

基于选择性贝叶斯分类器的变压器故障诊断

赵文清

(华北电力大学 控制与计算机工程学院, 河北 保定 071003)

摘要: 电力变压器故障诊断中的测试数据信息不完备、有偏差, 而贝叶斯网络处理不确定性问题能力强。提出了一种基于选择性贝叶斯分类器的、溶解气体分析结合其他电气试验结果的变压器故障诊断方法, 并建立了变压器选择性贝叶斯故障诊断模型。详细阐述并验证了该方法解决信息不完备问题的优越性。该模型还可以通过不断积累完善训练样本, 自动修正网络结构参数和概率分布参数。实验表明提出的选择性贝叶斯分类器适于变压器故障诊断。

关键词: 变压器; 故障诊断; 贝叶斯网络; 选择性分类器

中图分类号: TM 41; TP 206⁺.3

文献标识码: A

文章编号: 1006-6047(2011)02-0044-04

0 引言

针对传统的变压器故障诊断方法的不足, 各种智能技术被引入变压器故障诊断中, 并取得了比较好的效果^[1-5]。变压器检修试验所获取的数据包括离线试验数据和连续监测的在线数据, 在实际应用中, 由于受到环境和条件的限制, 大部分情况下得到的数据样本或者包含不完整数据或者变压器故障样本数很少, 如何充分利用有限的变压器数据样本和不完整数据样本中包含的信息, 始终是变压器诊断领域研究的热点问题。

贝叶斯网络用概率测度的权重来描述数据间的相关性, 从而解决了数据间的不一致性问题 and 信息不完备问题^[6-8]。针对变压器故障诊断中原始试验数据的属性缺失问题, 文献[9]使用粗糙集进行属性的约简, 然后再利用贝叶斯网络分类器进行故障甄别, 取得了较好的效果; 但即使是经过属性约简之后的试验数据, 可能还有部分数据不完整。本文将使用选择性贝叶斯分类器解决该问题, 直接从不完整数据计算贝叶斯网络结构的测度值来学习一般的贝叶斯网络。选择性分类器通过删除数据集中的无关属性和冗余属性可有效地提高分类精度和效率。

在不完整数据条件下, 一般贝叶斯网络的学习是一个难点问题。虽然在这方面取得了一定的成果, 如近似边缘分布方法^[10]、AutoClass 方法^[11]、SEM (Structural EM) 算法^[12]等。但这些方法的共同策略是对不完整数据进行修补, 采用循环迭代、逐步求精的学习过程, 这需多次迭代才能收敛, 并且每向前迭代一步都要进行一次数据修补。变压器故障诊断中, 即使经过粗糙集约简, 不完整数据仍有可能比较

多, 若采用以上方法学习贝叶斯网络, 将是很耗时甚至是不可行的; 同时, 这些方法都基于缺值数据满足 MAR (Missing At Random) 假设^[13]。当不满足这个假设时, 其数据修补方法的精度会显著下降, 由此构造的分类器的精度也会下降^[14]。为避免 MAR 假设, 文献[15]提出了一种 RBC (Robust Bayes Classifier), 该方法在不要求缺值数据满足 MAR 假设的前提下, 可以直接从不完整数据构造贝叶斯分类器。文献[16]提出了用于不完整数据的选择性贝叶斯分类器 SRBC (Selective Robust Bayes Classifiers), SRBC 是在 RBC 基础上利用包装法构建的用于不完整数据的选择性贝叶斯分类器; 并证明该方法在大幅度减少属性数目的同时, 能显著提高分类准确率和稳定性, 分类效果优于朴素贝叶斯和其他选择性分类器。

1 基于 SRBC 的变压器故障诊断模型

1.1 属性变量、故障类、训练样本集的确定

变压器故障属性集、故障类、训练样本集分别如表 1、表 2、表 3 所示^[9]。表 3 中 $x_2=0$ 、 $x_2=1$ 、 $x_2=2$ 、 $x_5=0$ 、 $x_5=1$ 、 $x_5=2$ 分别表示油中溶解气体分析 (DGA) 结果为无过热、低温过热、高温过热、无放电、低能量放电、高能量放电; x_1 、 x_3 、 x_4 、 x_7 、 x_8 、 x_9 、 x_{10} 、 x_{11} 、 x_{12} 、 x_{13} 、 x_{14} 的值为 0、1、2 分别表示属性值和趋势值均不超标、属性值或趋势值超标、属性值和趋势值均超标; * 表示因试验未做而不能确定的值。

1.2 模型的建立

本文利用文献[16]提出的用于不完整数据的 SRBC 直接从不完整数据计算贝叶斯网络结构的测度值来学习一般的贝叶斯网络, 从而构建针对变压器故障属性缺失之下的故障诊断模型。文献[16]在利用包装法构建 SRBC 的过程中, 采用了搜索效果好而复杂度相对较低的最好优先前向搜索方法对属性空间进行搜索。

表 1 属性集
Tab.1 Attributes set

属性变量	征兆类型
x_1	铁心接地电流
x_2	DGA 诊断结果为过热
x_3	绕组直流电阻三相不平衡
x_4	本体油中微水含量
x_5	DGA 诊断结果为放电
x_6	瓦斯继电器保护动作
x_7	绕组变比
x_8	局部放电参数
x_9	油中气体比值 $\varphi(\text{CO}_2)/\varphi(\text{CO})$
x_{10}	绕组间或绕组对地电容量
x_{11}	绕组直流电阻
x_{12}	绝缘介质损耗
x_{13}	空载电流及损耗
x_{14}	绕组吸收比、极化指数

表 2 故障类
Tab.2 Fault classes

类	故障类型
C_0	变压器正常
C_1	铁心多点接地或局部短路
C_2	绝缘老化
C_3	漏磁引起发热或磁屏蔽过热
C_4	匝绝缘损伤并匝间短路
C_5	绝缘受潮
C_6	分接开关及引线缺陷
C_7	悬浮放电
C_8	围屏放电
C_9	线圈变形并匝间短路
C_{10}	裸金属过热
C_{11}	变压器油流受阻引起过热
C_{12}	有载分接开关油渗漏

表 3 训练样本集
Tab.3 Training samples set

序号	数量	属性变量														类
		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	
1	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	C_0
2	14	2	2	0	1	0	1	*	*	*	*	0	1	0	2	C_1
3	5	1	2	0	0	*	*	*	*	1	1	*	*	*	2	C_1
4	6	*	1	0	1	0	0	*	0	2	0	0	*	*	1	C_2
5	10	0	2	0	0	0	*	*	*	*	0	0	0	1	*	C_3
6	18	0	2	*	0	0	1	0	0	0	*	*	*	*	0	C_3
7	12	0	0	1	0	2	1	0	1	1	*	1	0	2	1	C_4
8	6	*	0	1	1	0	*	*	0	1	*	*	0	*	0	C_5
9	2	*	2	2	1	0	1	*	1	0	*	*	*	*	0	C_6
10	6	1	0	1	0	2	1	0	2	0	0	0	*	*	1	C_7
11	13	0	0	0	1	2	1	0	2	2	0	0	0	*	*	C_8
12	16	1	0	*	0	2	*	0	0	0	*	1	0	0	2	C_9
13	12	*	2	0	*	0	*	*	*	1	0	1	0	1	*	C_{10}
14	2	0	1	0	0	0	0	0	0	0	0	1	1	0	*	C_{11}
15	4	0	1	0	1	2	0	*	1	0	*	*	0	*	1	C_{12}

设 $X=\{x_1,x_2,\dots,x_N\}$ 为整个变压器故障属性的集合, N 为 X 中属性的个数, 本文取 $N=14$; Q 为一个队列, 用来存放曾是最优的变压器故障属性子集及其对应的故障诊断正确率; S_b 为当前最优变压器故障属性子集, $d(S)$ 表示 RBC 在变压器故障属性子集 S 上的故障诊断正确率; 阈值 T 为用来控制搜索过程是

否停止的参数, 即如果连续 T 次对 Q 的头结点进行扩展都没有使当前最高故障诊断正确率改善, 则搜索过程结束。

SRBC 算法可描述如下:

- a. 初始化, 设置参数 T , 令整数 $t=0$, 令变压器故障属性 $x_s=\arg \max_{1 \leq i \leq N}\{d(\{x_i\})\}$, 当前最高分类精度 $d_{\max}=d(\{x_s\})$, 将变压器故障属性子集 $\{x_s\}$ 作为一个结点加入到队列 Q 中;
- b. 当 $t < T$ 时执行步骤 c、d 和 e, 否则执行步骤 f;
- c. 取出 Q 的头结点 S_h (为一属性子集), 令 $added=false$ ($added$ 用来标志在对 Q 的头结点的扩展中, 是否向 Q 中加入新的结点), 对每一属性 $x \in X-S_h$, 如果 $S_h \cup \{x\}$ 没有被评价过, 且 $d(S_h \cup \{x\}) > d_{\max}$, 则令 $added=true$, $S_b=S_h \cup \{x\}$, $d_{\max}=d(S_h \cup \{x\})$, 以及 $t=t+1$, 并且将 S_b 作为一个新结点加入到队列 Q 中;
- d. 如果 $added=false$, 那么 $t \leftarrow t+1$;
- e. 转到步骤 b 继续执行;
- f. 在最终的变压器故障属性子集 S_b 上构建 RBC。

2 实验对比与实例分析

2.1 实验对比

为了与文献[9]的方法进行对比, 本文利用文献[9]整理搜集的变压器测试数据, 同样地, 取 200 个样本中的 150 个实例作为训练集, 50 个实例作为检测集, 构建了 SRBC 变压器故障诊断模型。

a. 属性信息缺失时选择性贝叶斯分类器与各种分类器诊断效果的比较。

工程实际中样本的属性信息经常是残缺的, 为了模拟属性残缺情况, 文献[9]将上述 50 个待检测样本(多数存在属性缺失现象)作为检测样本集的基础, 随机去掉 2~10 个属性值的测试样本各 500 个。本文利用相同的数据集, 建立 SRBC 模型对测试样本进行诊断, 同时与文献[9]所建立的 NB、TAN、BAN 模型和粗糙集决策表匹配方法进行比较, 比较结果见表 4。结果表明 SRBC 具有较强的处理不完备信息的能力, 其诊断效果优于 TAN、NB、BAN 及粗糙

表 4 属性不完备的测试样本的诊断结果
Tab.4 Results of diagnosis based on test samples with incomplete attributes

样本缺失属性数量	样本数量	正判率/%				
		NB 模型	TAN 模型	BAN 模型	粗糙集	SRBC 模型
0	50	100	100	100	100	100
1	210	100	100	100	100	100
2	500	100	100	100	99.4	100
3	500	100	89.6	100	97.8	100
4	500	99.6	89.0	99.6	96.4	99.7
5	500	98.2	86.0	98.2	90.0	98.3
6	500	89.2	66.6	89.6	78.0	92.1
7	500	76.2	26.2	79.0	63.6	83.5
8	500	59.6	24.6	65.0	49.6	71.2
9	500	45.4	22.0	48.6	29.2	55.9
10	500	36.6	18.8	37.8	14.8	43.6

集决策表匹配方法。

b. 与文献[9]所建立的 NB、TAN、BAN 粗集分类器模型诊断效果的比较。

为了与各种诊断模型在诊断效果方面进行比较,与文献[9]一样,本文在 50 个待检测样本的基础上形成了 30 个缺失一个关键属性的样本(关键属性个数为 3~5 个),诊断结果对比见表 5。

表 5 缺失一个关键属性的诊断结果

Tab.5 Results of diagnosis based on samples missing one key attribute

诊断方法	正判率/%
NB 分类器	46.67
TAN 分类器	22.67
BAN 分类器	46.67
NB 粗集分类器	80.00
TAN 粗集分类器	80.00
BAN 粗集分类器	80.00
SRBC	86.21

2.2 实例分析

某主变(SFSZ7-40000/110),所用有载开关为 ZYI 系列国产开关,1995 年 5 月 18 日安装投运,由于供电用户对电压合格率要求高,调压频率比较频繁,变压器 DGA 数据如表 6 所示。

表 6 实例的 DGA 数据

Tab.6 DGA data of case

$\varphi(H_2)$	$\varphi(CH_4)$	$\varphi(C_2H_6)$	$\varphi(C_2H_4)$	$\varphi(C_2H_2)$	$\varphi(CO)$	$\varphi(CO_2)$	φ (总烃)
76	12.6	6.34	6.1	8.93	635	4276	32.63

由于气体组分中 C_2H_2 占主要成分,且超出注意值,IEC 三比值码为 100,初步怀疑变压器内部存在高能量的放电故障。由表 6 和试验规程可知 $x_2=0$, $x_5=2$, $x_9=0$ 。油中微水含量为 $11.5 \mu L/L$, $x_4=0$ 。另外,现场停电对变压器进行全面电气检查试验,变压器绕组直流电阻、空载损耗、绕组及铁心绝缘、变压比、绕组泄漏电流等试验项目均无异常,所以 $x_1=0$, $x_7=2$, $x_{13}=0$ 。

使用 SRBC 模型和文献[9]中的模型分别进行故障诊断,结果如图 1 所示(图中, $i=0,1,\dots,12$)。很明显,在已知数据 E 的前提下 SRBC 模型、NB 粗集、TAN 粗集和 BAN 粗集模型计算所得的类 C_{12} 的后验概率最大,分别为 1.00、0.71、0.82、0.79;类 C_0 的后验概率分别为 0、0.0021、0.035、0.21。所以可判断该

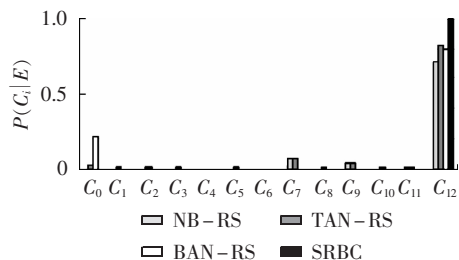


图 1 各种模型对实例的诊断结果

Fig.1 Results of diagnosis by different models for case

变压器故障为 C_{12} ,即有载分接开关油渗漏。与当时该变压器吊芯检查的结果相符。

3 结论

针对已有变压器故障诊断方法在信息缺失多、有偏差的情况下误判率高的问题,提出了用于变压器故障诊断的选择性贝叶斯分类器模型,该模型直接从不完整的变压器试验数据估算变量的概率分布,进而完成故障诊断。另外,该模型的结构和参数能够随着样本集的扩大而完善,因而具有学习功能。实验表明,该方法具有处理不完备信息的能力和容错特性,其性能明显优于其他贝叶斯网络分类器方法。

参考文献:

[1] WEI Shouzhi, JIN Ningde. Fault diagnosis system based on information fusion and embedded internet[C]//2007 IEEE International Conference on Integration Technology. Shenzhen, China: IEEE Computer Society, 2007: 203-207.

[2] WANG Chenhao, HUANG Huixian, XIAO Yewei, et al. Fault diagnosis of power transformers based on BP network with colonial selection algorithm[C]//Third International Conference on Natural Computation. Haikou, China: IEEE Computer Society, 2007: 13-16.

[3] JIA Honghong, DAI Wenzhan. Fault diagnosis against oil-immersed transformer based on PNN and GM(1,1)[C]//2007 IEEE International Conference on Integration Technology. Shenzhen, China: IEEE Computer Society, 2007: 394-397.

[4] BABU P, KALAVATHI S. Use of wavelet and neural network (BPFN) for transformer fault diagnosis[C]//2006 IEEE Conference on Electrical Insulation and Dielectric Phenomena. Kansas City, Missouri, USA: IEEE Electrical Insulation Magazine, 2006: 93-96.

[5] 董明, 屈彦明, 周孟戈, 等. 基于组合决策树的油浸式电力变压器故障诊断[J]. 中国电机工程学报, 2005, 25(16): 36-41.

DONG Ming, QU Yanming, ZHOU Mengge, et al. Fault diagnosis of oil-immersed power transformer using combinatorial decision tree[J]. Proceedings of the CSEE, 2005, 25(16): 36-41.

[6] NILSSON N. Artificial intelligence, a new synthesis[M]. San Francisco, USA: Morgan Kaufmann Publishers, 1998: 196-240.

[7] 王华伟, 周经伦, 何祖玉, 等. 基于贝叶斯网络的复杂系统故障诊断[J]. 计算机集成制造系统, 2004, 10(2): 230-234.

WANG Huawei, ZHOU Jinglun, HE Zuyun, et al. Fault diagnosis of complex system based on Bayesian networks[J]. Computer Integrated Manufacturing Systems, 2004, 10(2): 230-234.

[8] 李俭川, 胡莒庆, 秦国军, 等. 贝叶斯网络理论及其在设备故障诊断中的应用[J]. 中国机械工程, 2003, 14(10): 896-900.

LI Jianchuan, HU Niaoqing, QIN Guojun, et al. Bayesian network and its applications for device fault diagnosis[J]. China Mechanical Engineering, 2003, 14(10): 896-900.

[9] 吴立增. 变压器状态评估方法的研究[D]. 保定: 华北电力大学电气工程学院, 2005.

WU Lizeng. Assessing approach of transformer condition[D]. Baoding: North China Electric Power University, 2005.

[10] CHICKERING D M, HECKERMAN D, MEEK C. Large sample learning of Bayesian networks is NP-hard[J]. Journal of Machine Learning Research, 2004, 39(5): 1287-1330.

[11] CHEESEMAN P, STUTZ J. Bayesian classification (autoclass): theory and results[M]//FAYYAD U M, PIATETSKY-SHAPIRO

- G,SMYTH P,et al. Advances in knowledge discovery and data mining. Menlo Park,California,USA:AAAI Press,1996:153-180.
- [12] FRIEDMAN N. The Bayesian structural EM algorithm [C]//Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence. San Francisco,USA:Morgan Kaufmann Publishers, 1998:129-138.
- [13] LITTLE R J A,RUBIN D B. Statistical analysis with missing data[M]. New York,USA:Wiley,1987:120-149.
- [14] SPIEGELHALTER D J,COWELL R G. Learning in probabilistic expert systems[C]//Proceedings of the Fourth Valencia International Meeting on Bayesian Statistics. Oxford,England:Oxford University Press,1992:447-466.
- [15] RAMONI M,SEBASTIANI P. Robust Bayes classifiers[J]. Artificial Intelligence,2001,125(1-2):209-226.
- [16] 陈景年,黄厚宽,田凤占,等. 用于不完整数据的选择性贝叶斯分类器[J]. 计算机研究与发展,2007,44(8):1324-1329.
- CHEN Jingnian,HUANG Houkuan,TIAN Fengzhan,et al. Selective Bayes classifiers for incomplete data[J]. Journal of Computer Research and Development,2007,44(8):1324-1329.
- (实习编辑:李 莉)

作者简介:

赵文清(1973-),女,山西朔州人,副教授,博士,研究方向为智能技术在电力系统中的应用、数据挖掘(E-mail:jbwq@126.com)。

Transformer fault diagnosis based on selective Bayes classifier

ZHAO Wenqing

(North China Electric Power University,Baoding 071003,China)

Abstract: As the test data of electric power transformer fault diagnosis are incomplete and biased,a transformer fault diagnostic method is proposed based on selective Bayes classifier,which,with the ability to process uncertain information,combines DGA(Dissolved Gas Analysis) with other electrical test results. The transformer fault diagnosis model is built based on selective Bayes classifier. Its superiority in uncertain information processing is elaborated in detail. With the accumulated and improved training samples,it automatically modifies the parameters of network structure and probability distribution. Experimental results show that Bayes classifier is suitable for the transformer fault diagnosis.

This work is supported by the National Natural Science Foundation of Hebei Province(E2009001392) and the Chinese Universities Scientific Fund(09QG33).

Key words: transformer; fault diagnosis; Bayes network; selective classifier