 \leq i \leq N} m_i(A_i) \quad k \neq 1 \quad (4)$$

除组合规则外,决策方式是 DS 理论融合信息时用于最终决策判断的另一关键。基于信任函数的决策、基于最小风险的决策以及基于概率赋值的决策是 DS 理论中常用的 3 种决策方式。基于信任函数的判决是一种软判决,一般是基于组合后的 m 求出信任函数,该函数即判决结果;基于最小风险的判决是利用风险函数与基本概率赋值函数做出最优决策。为减小决策融合工作量,本文采用基于基本概率赋值的决策方式,决策判断式如下:

定义 4: 设 $A_1, A_2 \subset U$, 满足:

$$\begin{cases} m(A_1) = \max \{m(A_i), A_i \subset U\} \\ m(A_2) = \max \{m(A_i), A_i \subset U \text{ 且 } A_i \neq A_1\} \end{cases} \quad (5)$$

若有

$$\begin{cases} m(A_1) - m(A_2) > \varepsilon_1 \\ m(U) < \varepsilon_2 \\ m(A_1) > m(U) \end{cases} \quad (6)$$

式中, $\varepsilon_1, \varepsilon_2$ 为预选设定门限。

则 A_1 为判决结果。

3.2 证据体构造

进行信息融合前,需要将不同信源按 3.1 节要求转换为辨识框架下的 BPA,即证据体构造。有研究表明^[15,21],BPNN、SVM 以及 Bayesian 网络在构造证据体时,不仅诊断可靠性相比其他智能算法高,它们的诊断输出也更贴近证据体构造要求。因此,利用其诊断结果构造融合需要的多个证据体。证据体构造时,令 $q_i(k)$ 为第 i 种算法的第 k 个输出值; β_i 为单一智能诊断方法的诊断准确率; $m(\theta)$ 为此种方式下的诊断不确定度^[15]。最后,第 i 种智能诊断算法赋予其第 k 个诊断输出的基本概率赋值与不确定度为:

$$\begin{cases} m_i(k) = \frac{q_i(k)\beta_i}{\sum_{k=1}^p q_i(k)} & i = 1,2,3 \quad k = 1,2,\cdots,7 \\ m_i(\theta) = 1 - \beta_i \end{cases} \quad (7)$$

4 基于粗糙集理论与多信息融合的证据体构造

本文建立了一种基于粗糙集与多信息融合的变压器故障诊断模型,诊断流程图如图 1 所示。其中粗糙集约简与信息决策融合是本模块的核心。

4.1 初始决策表建立

DGA 技术由于能准确、可靠地发现变压器内部潜伏性故障,成为目前变压器故障诊断中应用较广的方法之一^[16]。DGA 方法以检测 H_2 、 CH_4 、 C_2H_2 、 C_2H_4 、 C_2H_6 五种特征气体含量为主,但由于变压器内部机理复杂,且故障特征量与故障类型间映射模糊,使得利用单一特征气体含量作为特征参量的故障诊断方式可靠性不高。大量研究表明,特征气体比值与故障联系更紧密^[22]。因此,为进一步提高故障诊断准确率,揭示特征气体比值与故障类关系,本文根据三比值法、Rogers 四比值法、Doernenburg 法以及无编码比值法等^[23]广泛应用的比值法方法总结出 CH_4/H_2 、 CH_4/C_2H_4 、 C_2H_2/H_2 、 C_2H_2/CH_4 、 C_2H_2/C_2H_4 、 C_2H_2/C_2H_6 、 C_2H_4/H_2 、 C_2H_4/C_2H_6 、

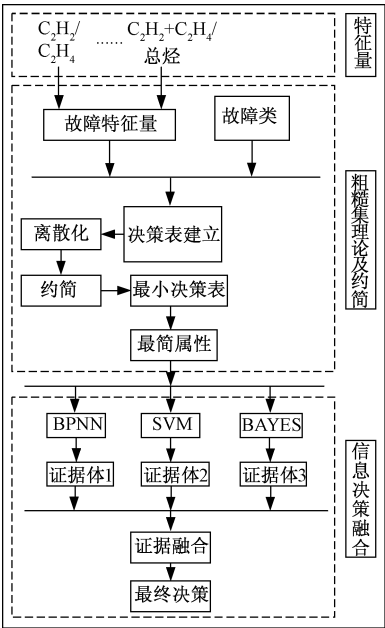


图 1 故障模型
Fig.1 Failure model

C_2H_6/H_2 、 C_2H_6/CH_4 、 $H_2/(总烃 + H_2)$ 、 $CH_4/总烃$ 、 $C_2H_2/总烃$ 、 $C_2H_4/总烃$ 、 $C_2H_6/总烃$ 以及 $(CH_4 + C_2H_4)/总烃$ 16 种气体比值作为初始特征参量,即决策表条件属性,编号依次为 S_1, S_2, \cdots, S_{16} 。

由于变压器故障类型复杂多变,难以进行细致化解析与划分。为此,本文参照 IEC60599 及 GB/T7252 导则,将中低温过热 ($< 700^{\circ}C$)、高温过热 ($> 700^{\circ}C$)、局部放电、低能放电、高能放电 5 种变压器故障类型作为决策表决策属性,编号依次记为 F_1, F_2, \cdots, F_5 。

本文搜集已明确故障类型的数据 301 条,抽取其中 126 条作为原始样本集,175 条作为测试集,各故障样本皆均匀分布,以此建立初始决策信息表与故障诊断模型。

4.2 粗糙集约简

构建的初始决策表是连续系统,需对其进行数据离散化处理。其中, S_1, S_5, S_8 采用 IEC 三比值法规定范围进行离散化,而 $S_{10}, S_{12}, S_{13}, S_{14}, S_{15}$ 依据文献[13]给出的基于 DGA 先验知识的方式进行离散化。除以上 8 条属性外,其余属性皆采用 2.1 节所述的等频划分算法离散,频数设定为 3。最终形成的离散化规则如表 1 所示。采用基于粗糙集知识约简方法对离散化后的决策表进行知识提取。依据 2.2 节约简步骤,在 Matlab 中进行程序的编写与运行,最终得到最小决策表,如表 2 所示。

表 1 离散化规则

Tab. 1 Discretization rules

属性 编码	编码规则			属性 编码	编码规则		
	0	1	2		0	1	2
S_1	[0. 10000, 1. 00000)	[0, 0. 10000)	[1. 00000, + ∞)	S_9	[0, 0. 05900)	[0. 05900, 0. 60000)	[0. 60000, + ∞)
S_2	[0, 0. 71572)	[0. 71572, 1. 90622)	[1. 90622, + ∞)	S_{10}	[0, 1. 00000)	[1. 00000, + ∞)	—
S_3	[0, 0. 00625)	[0. 00625, 0. 22930)	[0. 22930, + ∞)	S_{11}	[0, 0. 30000)	[0. 30000, 0. 90000)	[0. 90000, + ∞)
S_4	[0, 0. 00501)	[0. 00501, 0. 52304)	[0. 52304, + ∞)	S_{12}	[0, 0. 30000)	[0. 30000, 0. 90000)	[0. 90000, + ∞)
S_5	[0, 0. 10020)	[0. 10020, 3. 00020)	[3. 00020, + ∞)	S_{13}	[0, 0. 30000)	[0. 30000, 0. 90000)	[0. 90000, + ∞)
S_6	[0, 0. 02203)	[0. 02203, 0. 61300)	[0. 61300, + ∞)	S_{14}	[0, 0. 20000)	[0. 20000, 0. 70000)	[0. 70000, + ∞)
S_7	[0, 0. 16500)	[0. 16500, 2. 17003)	[2. 17003, + ∞)	S_{15}	[0, 0. 20000)	[0. 20000, 0. 70000)	[0. 70000, + ∞)
S_8	[0, 1. 00300)	[1. 00300, 3. 00100)	[3. 00100, + ∞)	S_{16}	[0, 0. 59949)	[0. 59949, 0. 83228)	[0. 83228, + ∞)

表 2 最小决策表

Tab. 2 Minimum decision table

序号	S_1	S_2	S_5	S_8	S_{11}	S_{12}	S_{14}	S_{13}	S_{16}	类型	序号	S_1	S_2	S_5	S_8	S_{11}	S_{12}	S_{14}	S_{13}	S_{16}	类型
1	0	1	0	1	1	1	0	1	2	F_1	17	1	2	0	0	1/2	1	0	0	1	F_3
2	2	2	0	0	0	1	0	0	1	F_1	18	1	0/1	2	2	1	0	1	0	0	F_3
3	2	1/2	0	1	0	1	0	1	2	F_1	19	2	2	1	0	0	1	0	0	0	F_3
4	2	2	0	0	0	2	0	0	2	F_1	20	0	2	1	1	1	1	0	1	1	F_4
5	2	2	1	0	0	1	0	0	0	F_1	21	1	2	0	0	1	1	0	0	0/1	F_4
6	2	2	0	1	0	1	0	1	1/2	F_1	22	1	2	0	0	2	1	0	0	1	F_4
7	2	1	0	1	0	1	0	1	1/2	F_1	23	1	0	0	1	2	0	0	1	0	F_4
8	2	1	1	1	0	1	0	1	1	F_1	24	1	1	1	2	1	0	1	1	0	F_4
9	2	1	0	0	0	0	0	1	0	F_1	25	1	2	0/1	0	1/2	1	0	0	1	F_4
10	0/2	0	0	2	0	0	0	1	2	F_2	26	2	2	1	1	1	1	0	0	1	F_4
11	2	1	0	2	0	1	0	1	2	F_2	27	0	1	1	1/2	1	1	0	1	1	F_5
12	2	1	0/1	1	0	1	0	1	1	F_2	28	0	0	1	2	0/1	0	1	1	0	F_5
13	2	0	1	2	0	0	0	1	1	F_2	29	1	0	1	2	1	0	1	1	0	F_5
14	2	0	0	2	0	0	0	1/2	2	F_2	30	1	2	1	0	2	1	0	0	0	F_5
15	0	1	1	2	1	0	2	1	0	F_3	31	2	0	1	2	0	0	1	1	0	F_5
16	1	2	0	0	1	1	0	0	0	F_3	32	2	2	1	2	0	1	0	0	2	F_5

可见,经粗糙集约简后,删除了7条冗余属性获得9条与故障相关的必要属性。由此,不仅使故障特征维度大幅度降低,也一定程度降低了数据采集量。其次,将最小决策表中的 $\{S_1、S_2、S_5、S_8、S_{11}、S_{12}、S_{13}、S_{14}、S_{16}\}$ 9条属性作为BPNN、SVM以及Bayesian网络三种智能方法的特征输入,对于算法结构简化具有一定意义。

4.3 信息决策融合

DS证据理论的信息融合方法可通过信度分布来描述故障特征对各种故障模式的支持度,因此可利用DS组合规则融合多源诊断证据并进行故障决策,从而获得更为精确的融合结果,克服单一诊断方式的片面性^[24]。基于此,本文采用BP神经网络、SVM以及Bayesian网络三种方式作为初步诊断模

块,并设定DS证据理论辨识框架 $\Omega = \{F_1, F_2, F_3, F_4, F_5\}$ 。如若三种方式诊断结果一致,则可直接得出诊断结果;若诊断结果不一致,依据式(7)构造每种诊断输出BPA后,利用式(3)、式(4)的DS组合规则对其进行信息融合后再进行最终诊断结果的判定。进行判定时,判定的运行模式与其他基本概率赋值最小阈值 ε_1 一般情况下要求大于 $1/n$ (n 为故障类型数量),此处取0.2,不确定度 $m(\theta)$ 存在最大阈值 ε_2 ,本文取0.2。最终由式(5)、式(6)与给定阈值得到最终诊断结果。

其中,各算法输入特征量皆为9种约简后的属性,输出为5种故障类型。BPNN采用Matlab中的神经网络工具箱进行编程,取输入层、隐含层、输出层3层,隐含层包括20个神经元,取tansig作为传

递函数;SVM 取径向基函数(RBF)作为核函数,相应参数 $\sigma=0.5$;Bayesian 网络则设定 9 个父节点,1 个子节点。

5 实例分析与比较

5.1 实例分析

以测试样本中的两组样本为例对其进行本文方法的验证,其中 1 号样本实际为高温过热故障,2 号样本为低能放电故障,两种样本的油色谱数据如表 3 所示。以降维后的 9 种特征属性作为三种诊断方式的特征输入,得到各诊断输出 BPA 及信息融合结果,如表 4 所示。

由表 4 可知,BPNN、SVM 与 Bayesian 网络三种单一智能诊断方法对两组样本的诊断结果依次为 $\{F_1, F_4\}$ 、 $\{F_2, F_4\}$ 、 $\{F_2, F_2\}$,可见,其对 2 组样本的判断并不一致,且存在误判。从构造的 BPA 结果

表 3 样本数据

Tab. 3 Sample data						
编号	特征气体浓度/($\mu\text{L/L}$)					故障类型
	H_2	CH_4	C_2H_2	C_2H_4	C_2H_6	
1	21.00	124.00	5.00	183.00	45.00	F_2
2	68.70	9.05	34.88	11.38	1.71	F_4

可知,三种单一算法的不确定度皆偏大,且判定的故障与不确定度间的 BPA 相差过小,BPNN 的不确定度甚至超过了 0.2,增加了最终判定结果的不真实性。特别地,在 Bayesian 网络对 2 号样本的判定输出概率赋值中,形成每种输出与不确定度皆相近的现象,导致最终决策可信度下降。而利用 DS 组合规则对其进行信息融合后,其 BPA 分布清楚明晰。对比表 3 与表 4 可知,最终判断也与实际相符,且不确定度相比于融合前明显下降,决策可信度增加。

表 4 基本概率分配及融合结果

Tab. 4 Basic probability assignment and fusion results								
方式	编号	$m(F_1)$	$m(F_2)$	$m(F_3)$	$m(F_4)$	$m(F_5)$	$m(\theta)$	判断
BPNN	1	0.4597	0.2235	0.0162	0.0349	0.0239	0.2515	F_1
	2	0.0668	0.0082	0.0996	0.5469	0.0673	0.2285	F_4
SVM	1	0.1025	0.5549	0.0254	0.0084	0.0844	0.1554	F_2
	2	0.0034	0.0369	0.0444	0.5838	0.0148	0.1428	F_4
Bayesian	1	0.0412	0.6944	0.0029	0.0185	0.0797	0.1302	F_2
	2	0.1391	0.2925	0.2136	0.1472	0.0615	0.1253	F_2
DS	1	0.0193	0.8702	0.0002	0.0003	0.0018	0.0685	F_2
	2	0.0008	0.0033	0.0545	0.8082	0.0096	0.1168	F_4

5.2 不同诊断方式对比分析

为验证本文方法的有效性,利用相同样本集对 3 种单一诊断方法进行故障诊断,并分别将不同故障类型下的诊断准确率和总体准确率与本文方法进行对比。图 2 为各故障诊断准确率。各方式总体诊断准确率如表 5 所示。

表 5 单一算法与本文方法诊断准确率

Tab. 5 Accuracy rate of some single algorithm and method of this article						
算法	三比值	BPNN	SVM	Bayesian	本文方法	
诊断率(%)	71.00	73.58	78.86	83.12	89.03	

由图 2 与表 5 可以看出,无论是单一诊断还是总体,本文方法诊断准确率皆优于单一手段。这是由于单一诊断方式皆存在不同程度的缺陷:①三比值法比值编码不全,且编码边界过于绝对;②BPNN 要获得较理想效果需要在大量样本下训练;③SVM 核函数参数确定困难;④Bayesian 在关键属性缺失时诊断效果不理想。而本文将各单一诊断结果进行融合,由此实现各单一方法的优势互补,降低了各单一算法缺陷对诊断结果的不利影响。同时,高能放电由于往往兼有过热现象,相较于其他故障,其故障识别率偏低,但依旧在 85% 以上。综上所述,本文

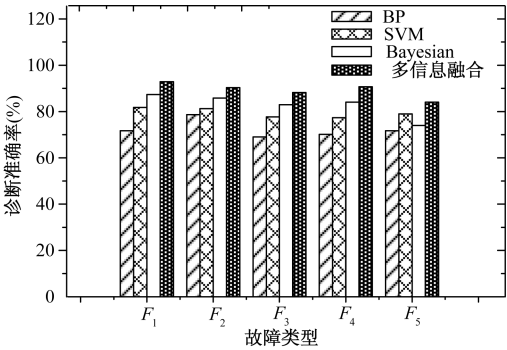


图 2 不同故障下各诊断方式对比

Fig. 2 Comparison of different diagnostic methods

方式可取得较单一方法更可靠有效的诊断结果。另外,经降维后的比值属性集既包含了三比值法的三种气体比值,又含有其他 5 种比值,可作为三比值法的一种扩充,可为完善三比值法提供一种新思路。

6 结论

针对变压器故障诊断特征信息冗余度高与诊断方法单一问题,本文构建了一种基于粗糙集与 DS 证据理论相结合的诊断方法,得到如下结论:

(1)利用粗糙集约简知识获得了 9 条必要特征属性,降低了故障特征集冗余度,使得故障特征判据相关性有所提高。

(2)利用 DS 组合规则对 BPNN、SVM 以及 Bayesian 网络三种诊断方法的不同判断结果进行决策融合,实现了各冲突证据的有效融合,可获得比单一诊断方法更有效的诊断结果。

(3)经实例证明,本文方法能有效减小故障特征冗余信息,并解决单一方法结论片面性问题,使得诊断准确率获得较大提升。

参考文献 (References):

- [1] Li S B, Wu G N, Gao B, et al. Interpretation of DGA for transformer fault diagnosis with complementary SaE-ELM and arctangent transform [J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2016, 23 (2): 586-595.
- [2] Ghoneim S S M, Taha I B M. A new approach of DGA interpretation technique for transformer fault diagnosis [J]. International Journal of Electrical Power & Energy Systems, 2016, 81: 265-274.
- [3] 陈伟根, 潘翀, 云玉新, 等 (Chen Weigen, Pan Chong, Yun Yuxin, et al.). 基于小波网络及油中溶解气体分析的电力变压器故障诊断方法 (Fault diagnosis method of power transformers based on wavelet networks and dissolved gas analysis) [J]. 中国电机工程学报 (Proceedings of the CSEE), 2008, 28 (7): 121-126.
- [4] 高骏, 何俊佳 (Gao Jun, He Junjia). 量子遗传神经网络在变压器油中溶解气体分析中的应用 (Application of quantum genetic ANNs in transformer dissolved gas-in-oil analysis) [J]. 中国电机工程学报 (Proceedings of the CSEE), 2010, 30 (30): 121-127.
- [5] 赵文清, 李庆良, 王德文 (Zhao Wenqing, Li Qingliang, Wang Dewen). 基于多模型的变压器故障组合诊断研究 (Combinational diagnosis for transformer faults based on multi-models) [J]. 高电压技术 (High Voltage Engineering), 2013, 39 (2): 302-309.
- [6] 郑含博, 王伟, 李晓纲, 等 (Zheng Hanbo, Wang Wei, Li Xiaogang, et al.). 基于多分类最小二乘支持向量机和改进粒子群优化算法的电力变压器故障诊断方法 (Fault diagnosis method of power transformers using multi-class LS-SVM and improved PSO) [J]. 高电压技术 (High Voltage Engineering), 2014, 40 (11): 3424-3429.
- [7] 赵文清 (Zhao Wenqing). 基于选择性贝叶斯分类器的变压器故障诊断 (Transformer fault diagnosis based on selective Bayes classifier) [J]. 电力自动化设备 (Electric Power Automation Equipment), 2011, 31 (2): 44-47.
- [8] 谢龙君, 李黎, 程勇, 等 (Xie Longjun, Li Li, Cheng Yong, et al.). 融合集对分析和关联规则的变压器故障诊断方法 (A fault diagnosis method of power transformers by integrated set pair analysis and association rules) [J]. 中国电机工程学报 (Proceedings of the CSEE), 2015, 35 (2): 277-286.
- [9] 郑元兵, 孙才新, 李剑, 等 (Zheng Yuanbing, Sun Caixin, Li Jian, et al.). 变压器故障特征量可信度的关联规则分析 (Association rule analysis on confidence of features for transformer faults) [J]. 高电压技术 (High Voltage Engineering), 2012, 38 (1): 82-88.
- [10] 赵峰, 李硕 (Zhao Feng, Li Shuo). 基于熵权和灰关联度的变压器故障诊断 (Fault diagnosis in transformer based on entropy weighted and grey incidence) [J]. 电工电能新技术 (Technology of Electrical Engineering & Energy), 2015, 34 (1): 57-61.
- [11] 陈小青, 刘觉民, 黄英伟, 等 (Chen Xiaoqing, Liu Juemin, Huang Yingwei, et al.). 采用改进人工鱼群优化粗糙集算法的变压器故障诊断 (Transformer fault diagnosis using improved artificial fish swarm with rough set algorithm) [J]. 高电压技术 (High Voltage Engineering), 2012, 38 (6): 1403-1409.
- [12] 李志斌, 陈成优 (Li Zhibin, Chen Chengyou). 基于粗糙集和贝叶斯网络的变压器故障诊断 (Fault diagnosis of power transformer based on rough set theory and Bayesian network) [J]. 上海电力学院学报 (Journal of Shanghai University of Electric Power), 2013, 29 (1): 30-34.
- [13] 张景明, 肖倩华, 王时胜 (Zhang Jingming, Xiao Qianhua, Wang Shisheng). 融合粗糙集和神经网络的变压器故障诊断 (Transformer fault diagnosis by combination of rough set and neural network) [J]. 高电压技术 (High Voltage Engineering), 2007, 33 (8): 122-125.
- [14] Zhang L, Yuan J, Zhao C. Improvement of transformer gas-in-oil diagnosis based on evidence theory [A]. 2012

- Asia-Pacific Power and Energy Engineering Conference (APPEEC) [C]. 2012. 1-4.
- [15] 陈伟根, 刘娟, 曹敏 (Chen Weigen, Liu Juan, Cao Min). 基于信息融合的变压器内部故障诊断方法 (Diagnosis method of internal fault for transformers based on information fusion) [J]. 高电压技术 (High Voltage Engineering), 2015, 41 (11): 3797-3803.
- [16] 董海鹰, 李娜 (Dong Haiying, Li Na). 基于 D-S 证据理论信息融合的转辙机故障诊断方法研究 (Fault diagnosis method for switch machine based on D-S evidence theory information fusion) [J]. 测试技术学报 (Journal of Test & Measurement Technology), 2013, 27 (1): 1-7.
- [17] 杨廷方, 周力行, 李景禄, 等 (Yang Tingfang, Zhou Lixing, Li Jinglu, et al.). 基于最优权值的组合模型诊断变压器故障 (Power transformer fault diagnosis based on optimal weights combined model) [J]. 电网技术 (Power System Technology), 2013, 37 (1): 190-194.
- [18] 周献中, 黄兵, 李华雄, 等 (Zhou Xianzhong, Huang Bing, Li Huaxiong, et al.). 不完备信息系统知识获取的粗糙集理论与方法 (Rough set theory and method of knowledge acquisition in incomplete information system) [M]. 南京: 南京大学出版社 (Nanjing: Nanjing University Press). 2010. 12.
- [19] 潘泉 (Pan Quan). 多源信息融合理论及应用 (Multi-source information fusion theory and its application) [M]. 北京: 清华大学出版社 (Beijing: Tsinghua University Press), 2013.
- [20] Shafer G. A mathematic theory of evidence [M]. Princeton, USA: Princeton University Press, 1976.
- [21] 阮羚, 谢齐家, 高胜友, 等 (Ruan Ling, Xie Qijia, Gao Shengyou, et al.). 人工神经网络和信息融合技术在变压器状态评估中的应用 (Application of artificial neural network and information fusion technology in power transformer condition assessment) [J]. 高电压技术 (High Voltage Engineering), 2014, 40 (3): 822-828.
- [22] 董明, 孟源源, 徐长响, 等 (Dong Ming, Meng Yuanyuan, Xu Changxiang, et al.). 基于支持向量机及油中溶解气体分析的大型电力变压器故障诊断模型研究 (Fault diagnosis model for power transformer based on support vector machine and dissolved gas analysis) [J]. 中国电机工程学报 (Proceedings of the CSEE), 2003, 23 (7): 88-92.
- [23] Duval M, Depabla A. Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases [J]. IEEE Electrical Insulation Magazine, 2001, 17 (2): 31-41.
- [24] Xu X B, Liu P, Sun Y B, et al. Fault diagnosis based on the updating strategy of interval-valued belief structures [J]. Chinese Journal of Electronics, 2014, 23 (4): 753-760.

Fault diagnosis of transformer based on rough set theory and decision information fusion

LI Chun-mao, ZHOU Mo-mo, YUAN Hai-man, GAO Bo, WU Guang-ning

(School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: Aiming at the problem of information uncertainty and redundancy of transformer fault feature and the singularity of traditional method of fault diagnosis, a fault diagnosis model of transformer based on rough set theory and multi information fusion is proposed. Firstly, 16 groups of characteristic gas ratios are considered as fault characteristic parameters, and information is extracted by discretization rules and rough set to reduce feature information redundancy effectively. Additionally, the reduced attribute set is used as the input of BP neural network (BPNN), support vector machine (SVM) and Bayesian network to diagnose the fault types. Finally, the decision fusion of three kinds of preliminary judgment results is made by DS information fusion rules to obtain more efficient fault judgment conclusion. Example analysis shows that the proposed method can effectively reduce the impact of redundant feature information on the diagnosis result and solve the conflict of evidence fusion, and improve the accuracy of fault recognition. It's easy to see the performance of the method is superior to the single diagnosis method.

Key words: transformer; Dempster-Shafer theory; attribute reduction; discretization; rough set theory; information fusion; fault diagnosis