

文章编号: 1001-0920(2004)08-0943-04

基于粗糙集理论的电力变压器故障诊断方法研究

倪远平, 周建华, 李彬华, 邹金慧

(昆明理工大学 信息与自动化学院, 云南 昆明 650051)

摘 要: 为了能从完备信息中得出正确的诊断结论, 并能从不完整的信息中得出满意的诊断结果, 提高故障识别率, 提出一种基于粗糙集理论的电力变压器故障诊断方法 给出了粗糙集的基本概念、约简计算方法和约简过程, 并给出了故障诊断原理和诊断实例 实验结果表明, 该方法是有用的, 达到了预期目的, 为电力变压器故障诊断提供了新的思路, 对其他电气设备的故障诊断具有参考价值

关键词: 粗糙集; 约简; 决策表; 故障诊断; 变压器

中图分类号: TP391

文献标识码: A

Method of fault diagnosis for power transformer based on rough set theory

N I Yuan-ping, ZHOU Jian-hua, LI Bin-hua, ZOU Jin-hui

(Faculty of Information and Automation, Kunming University of Science and Technology, Kunming 650051, Correspondent: N I Yuan-ping, Email: niyp-me@yahoo.com.cn)

Abstract: In order to get right conclusions of diagnosis from maturity information and satisfied results of diagnosis from incomplete information, and to improve the recognizing rate of faults, a method of fault diagnosis for transformer based on rough set theory is presented. The basic concepts of rough set, the algorithm of reduction, the principle of fault diagnosis and the practical examples are given out. Experiment results show that the method is effective and reaches the expect aim.

Key words: rough set; reduction; decision table; fault diagnosis; transformer

1 引 言

现代工业系统在整个国民经济中占有举足轻重的地位, 而电力变压器是现代工业系统中的重要设备, 它是否可靠安全运行, 直接影响电网运行的可靠性、安全性以及企业的经济效益和社会效益。对电力变压器故障进行监测和诊断, 是一个重要的技术保障措施。

电力变压器故障诊断实质上是一个模式分类问题。近年来, 研究工作者提出了许多故障诊断方法,

如专家系统方法、神经网络方法、模糊逻辑方法等。这些方法为建立智能监测与诊断, 提高决策效率非常有益。本文以变压器特征气体信息为依据, 尝试应用粗糙集理论, 提出一种电力变压器故障诊断的新方法。仿真实验表明, 该方法可行, 收到了良好的效果。

2 粗糙集基本理论

粗糙集理论^[1] (Rough Set) 是由波兰数学家 Pawlak 于 1982 年提出的一种新型数学工具, 主要

收稿日期: 2003-09-01; 修回日期: 2003-11-14

基金项目: 云南省科技厅应用基础基金资助项目(2000F004M); 云南省“控制理论与控制工程”省级重点学科建设基金资助项目(14039043)。

作者简介: 倪远平(1957—), 女, 重庆人, 教授, 从事智能控制与智能故障诊断的研究; 周建华(1963—), 男, 云南昆明人, 讲师, 从事智能控制与智能故障诊断的研究。

用于刻画不完整数据和不确定知识的表达、学习和归纳。它从新的角度对知识进行定义,把知识看作是论域的划分。这一理论的特点是无需任何附加信息或先验知识,能有效地分析和处理不精确、不完整数据,发现数据间隐含的关系,揭示潜在的规律。经过近 20 年的发展,该理论正日益受到国际上的广泛关注,目前已经在人工智能、知识与数据发现、人面检测、模式识别与分类、决策支持与分析等方面得到了成功的应用^[2,3]。

2.1 粗糙集的基本概念

定义 1 设 U 是一个非空有限论域,任何子集 $X \subseteq U$ 称为 U 中的一个概念, U 中的任何概念族称为关于 U 的知识, U 上的一族划分称为 U 上的一个知识库

定义 2 设 r 为 U 上的一个等价关系, U/r 表示 r 在 U 的等价分类所构成的集合。一个知识库就是一个关系系统 $K = (U, R)$, R 是 U 上的一族等价关系

定义 3 令 R 为一族等价关系, $r \in R$, 如果 $\text{ind}(R) = \text{ind}(R - \{r\})$, 则称 r 为 R 中不必要的, 否则 r 为 R 中必要的。如果每一个 $r \in R$ 都为 R 中必要的, 则称 R 为独立的, 用 $\text{ind}(R)$ 表示; 否则称 R 为依赖的

定义 4 设 $Q \subseteq P$, 如果 Q 是独立的, 且 $\text{ind}(Q) = \text{ind}(P)$, 则称 Q 为 P 的一个约简, 用 $\text{red}(P)$ 表示。显然, P 可以有多种约简, 全部约简的交集定义为 P 的核, 记作 $\text{core}(P) = \bigcap \text{red}(P)$ 。

2.2 知识约简及决策表^[4,5]

知识约简是粗糙集理论的核心内容之一。人们在对一个事物作判断和决策时,不是依据被判断事物的全部特性,而是依据最主要的一个或几个特点作出判断。所谓知识约简,就是在保持知识库分类能力不变的条件下,删除其中不相关或不重要的冗余知识,以简化判断规则。

知识表达在智能数据处理中占有十分重要的地位。四元组 $S = (U, A, V, f)$ 是一个知识表达系统,其中: U 是论域, A 是属性集合, V 是属性值的集合, $f: U \times A \rightarrow V$ 是一个信息函数,它指定 U 中每一对象的属性值。 $A = C \cup D$, C 称为条件属性集, D 称为决策属性集。具有条件属性和决策属性的知识表达系统称为决策表——特殊的信息表。一个关于某些病人的决策表如表 1 所示。决策表中一个属性对应一个等价关系,一个表可看作定义的一族等价关系,即知识库。决策表中并非所有的条件属性都是必

表 1 关于某些病人的决策表

病人	条件属性			决策属性
	头痛	肌肉痛	体温	流感
1	是	是	正常	否
2	是	是	高	是
3	是	是	很高	是
4	否	是	正常	否
5	否	否	高	否
6	否	是	很高	是

要的,其中有些是多余的,知识约简可转化为属性约简。

决策表的简化就是简化表中的条件属性,即去除冗余的条件属性。删除这些属性后不会影响原有的表达效果。

决策表的简化步骤如下:

Step 1: 对决策表中可省略的条件属性进行约简;

Step 2: 对每一决策规则中的冗余属性进行约简;

Step 3: 去掉多余的规则。

表 2 给出了一个决策表的简化例子。论域 U 中有 7 个对象, C_1, C_2, C_3, C_4 为条件属性, d 为决策属性。根据 d 取值的不同,可将这 7 个对象分为 3 类。按决策表的简化步骤,逐一去除条件属性,看是否正确对 d 的分类产生影响。当分别去掉 C_1, C_2 和 C_4 时,发现对 d 的正确分类有影响,即有一部分“条件-决策”是不相容的,因而不能省略。当去掉 C_3 时,对 d 的正确分类没有影响,即:“条件-决策”是相容的, C_3 是可省略的。因此, $\{C_1, C_2, C_4\}$ 是必需的条件属性,称为原条件属性 $\{C_1, C_2, C_3, C_4\}$ 的一个约简。一个决策表的条件属性可能存在多个约简,本例只有一个,如表 3 所示。

表 2 某决策表

u	C_1	C_2	C_3	C_4	d
1	1	0	0	1	1
2	1	0	0	0	1
3	0	0	0	0	0
4	1	1	0	1	0
5	1	1	0	2	2
6	2	1	0	2	2
7	2	2	2	2	2

下面针对表 3 计算各规则的约简。先计算各规则的核值。考虑第 1 个对象的“条件-决策”关系,也就是第 1 条规则:

$$C_1 C_2 C_4 \rightarrow d, \text{ 即 } 1 \ 0 \ 1 \ 1$$

去掉 $C_1=1$, 得 $0 \ 1 \ 1$ 相容,故 $C_1=1$ 是多余的; 去掉

表 3 约简(C₁, C₂, C₄) 决策表

U	C ₁	C ₂	C ₄	d
1	1	0	0	1
2	1	0	0	1
3	0	0	0	0
4	1	1	0	0
5	1	1	0	2
6	2	1	0	2
7	2	2	2	2

C₂ = 0, 得 11 1 不相容, 故 C₂ = 0 是核值; 去掉 C₃ = 1, 得 10 1 相容, 故 C₃ = 1 是多余的

注意: 这里规则 01 1 和 10 1 已是规则 101 1 的最简形式了。

依次对每一个规则求其核值并获得所有最简形式, 消去每一决策规则中的冗余属性, 简化后得到的最终决策如表 4 所示。符号 * 表示该属性的取值对分类没有影响

表 4 简化后的最终决策表

U	C ₁	C ₂	C ₄	d
1	1	0	*	1
2	0	*	*	0
3	*	1	1	0
4	*	*	2	2

显然表 4 比表 1 简单得多, 它仅包含那些在决策时所必需的条件属性值, 但它具有原始知识系统的所有知识和分类能力

3 基于粗糙集理论的故障诊断方法

电力变压器的故障诊断可描述成一个模式分类问题, 因此很适合于应用粗糙集理论的决策表方法进行处理

对于电力变压器内部故障, 通常通过分析油中的气体进行检测。当变压器内部发生放电故障或潜伏性过热时, 在热和电的作用下, 会分解产生出 CO, H₂, CH₄, C₂H₆, C₂H₄ 和 C₂H₂ 等主要特征气体。故障特征气体的组成和含量与故障的类型和故障的严重程度密切相关。分析溶解于油中的特征气体及比值关系, 便能判别出变压器内部发生了何种故障及严重程度。下面以故障主要特征气体及比值关系为依据, 建立变压器故障诊断决策表。通过决策表的化简得到简化的诊断规则, 从而实现电力变压器的故障诊断

由于变压器故障十分复杂, 限于篇幅, 本文只讨论变压器的几种大致故障分类, 如图 1 所示

这里将故障特征气体 H₂, CH₄, C₂H₆, C₂H₄, C₂H₂ 彼此间的比值按大小划分, 作为故障分类的条

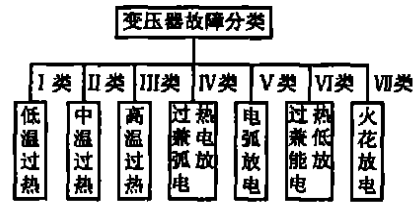


图 1 变压器故障分类

件属性。在改良的 IEC 法 (国际电工委员会推荐的比值法) 的基础上, 这些比值划分为:

- 1) 当 $C_2H_2/C_2H_4 < 0.1$, 则 $C_1 = 0$;
- 2) 当 $0.1 < C_2H_2/C_2H_4 < 3$, 则 $C_1 = 1$;
- 3) 当 $C_2H_2/C_2H_4 \geq 3$, 则 $C_1 = 2$;
- 4) 当 $CH_4/H_2 < 0.1$, 则 $C_2 = 1$;
- 5) 当 $0.1 < CH_4/H_2 < 1$, 则 $C_2 = 0$;
- 6) 当 $CH_4/H_2 \geq 1$, 则 $C_2 = 2$;
- 7) 当 $C_2H_4/C_2H_6 < 1$, 则 $C_3 = 0$;
- 8) 当 $1 < C_2H_4/C_2H_6 < 3$, 则 $C_3 = 1$;
- 9) 当 $C_2H_4/C_2H_6 \geq 3$, 则 $C_3 = 2$

根据比值划分, 多年的历史数据以及实际故障分析, 可建立变压器故障诊断决策表如表 5 所示

表 5 变压器故障诊断决策表

U	C ₂ H ₂ /C ₂ H ₄ C ₁	CH ₄ /H ₂ C ₂	C ₂ H ₄ /C ₂ H ₆ C ₃	诊断结果
1	0	0	1	I
2	0	2	0	I
3	0	2	1	II
4	0	2	2	III
5	0	0	2	III
6	0	1	2	III
7	1	2	0	IV
8	1	2	2	IV
9	1	2	1	IV
10	1	0	0	V
11	1	1	0	V
12	1	0	1	V
13	1	1	1	V
14	1	0	2	V
15	1	1	2	V
16	2	2	1	VI
17	2	2	0	VI
18	2	2	2	VI
19	2	1	0	VII
20	2	0	0	VII
21	2	1	2	VII
22	2	0	2	VII
23	2	1	1	VII
24	2	0	1	VII
25	0	0	0	No Fault

表 5 中的条件属性共有 3 个, 即 C₁, C₂ 和 C₃。按

决策表的简化步骤, 首先逐一去除条件属性进行简化. 当分别去掉 C_1 , C_2 和 C_3 时, 通过对比计算, 发现对诊断结果的正确分类有影响, 有一部分“条件决策”是不相容的, 因而 C_1 , C_2 和 C_3 均不能省略. 然后, 通过计算, 消去表 5 中每一决策规则的冗余属性, 并去掉多余的规则, 便得到了约简的最终决策表, 如表 6 所示.

表 6 变压器约简的最终决策表

U	C_2H_2/C_2H_4		CH_4/H_2	C_2H_4/C_2H_6		诊断结果
	C_1	C_2	C_3	C_3	C_3	
1	0	0	1			I
2	0	2	0			I
3	0	2	1			II
4	0	*	2			III
5	1	2	*			IV
6	1	1/0	*			V
7	2	2	*			VI
8	2	1/0	*			VII
9	0	0	0			No Fault

将表 6 与表 5 相比较, 可明显地看到本文方法的优化效果. 表 6 的结构非常简单, 仅包含决策时所必需的条件属性值和决策规则. 它不仅具有原决策表的知识 and 分类能力, 能对 8 类故障进行正确诊断, 而且揭示了故障特征气体信息的冗余性. 符号 * 表示该属性的取值对故障分类没有影响.

表 6 有 9 条简化的故障诊断规则, 例如规则 4: If $C_1 = 0$ and $C_3 = 2$, Then III 类故障. 即: 如果特征气体 $C_2H_2/C_2H_4 < 0.1$ 和特征气体 $C_2H_4/C_2H_6 = 3$, 那么变压器发生了高温过热故障, 这里不需要特征气体 CH_4/H_2 的比值. 本文方法删除了冗余的特征气体信息, 简单明了, 保持了与原决策表完全相同的分类能力, 为处理不完整特征信息提供了一条新思路.

实例 1 淮北供马庄 1 号变压器, 油中气体含量如下 (ppm): $H_2 = 1100$, $CH_4 = 7800$, $C_2H_6 = 6500$, $C_2H_4 = 36200$, $C_2H_2 = 200$. 用基于粗糙集理论的故障诊断方法进行判断, 规则 4 可决策出该故障属于高温过热. 实际吊芯检查发现分接开关烧坏, 与实际结果吻合.

实例 2 银川变电站 1 号主变 (110kV, 40MVA), 油中气体含量如下 (ppm): $H_2 = 1198$,

$CH_4 = 32$, $C_2H_6 = 1.4$, $C_2H_4 = 32$, $C_2H_2 = 0.5$. 此时采用普通的 IEC 法诊断不出故障类型, 因为在 IEC 法编码规则中没有包含这类故障的编码. 而用基于粗糙集理论的故障诊断方法进行诊断, 符合规则 6, 得出结论是变压器存在电弧放电故障, 诊断与吊芯检测结果完全吻合.

实验说明, 对完备或不完整信息的上述 9 类故障都可以正确诊断, 解决了 IEC 法故障识别率低的问题. 但是, 因为在 IEC 法基础上确定的条件属性只有 3 个, 化简后的约简只有一个, 因而对新故障的诊断存在局限性. 这需增加条件属性, 并得到多个约简才能较好解决.

4 结 论

电力变压器故障诊断实质上是一个模式分类问题. 本文提出的基于粗糙集理论的变压器故障诊断方法不仅可以直接从完备的故障特征气体中导出正确的诊断结论, 而且能从不完整的故障特征气体中导出满意的诊断结论. 它揭示了故障特征气体信息的冗余性, 提高了故障的识别率. 实例诊断结果验证了该方法的有效性, 然而对新故障的诊断尚需进一步探讨.

参考文献 (References):

- [1] Pawlak Z. Rough sets approach to knowledge-based decision support [J]. *European J of Operational Research*, 1997, 99(1): 48-57.
- [2] 范劲松, 方廷健. 基于粗糙集理论和 SVM 算法的模式分类方法 [J]. *模式识别与人工智能*, 2000, 13(4): 419-423. (Fan J S, Fang T J. A method of pattern classification on line based on rough set and SVM algorithm [J]. *Pattern Recognition and Artificial Intelligence*, 2000, 13(4): 419-423.)
- [3] 胡涛, 吕炳朝, 陈光禹. 基于粗糙集理论的旋转机械故障诊断方法研究 [A]. *WICAC [C]*. 合肥: 中国科技大学出版社, 2000. 685-689.
- [4] 张文修, 吴伟志, 梁吉业, 等. 粗糙集理论与方法 [M]. 北京: 科学出版社, 2001, 1-115.
- [5] 徐德友, 胡寿松. 一种基于粗糙集的近似质量求取属性约简的决策算法 [J]. *控制与决策*, 2003, 18(3): 313-316. (Xu D Y, Hu S S. Decision algorithm for finding reduct based on approximation quality of rough set [J]. *Control and Decision*, 2003, 18(3): 313-316.)