

# 基于主成分分析法与贝叶斯网络的汽轮机故障诊断方法

韩 璞, 张德利, 韩晓娟, 焦嵩鸣

(华北电力大学 控制科学与工程学院, 河北 保定 071003)

**摘 要:** 在利用贝叶斯网络进行汽轮机故障诊断时, 汽轮机故障诊断模型的建立直接影响着诊断过程的复杂程度, 因此建立贝叶斯网络模型是首先要解决的问题, 反映故障状态的特征参数提取是建立模型的重要环节。通过对主成分分析方法提取故障特征的讨论, 提出了基于主成分分析方法和贝叶斯网络的汽轮机故障诊断模型建立方法, 并与传统的频率特征建模方法进行了比较。结果表明, 应用主成分分析法和贝叶斯网络建立的汽轮机故障诊断模型简洁, 易于推理, 提高了汽轮机故障诊断的效率。

**关 键 词:** 汽轮机; 故障诊断; 主成分分析法; 贝叶斯网络

中图分类号: TK268 文献标识码: A

## 引 言

汽轮发电机组结构及振动情况的复杂性和耦合性使其故障具有多层次性、随机性等特点, 难以通过理论分析的方法在故障原因与故障征兆之间建立对应关系<sup>[1]</sup>。近年来备受关注的贝叶斯网络对于解决复杂系统不确定因素引起的故障具有很大的优势, 取得了较好的诊断效果。

为了尽可能全面地刻画汽轮发电机组振动故障模式, 提高诊断精度, 常提取大量的故障征兆<sup>[2]</sup>, 这些征兆在故障诊断过程中的重要性并不相同, 某些征兆是次要的。次要信息的存在一方面是对诊断资源的浪费; 另一方面直接影响生成简洁高效的贝叶斯网络诊断模型, 影响了故障诊断的效率和实时性。

可见, 次要性是汽轮发电机组振动故障问题的一个普遍特征, 如果能提供一种解决问题的方法, 即使对包含次要信息的故障诊断数据仍能建立简洁高效的贝叶斯网络诊断模型, 则无疑是很有意义的。为了根据含有次要信息的故障诊断数据中建立简单有效的贝叶斯网络, 本文提出了基于主成分分析方法和贝叶斯网络的汽轮机故障诊断方法。实例表明, 应用该方法建立的贝叶斯网络模型具有简洁, 易

于推理, 减少了故障诊断的计算量, 缩短了故障诊断时间, 提高了汽轮机故障诊断的效率。

## 1 主成分分析法

主成分分析 (Principal Component Analysis, PCA) 方法由 Pearson 于 1901 年提出, 这种方法作为一种有效数学工具被归类于标准的多元统计方法中。PCA 的基本算法是要找一种空间变换方式, 让经标准化后的原始变量线性组合成若干个向量, 要求它们之间相互正交, 且第一个向量能反映样本间自变量的最大差异。也就是说, 样本集在该向量上的投影坐标是按照样本间自变量差异的大小确定的, 其它向量所反映的这种差异程度依次降低, 称这些向量为主成分 (Principal Component)。这种空间变换方式, 又称为 Karhunen-Loeve 变换, 简称 K-L 变换<sup>[3]</sup>。

### 1.1 主成分分析算法

设  $X$  是经标准化后含  $M$  个变量  $N$  个样本的本集。第  $k$  个主成分向量 (也称为主成分负载 loading) 为:

$$P_k = (p_{1k}, p_{2k}, \dots, p_{Mk})^T \quad k=1, 2, \dots, M \quad (1)$$

式中:  $P$ —主成分向量矩阵。  $X$  在  $P$  的投影为样本的主成分 (也称为主成分得分 score) 矩阵:

$$T = XP \quad (2)$$

主成分矩阵元素为:

$$t_{ij} = \sum_{k=1}^M p_{kj} x_{ik} \quad (i=1, 2, \dots, N; j=1, 2, \dots, M) \quad (3)$$

第  $j$  个主成分 ( $j=1, 2, \dots, M$ ):

$$T_j = \sum_{k=1}^M X_k p_{kj} \quad (4)$$

要求主成分之间正交, 即下式成立:

$$T_j T_{j'} = \sum_{k=1}^M \sum_{l=1}^M X_k X_l p_{kj} p_{l j'} = 0 \quad (j \neq j') \quad (5)$$

写成矩阵形式为:

$$T^T T = (XP)^T XP = P^T X^T XP = P^T DP = \Lambda \quad (6)$$

式中:

$$D = X^T X \quad (7)$$

是样本在原始空间的协方差矩阵。又:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_M \end{bmatrix} \quad (8)$$

是对角化矩阵。

满足式(6)有  $M(M-1)/2$  个条件。在增加下面的限制条件, 即  $P$  阵的列向量两两正交:

$$\sum_{i=1}^M p_{ij} p_{ik} = \delta_{jk} = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases} \quad (9)$$

写成矩阵形式为:

$$P^T P = P P^T = I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (10)$$

对式(6)两边乘  $P$  矩阵得:

$$P P^T D P = P A \quad (11)$$

应用式(10)得:

$$D P_j = P_j \Lambda \quad (12)$$

式(12)是一个线性代数中的本征值方程。 $D$  为本征矩阵。当  $D$  的秩是  $M$  时(即  $D$  矩阵的真正维数);  $P$  的大小为  $M \times M$  的本征向量矩阵;  $\Lambda$  是对应的本征值的对角化矩阵, 当只考虑第  $j$  个本征值时:

$$D P_j = \lambda_j P_j \quad (13)$$

或:

$$(D - \lambda_j I) P_j = 0 \quad (14)$$

为得到有意义的解, 要求:

$$|D - \lambda_j I| = 0 \quad (15)$$

解出  $\lambda_j$  后, 代入式(13), 可求出本特征向量  $P_j$ 。

解本征方程已有多种成熟的计算方法, 如 Jacobi、QL、NIPALS 等方法。

### 1.2 基于主成分分析法的故障样本特征提取

通过对振动信号的频率划分得到的特征空间为八维或九维形式, 传统的振动故障诊断系统都是基于这些八维或九维样本数据进行的<sup>[4-5]</sup>, 无论从计算的复杂程度还是分类器的性能来看都是不适宜的, 因此需要进行特征选择或特征提取。

这里使用主成分分析的方法进行样本的特征提取, 其目的是将含在原始特征值中有用的信息投影到几个复合特征变量上, 同时忽略多余的和不相干的信息。主成分的求解可化为求样本总体的协方差

矩阵的本征值和标准正交本征向量的问题, 归纳为如下 6 个步骤:

- (1) 数据标准化;
- (2) 求协方差矩阵;
- (3) 求协方差矩阵的本征值和本征向量;
- (4) 计算变量的主成分贡献率;
- (5) 计算训练样本的主成分;
- (6) 二维主成分映照。

## 2 贝叶斯网络及其建模方法

贝叶斯网络是用来表示不确定变量集合联合概率分布的图形模式, 是图形表示方法和概率知识的有机结合, 它反映了变量间潜在的依赖关系, 揭示了领域对象的内在联系, 是复杂全概率分布的紧凑表达方式。贝叶斯网络在专家系统中能够有效地协调双向推理, 并能很好地表示不确定信息, 同时其最大的特点在于它的因果关系组织, 并能实时的表示和适应不断变化的外界环境。一个贝叶斯网络包括一个节点集合, 所有的节点由一系列的有向连接弧连接起来构成一个完整的网络图形。两个节点之间的有向连接弧代表它们之间所存在的某种因果关系, 而因果关系的不确定性则用条件概率的形式加以量化<sup>[9]</sup>。

例如在汽轮机的故障诊断系统中, 转子不平衡故障将导致机组振动信号中的一倍频振幅加大。这样可以根据已有的知识定义两个节点: 一个是“转子不平衡(F1)”; 另一个是“一倍频振幅(C5)”, 同时由节点“Imbalance”指向节点“C5”建立一条有向连接弧, 箭头的方向表示由“原因”产生“结果”或由“故障”产生“故障征兆”。图 1 给出了几种常见汽轮机故障的贝叶斯网络诊断模型。

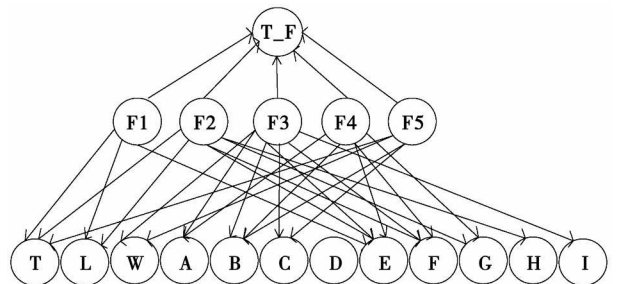


图 1 汽轮机故障的贝叶斯网络诊断模型

## 3 基于主成分分析方法和贝叶斯网络的汽轮机故障诊断

### 3.1 主成分分析用于汽轮机故障样本特征提取

从实验台中选取 4 组故障 14 个样本进行试验分析, 分别是不平衡、碰磨、裂纹和无故障。原始数据为九维的频率数据, 如表 1 所示; 原始数据为九维

的频率数据(频域征兆: A: (0.01 ~ 0.39) f1; B: (0.40 ~ 0.49) f1; C: 0.5 f1; D: (0.51 ~ 0.99) f1; E: 1 f1; F: 2f1; G: (3 ~ 5) f1; H: odd f1; I: > 5f1。其中 f1 为转速工频, odd f1 为奇数倍 f1。)经主成分变换之后的二维数据如表 2 所示, 得到的矩阵本征值与本征向量

如表 3 所示。图 2 表示两维最佳特征空间的模式向量投影, 其中圆形数据点为不平衡数据, 正十字形数据点为碰磨数据, 斜十字数据点为裂纹数据, 方形数据点为无故障数据, 类别可分性很直观, 因此选维特征作为最佳特征子集是完全可行的。

表 1 测试故障样本数据

征 兆									类 型
A	B	C	D	E	F	G	H	I	
0.002 56	0.001 22	0.009 93	0.018 26	0.811 23	0.079 04	0.049 58	0.049 58	0.003 9	不平衡
0.051 29	0.002 67	0.002 27	0.018 46	0.757 80	0.093 88	0.033 73	0.033 73	0.005 9	
0.004 93	0.001 62	0.001 31	0.010 49	0.841 74	0.052 99	0.019 62	0.019 62	0.003 2	
0.118 05	0.015 98	0.008 31	0.125 27	0.566 43	0.017 25	0.046 53	0.023 56	0.078 6	碰磨
0.030 12	0.012 75	0.021 75	0.169 04	0.612 79	0.019 77	0.056 57	0.025 18	0.052 0	
0.116 70	0.005 45	0.005 23	0.174 01	0.563 65	0.021 07	0.053 58	0.0128 8	0.053 4	
0.003 44	0.003 44	0.005 53	0.007 23	0.540 74	0.154 88	0.128 93	0.128 93	0.026 8	裂纹
0.001 78	0.001 78	0.003 23	0.005 66	0.580 58	0.156 24	0.114 22	0.114 22	0.022 3	
0.013 20	0.002 61	0.002 81	0.006 42	0.634 13	0.149 74	0.076 67	0.076 67	0.037 7	
0.024 75	0.182 73	0.392 01	0.196 42	0.057 36	0.096 57	0.022 54	0.022 54	0.005 1	无故障
0.004 82	0.240 00	0.505 75	0.072 14	0.085 49	0.0352 6	0.025 35	0.025 35	0.005 9	
0.023 63	0.144 73	0.539 38	0.102 11	0.052 16	0.091 17	0.020 69	0.020 69	0.005 4	
0.007 55	0.261 29	0.481 80	0.076 10	0.084 15	0.034 98	0.023 31	0.023 31	0.005 5	
0.013 21	0.233 94	0.488 00	0.063 58	0.099 38	0.038 41	0.027 77	0.027 77	0.007 9	

表 2 主成分变换后的二维数据

第一主成分	第二主成分	序号	第一主成分	第二主成分
1.158 7	-0.527 8	8	2.929 7	-1.573 9
0.822 5	0.189 5	9	2.216 1	-0.664 9
0.416 9	-0.042 4	10	-2.378 9	-0.494 6
0.084 7	3.448 8	11	-2.569 2	-1.209 0
-0.004 8	1.942 3	13	-2.193 6	-1.000 7
-0.225 5	3.250 2	13	-2.678 5	-1.161 0
3.204 5	-1.626 1	14	-2.383 0	-1.104 2

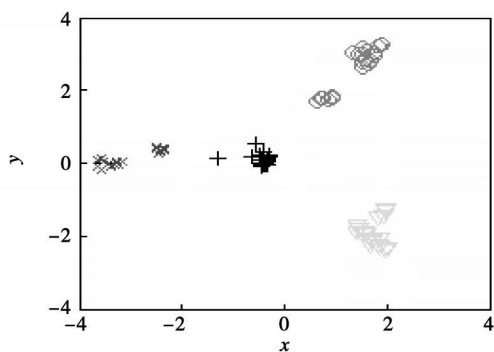


图 2 两维最佳特征空间的模式向量投影图

表 3 样本总体散布矩阵本征值与本征向量

本征值	本征向量	贡献率/%
3.995 7	-0.022 2, -0.412 9, -0.411 0, -0.270 1, 0.393 9, 0.312 3, 0.407 1, 0.382 0, 0.150 8	46.39
2.670 4	0.532 8, -0.238 5, -0.253 0, 0.348 2, 0.210 6, -0.354 3, -0.082 3, -0.277 9, 0.469 2	31.00
1.030 0	0.161 7, 0.177 6, 0.174 5, 0.355 6, -0.503 0, 0.204 5, 0.453 3, 0.349 5, 0.403 6	11.96
0.315 1	0.046 8, -0.238 9, 0.007 7, 0.637 0, 0.128 6, 0.600 1, -0.149 1, -0.123 9, -0.347 3	3.66
0.257 9	-0.539 6, 0.200 4, 0.169 8, 0.439 3, 0.516 9, -0.312 3, 0.247 1, 0.093 7, 0.111 3	2.99
0.221 4	-0.488 5, -0.339 9, -0.525 0, 0.204 4, -0.507 1, -0.246 2, 0.013 8, -0.083 5, -0.058 1	2.57
0.089 1	-0.394 3, -0.097 7, 0.154 1, -0.131 0, -0.004 2, 0.379 1, -0.390 7, -0.210 4, 0.673 4	1.03
0.029 2	0.015 6, 0.700 1, -0.635 0, 0.072 6, 0.081 2, 0.193 3, -0.213 5, 0.085 9, 0.064 9	0.34
0.004 2	-0.041 3, 0.169 4, -0.071 2, -0.127 9, -0.001 5, 0.189 2, 0.583 4, -0.756 2, -0.012 6	0.05

### 3.2 应用特征提取后的样本建立汽轮机故障贝叶斯网络模型

将图 1 贝叶斯网络中的频率特征节点替换为第一主成分 M1, 第二主成分节点 M2, 得到新的贝叶斯网络模型如图 3 所示。

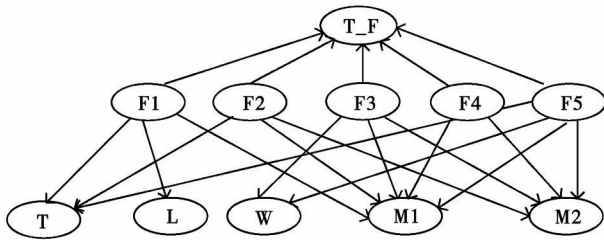


图 3 基于主成分分析方法和贝叶斯网络的汽轮机故障诊断模型

由于篇幅原因, 本文只讨论了几种典型汽轮机故障类型和相关的部分特征征兆。一个实际的较完整的汽轮机故障诊断模型要复杂的多。而汽轮机所有的特征一般包括频域特征、时域特征、热力特征以及其它特征等, 而每种特征几乎都存在主次之分。因此可通过主成分分析法, 将其次要信息剔除, 得到主要成分。

### 3.3 实例分析与讨论

某汽轮机组通过汽轮机的监控系统获得以下事实:

- (1) 振动信号的低倍频振幅增大;
- (2) 振动信号的一倍频和二倍频振幅增大;
- (3) 振动波形紊乱。

采用变量消元法对图 1 和图 3 所示的贝叶斯网络模型进行推理诊断, 其诊断结果如表 4 所示:

表 4 诊断结果

	诊断结果 / %					所需存储空间
	不平衡	裂纹	碰磨	不对中	轴承松动	
图 1	8.12	3.71	83.05	7.5	5.12	64kb
图 3	10.68	1.22	85.1	0.95	2.05	38kb

从诊断结果可以看出, 碰磨的可能性最大, 且明显高于其它故障类型。因此诊断结果为碰磨, 而本机组的实际诊断结果也是碰磨, 诊断结果与实际相符, 并且图 3 的诊断结果要好于图 1 的诊断结果。

贝叶斯网络的各种推理算法的复杂度一般为  $O(n^2)$  或者  $O(n)$ , 其中  $n$  是网络的节点个数, 因此一个贝叶斯网络的节点个数的多少从根本上影响着推理算法的复杂度。比较图 1 和图 2 2 个汽轮机故障诊断贝叶斯网络模型可以看出, 对于同样的汽轮

机故障诊断, 图 2 的节点个数明显小于图 1 的节点个数。因此通过使用主成分分析方法建立的汽轮机故障诊断贝叶斯网络模型能很好的降低诊断的开销, 缩短诊断时间, 提高诊断效率。

## 4 结 论

汽轮机故障诊断对电厂的安全运行有着十分重要的意义。本文基于智能互补融合的思想, 将主成分分析方法与贝叶斯网络有机结合在一起进行汽轮机故障诊断, 利用主成分分析法特征信息提取技术来实现对专家知识的简化与故障特征的压缩, 获得最小诊断规则, 基于最小诊断规则的贝叶斯网络模型可以有效降低网络结构的复杂性与故障特征获取的难度。同时利用贝叶斯网络实现概率推理, 便于描述故障特征的变化及对汽轮机故障原因的快速分析。故障实例分析表明, 文中所提出的方法能很好的降低诊断的开销, 缩短诊断时间, 提高诊断效率。

### 参考文献:

- [1] 陈长征, 刘 强. 概率因果网络在汽轮机故障诊断中的应用[J]. 中国电机工程学报, 2001, 21(3): 78-81.
- [2] 杨 萍, 吴 捷, 冯永新. 200 MW 汽轮发电机组振动故障的模糊诊断系统[J]. 电力系统自动化, 2001, 25(10): 45-49
- [3] 陈念胎, 钦 佩, 陈瑞亮, 等. 模式识别方法在化学化工中的应用[M]. 北京: 科学出版社 2000.
- [4] PLATT J, CRISTIANINI N, SHAWE TAYLOR J. Large margin DAG's for multi-class classification // Advances in neural information processing systems[C]. Massachusetts: MIT Press, 2000; 547~553
- [5] 李 勇, 陈梅倩. 汽轮机运行性能诊断技术及其应用[D]. 北京: 科学出版社, 1999
- [6] 华 斌, 周建中, 喻 菁. 贝叶斯网络在水电机组故障诊断中的应用研究[J]. 华北电力大学学报, 2004, 31(5): 33-36

(编辑 伟)

### · 书 讯 ·

## 《工程技术公式速查手册》

本书内容涵盖单位、平面图形、立体、计算、三角函数、解析几何、统计学、微分学、积分学、微分方程、静力学、运动学、动力学、水力学、热学、材料力学、机械零件、生产工程、电气工程、控制工程、化学、辐射物理等 21 个板块, 涉及 13 类专业领域。读者对象: 工程技术人员, 高等院校相关专业师生。

2008 年 2 月出版

火电厂直接空冷系统风洞热效应模拟实验研究 = **Simulation and Experimental Study of the Wind-tunnel Thermal Effect of a Directly Air-cooled System in a Thermal Power Plant** [刊, 汉] / LIU Pei-qing, ZHAO Wan-li (Education Ministry Key Laboratory on Hydromechanics, Beijing University of Aeronautics and Astronautics, Beijing, China, Post Code: 100083), XU Ze-lin (China Guodian Scientific Environment Protection Co. Ltd., Beijing, China, Post Code: 100034) // Journal of Engineering for Thermal Energy & Power. — 2008, 23(3). — 240 ~ 243

With constant-temperature hot water passing through an air-cooled condenser model to form a thermal circulation system, simulated simultaneously for the first time in a low speed wind tunnel was the coupling of complex flow with thermodynamics, including aerodynamics, forced convection in an axial compressor and buoyant effect of hot air. A concept of return flow rate that influences the heat dispersion effectiveness of the air-cooled condenser was presented and the influence of the change of coming flows on the averaged return flow rate at the bottom of the air-cooled island explored. In addition, PIV (Particle Image Velocimetry) velocity-measuring experiments were performed accordingly. It has been found that the average return-flow rate in question increases with an increase in air speed, assumes a basically identical variation tendency with the change in air flow direction at different air speeds. The above-mentioned return flow rate is comparatively large when the air flow direction is approximately west, southwest and northwest. The average return flow rate in question will decrease with an increase of the height of the windbreak wall and attains a minimum value when the windbreak wall is 14.64 m tall. **Key words:** direct air-cooling, return flow rate, PIV (Particle Image Velocimetry) experiment, measures aimed at improvement

基于主成份分析法与贝叶斯网络的汽轮机故障诊断方法 = **Steam Turbine Fault Diagnosis Methods Based on the Main Constituent Analysis Method and Bayesian Network** [刊, 汉] / HAN Pu, ZHANG De-li, HAN Xiao-juan, JIAO Song-ming (College of Control Science and Engineering, North China Electric Power University, Baoding, China, Post Code: 071003) // Journal of Engineering for Thermal Energy & Power. — 2008, 23(3). — 244 ~ 247

When Bayesian network is used to diagnose a fault, the establishment of a model for fault diagnosis of steam turbines has a direct bearing on the complexity of the fault diagnosis process. Therefore, to establish a model of Bayesian network becomes an issue of first priority and the collection of characteristic parameters reflecting the fault status constitutes an important link for setting up a model. Through a discussion of the collection of fault characteristics by using the main constituent analysis method, presented was the modeling method for steam turbine fault diagnosis based on the main constituent analysis and Bayesian network. In addition, the proposed method has been compared with the traditional frequency characteristics modeling method. The results show that the model in question for turbine fault diagnosis is simple and lends itself to easy reasoning, thus enhancing the efficiency of turbine fault diagnosis. **Key words:** steam turbine, fault diagnosis, main constituent analysis method, Bayesian network

船用正车油动机电液伺服系统改进 = **Electrohydraulic Control-based Improvement of a Marine Forward-drive Oil Servomotor** [刊, 汉] / XU Sheng-hang (Military Representative Office of Chinese Navy Resident at Harbin Turbine Works Co. Ltd., Harbin, China, Post Code: 150046), WANG Chong, LUAN Hai-feng, LIU Jiang (No. 703 Research Institute of CSIC, Harbin, China, Post Code: 150036) // Journal of Engineering for Thermal Energy & Power. — 2008, 23(3). — 248 ~ 251

To overcome the shortcomings of the original forward-drive oil servomotors, such as slow response speed and easy seizure of the electrohydraulic conversion portion, a comprehensive evaluation method was employed to perform a type selection of electrohydraulic converters, focusing on reliability and maintainability etc. An electrohydraulic control-based improvement scheme was finalized by using a sourceless self-capacitance type electrohydraulic actuator ReXA, which has a large output power and strong anti-seizure capacity. Being fully enclosed and capable of avoiding erosion of equipment internals by salt mist in the engine room, it is suited for use on board vessels. After the improvement, the transition duration of the forward-drive oil servomotor is shortened from the original 40 seconds to 18.2 seconds and even to a minimum of 3.3 seconds depending on specific needs. **Key words:** oil servomotor, electrohydraulic servo, improved design