DOI: 10.3901/JME.2020.04.035

基于多维高斯贝叶斯的机械设备失效/ 故障智能诊断及参数影响分析*

朱建新^{1,2} 陈学东^{1,2} 吕宝林^{1,2} 王溢芳^{1,2} 乔松^{1,2} 陈嘉宏^{1,2}

(1. 合肥通用机械研究院有限公司 合肥 230031;
2. 国家压力容器与管道安全工程技术研究中心 合肥 230031)

摘要: 采用多维高斯贝叶斯机器学习方法研究机械设备的失效/故障智能诊断方法, 基于极大似然分析基础上, 提出一种基于“马氏距离”估算的关键特征参量影响分析方法, 用于评价各特征参量对失效/故障分类的影响。该方法用于两个机械设备的失效/故障数据库的智能诊断分析, 在获得高诊断识别率(失效/故障模式正确识别率分别达到 96%和 86%)的同时, 也识别了影响失效/故障分类的关键特征参量。分析表明特定的失效/故障模式往往取决于少数关键特征参量, 而不确定的失效/故障模式的关键特征参量往往呈分散分布, 关键特征的分散性会影响多维高斯贝叶斯分类器的诊断识别率。该方法可用于机械设备的失效/故障的智能识别与关键特征参量的智能诊断, 也为失效/故障的影响因素分析指明方向。

关键词: 高斯贝叶斯; 失效/故障诊断; 特征参量; 影响因素; 马氏距离

中图分类号: TG156

Smart Failure/Fault Diagnosis and Influence Analysis for Mechanical Equipment with Multivariate Gaussian Bayesian Method

ZHU Jianxin^{1,2} CHEN Xuedong^{1,2} LÜ Baolin^{1,2} WANG Yifang^{1,2}
QIAO Song^{1,2} CHEN Jiahong^{1,2}

(1. Hefei General Machinery Research Institute Co. Ltd., Hefei 230031;

2. National Technology Research Center for Safety Engineering of Pressure Vessels and Pipelines, Hefei 230031)

Abstract: The failure/fault of mechanical equipment is diagnosed with multivariate Gaussian Bayesian classifier. Based on the maximum likelihood methodology, a novel influence analysis model based on "Mahalanobis distance" estimation is proposed. The method is then applied to two datasets for the mechanical equipment failure/fault mode identification. The results show that the proposed method obtains high diagnostic recognition rate (failure/fault mode recognition rate in two cases are 96% and 86%, respectively), as well as principal attributes that contribute to specific failure/fault modes. It is found that the specific failure/fault mode mainly depends on a few characteristic parameters, while the unspecified failure/fault mode always involves several diverse characteristic parameters. The desperation of key parameters will cause the unsatisfactory result of multivariate Gaussian Bayesian classifier. The model proposed in this paper is helpful for the intelligent diagnose of failure/fault mode of mechanical equipment and the analysis of key parameters that contribute to specific failure/fault mode, and so as to provide guidance for failure/fault reasoning.

Key words: Gaussian Bayesian; failure/fault diagnosis; characteristic parameter; influence analysis; Mahalanobis distance

0 前言

在流程工业设备运维领域, 经过百余年的发展, 设备运维技术已经逐渐由最早的事故后维修、预防

性维修发展到以可靠性为中心的维修, 相应地装备的可靠性也得到长足发展^[1-3]。无论是以耗散为特征的承压类设备的失效^[4], 还是以随机性为特征的转动类设备的失效, 设备的失效都有特定的规律。一般认为设备的失效过程是随时间变化的, 随着时间的增加, 设备或部件发生状态劣化(如裂纹萌生、扩展到开裂)、发生潜在失效直至最终功能失效。

设备故障诊断的目的, 一是依据设备当前运行发生的现象, 分析已经发生的失效或故障, 预测设

* 工信部智能制造综合标准化项目、安徽省科技重大专项(17030901014)、国机集团重大科技专项(SINOMAST-ZDZX-2017-01-05)和国家重点研发计划(2017YFF0207904)资助项目。20191021 收到初稿, 20200107 收到修改稿

备继续运行可能发生的失效或故障,从而为预知维修提供依据;二是通过建立因果分析或数学推理,查找引发失效或故障的最可能因素,从而为失效/故障的诊断与修复提供依据。由一种现象的发生预测另一种现象发生的可能性,其实质就是贝叶斯理论所述的条件概率。贝叶斯理论在失效/故障诊断中具有广泛应用,其中以朴素贝叶斯为基础的分类算法和贝叶斯网络最为著名。朴素贝叶斯的优点在于仅从统计意义而非机理上研究事件发生的可能性,这种数据驱动的分析方法在语音识别、邮件分类及设备故障诊断等领域具有广泛应用^[5]。贝叶斯网络则需要从事件发生的机理出发来构建网络,对于组成复杂、机理认识不明的系统,还需要研究复杂的算法来进行网络构造,因而其大多应用于失效机理明确的设备^[6]。文献[7]将朴素贝叶斯与神经网络相结合,从水力、机械、电气三方面构建了水电机组振动故障诊断模型,用于多因素耦合的复杂系统设备故障诊断。文献[8-9]介绍了贝叶斯网络用于转子、搅拌器等设备的故障诊断,表明贝叶斯网络可以描述设备故障与征兆间的因果关系。文献[10]以主观贝叶斯推理为基础,提出了一种基于不确定性推理的多传感器分布式检测融合的设备故障诊断方法。

近年来随着数据科学技术的发展,用数据驱动的方法进行复杂系统设备失效/故障诊断由于具备机器学习的特点,使得其应用日益广泛。巴西学者将朴素贝叶斯与定性趋势分析方法相结合用于输电线路的复合故障诊断,对分类故障诊断具有很高的精度^[11]。文献[12-13]报道了采用朴素贝叶斯结合小波分析用于机泵故障诊断的案例。

在设备失效/故障原因分析方面,传统的分析方法从失效机理出发,通过寻找机械设备外在特征与诱发失效的内在机制的对应关系来进行失效/故障原因分析,如采用断口形貌、腐蚀形态、材料力学性能等来进行失效原因分析,采用振动信号的频谱来分析设备故障等^[14-16]。这种诊断方法对机理清楚、失效影响因素简单的失效形式往往十分有效。然而对运行工况或结构复杂,影响因素多的设备的失效/故障的原因分析,这种方法还存在困难。

从国内外研究现状看,有关朴素贝叶斯分类器及贝叶斯网络在设备故障诊断中的应用已有不少学者进行研究。然而对于影响因素复杂且存在强耦合相关的设备失效/故障诊断问题还有待进一步研究。本文针对机械设备强耦合关联的特点,研究了多维高斯贝叶斯模型用于机械设备的失效分析/故障诊断技术,并在极大似然估计的基础上,提出了一种

基于“马氏距离”估算的设备故障影响分析方法。利用上述方法,以意大利都灵大学(University of Torino, Italy)公开发布的机械设备故障诊断数据库(Mechanical analysis data set, <https://archive.ics.uci.edu/ml/datasets/Mechanical+Analysis>)和意大利传播科学研究中心(Research center of sciences of communication)发表的钢板缺陷数据库(Steel plates faults data set, <https://archive.ics.uci.edu/ml/datasets/Steel+Plates+Faults>, Semeion)为例^[17-18],对提出的方法进行验证。

1 多维高斯贝叶斯分类模型

1.1 多维高斯贝叶斯原理

贝叶斯分类器(Bayesian classifier)是一种基于贝叶斯定理的分类算法,它利用已有的分类样本建立贝叶斯模型,确保同一分类中各影响因素(特征)的联合分布概率达到最大的一种算法,如下式所示

$$P(X|C_k) = \frac{P(C_k|X) \cdot P(X)}{P(C_k)} \quad (1)$$

式中, X 为影响因素矢量,它由 N 个相关或无关的影响因素组成, $X = (x_1, x_2, \dots, x_i, \dots, x_N)^T$ 。 C_k 为设备运行状态集 C 的一种 ($C_k \in C$), $C = \{C_1, C_2, \dots, C_i, \dots, C_K\}$, 用于设备的故障状态分类(失效、故障或工作正常)。

当式(1)中的影响因素相互独立时,称为朴素贝叶斯模型,计算上可进行简化,当影响因素不相互独立时,需要进行影响因素的相关性分析。特殊地,对某一失效/故障模式 C_k , 若对任一影响因素 x_i , $P(x_i|C_k)$ 服从高斯分布,则式(1)称为高斯贝叶斯。如对某特定的失效模式(如应力腐蚀开裂),假设多个影响参数中任一影响参数(如 Cl^- 离子含量)对该失效模式的影响呈高斯分布,则称该模型服从多维高斯贝叶斯分布。依据中心极限定理,客观世界中大量随机变量都近似服从高斯分布^[19],因而高斯贝叶斯模型具有广泛的应用范围。

对于同属于 C_k 分类的全部 N_k 个样本,可以通过式(2)计算隶属于 C_k 的 N_k 个样本的联合分布概率密度^[20]

$$P(x_1, x_2, \dots, x_N | C_k) = \frac{1}{\sqrt{(2\pi)^N |\Sigma_k|}} \exp\left(-\frac{1}{2}(X - \mu_k)^T \cdot \Sigma_k^{-1} \cdot (X - \mu_k)\right) \quad (2)$$

式中, $\mu_k = (\mu_1, \mu_2, \dots, \mu_N)^T$, 为 C_k 分类中各影响因素的期望值组成的矢量, Σ_k 为 C_k 分类中各影响参

数组成的协方差矩阵。上式对 $\mathbf{X}$ 中的各影响因素无相互独立的要求, 因而更适合实际的故障诊断。

式(2)取对数, 可得

$$\ln L = \ln(P(x_1, x_2, \dots, x_N | C_k)) = -\frac{1}{2} [\ln(|\Sigma_k|) + (\mathbf{X} - \mu_k)^T \cdot \Sigma_k^{-1} \cdot (\mathbf{X} - \mu_k) + N \ln(2\pi)] \quad (3)$$

依据极大似然理论, 对某一分类 C_k , 当待检样本中各影响因素的联合分布概率密度达到最大值时, 表明该样本属于 C_k 的可能性最大, 从而判定待检样本最可能隶属的分类。

1.2 影响的关键性分析

从式(3)可以看出, 影响极大似然值的主要因素一是某一分类的各影响因素协方差矩阵的行列式值, 另一项是 $(\mathbf{X} - \mu_k)^T \cdot \Sigma_k^{-1} \cdot (\mathbf{X} - \mu_k)$ 。后一项的实质是待检样本中影响因素归一化后构成的矢量 $(\mathbf{X} - \mu_k)^T$ 与该矢量依据符合 C_k 分类的多维高斯分布反向投影后得到的矢量 $\Sigma_k^{-1} \cdot (\mathbf{X} - \mu_k)$ 之间的马氏距离(Mahalanobis distance)。

考虑到对特定的分类, $\ln(|\Sigma_k|)$ 和 $N \ln(2\pi)$ 均为常数, 实际上影响 $\ln L$ 大小的是马氏距离, 故令 $y = (\mathbf{X} - \mu_k)^T \cdot \Sigma_k^{-1} \cdot (\mathbf{X} - \mu_k)$, 对于任一影响因素, 有

$$\frac{\partial y}{\partial x_i} = \frac{\partial (\mathbf{X} - \mu_k)^T}{\partial x_i} \cdot \Sigma_k^{-1} \cdot (\mathbf{X} - \mu_k) + (\mathbf{X} - \mu_k)^T \cdot \Sigma_k^{-1} \cdot \frac{\partial (\mathbf{X} - \mu_k)}{\partial x_i} \quad (4)$$

若协方差矩阵的逆阵 Σ_k^{-1} 中第 i 行第 j 列元素为 $\sigma_{i,j}$, $\mathbf{X} - \mu_k$ 中第 j 个元素记为 x_j , 则对任一分类, 省略分类编号字母 k 后, 式(4)可简化为

$$\frac{\partial y}{\partial x_i} = \sum_{j=1}^N \sigma_{i,j} \cdot (\sigma_{i,j} + \sigma_{j,i}) \quad (5)$$

式(5)反映了影响分类判定极大似然函数值的各特征参量中, 特征参量 x_i 变化对极大似然函数的影响。若 $\left| \frac{\partial y}{\partial x_i} \right|$ 越大说明该特征参量 x_i 对极大似然函数的影响越大, 从原理上说明该参数是影响样本是否属于该分类的最关键因素。对于机械设备而言, 该特征参量很可能是表征该失效/故障模式的关键因素, 这为后续开展故障诊断指明了方向。

2 应用案例验证

2.1 验证数据库简介

为验证上述方法的效果, 采用公开发表的两个数据库对上述算法进行了验证。

机械分析数据库是一个由意大利都灵大学 1990 年收集的数据库, 收录了某机电设备 164 个运行样本的 6 种故障模式(含 1 个无故障模式)数据, 包括连接件故障、轴承故障、机械松动、地基变形、不平衡和正常运行状态(表 1)。为开展设备的故障诊断, 对其进行了检测部件不固定(最多包含 78 个部件)、单个部件的检测项目不固定的检测, 检测项目为速度、振动等 10 类检测项目中的一种(表 2), 相同部件在不同样本中的检测项目也不固定, 因而该数据库不是一个传统意义上“结构化”的数据库。不同设备故障模式的样本数量及其比例如表 1 所示。

表 1 机械分析数据库中设备故障模式及比例

序号	故障模式分类	故障模式代码	样本数量	比例(%)
1	连接件故障	1	46	28.05
2	轴承失效	2	50	30.49
3	机械松动	3	8	4.88
4	基础变形	4	18	10.98
5	不平衡	5	16	9.76
6	无故障/正常运行	6	26	15.85

表 2 机械分析数据库中涉及的不同检测项目及比例

序号	检测项目	检测项目代码	检测项目数量	比例(%)
1	水平方向速度, 无滤波	vo	628	8.99
2	轴向速度, 无滤波	va	335	4.80
3	垂直方向速度, 无滤波	vv	433	6.20
4	水平振幅, 无滤波	ao	629	9.00
5	轴向振幅, 无滤波	aa	335	4.80
6	垂直方向振幅, 无滤波	av	434	6.21
7	水平方向速度, 有滤波	io	1 916	27.43
8	轴向速度, 有滤波	ia	882	12.63
9	垂直方向速度, 有滤波	iv	1 230	17.61
10	未明确定义	ma	164	2.35

机械分析数据库中数据的分散性很大, 图 1 为数据库不同样本中检测项目的分布, 横坐标为样本编号, 纵坐标为不同的检测项目编号, 图中数据标志点越大代表检测频次越高, 反之越低。从图中可

图 1 机械分析数据库不同样本的数据类型分布

以看出,除少数检测项目外,全部样本中部分检测项目如 iv、ia、io、av 和 vv 的检测频次非常分散。

钢板缺陷监测数据库是意大利传播科学研究中心收集的,库中记录了包括钢板麻点、Z-划伤、K-划伤、锈点、污渍、凸点及其他缺陷等 7 种钢板缺陷形式(失效模式),这些缺陷通过对钢种、反射亮度、厚度、污渍面积等 27 个影响因素的监测进行判断,整个数据库共包括 1 941 个样本。钢板缺陷监测数据中 7 个失效模式的分布如表 3 所示。

表 3 钢板缺陷监测数据库中失效模式及比例

序号	钢板失效模式	失效模式代码	样本数量	比例(%)
1	麻点	1	158	8.14
2	Z-划伤	2	190	9.79
3	K-划伤	3	391	20.14
4	锈点	4	72	3.71
5	污渍	5	55	2.83
6	凸点	6	402	20.71
7	其他缺陷	7	673	34.67

与机械分析数据库相比,钢板缺陷监测数据库是一个结构化的数据库,样本中监测的影响参数固定且数据完整。依据本文提出的算法对上述设备运行模式和钢板缺陷进行了智能诊断及影响因素分析。

2.2 数据的结构化预处理

为确保上述数据用于模式分类,对数据进行预处理。对于机械分析数据库,依据样本编号和检测项目对每一部件的检测结果进行重新排列,形成新的结构化样本数据集。依据部件编号及检测类别分别填入相应的检测数据,全部数据集是一个 164×780 的矩阵。对于钢板缺陷监测数据库,由于本身是一个结构化数据库,因而预处理过程相对简单,主要是对数据中各影响因素的相关性进行分析。

对于机械分析数据库,结构化预处理后的样本含有 780 个参量,考虑到检测的部件和检测项目不确定,780 个参量绝大多数是无检测数据的,而钢板缺陷监测数据库情况稍好,样本中各参量的数据是完整的。当样本中参量存在大量的数据缺失,或参量之间存在强烈的相关性(如机械分析数据库涉及振动的检测数据可能因振动的传递而存在强相关),求解过程得到的协方差矩阵可能为奇异矩阵导致无法求解,需要进行降维。常见的降维方法包括奇异值分解、拉普拉斯映射或主成分分析等,此处不再赘述。本文利用主成分分析方法,在保留一定比例的特征值的前提下,应用主成分分析方法使机械分析数据库的特征参量由 780 维降至 8 维,使钢板缺陷监测数据库的特征数量由 27 维降至 23 维,通过减少特征的数量,使多维高斯贝叶斯模型应用成为可能。

3 诊断结果分析

3.1 机械分析数据库诊断结果

利用本文提出的智能诊断方法,对机械分析数据库中的样本进行了训练与检验,将全部数据集集中的前 70%作为训练集(共 114 个)、后 30%作为检验数据集(共 50 个)对模型进行训练和检验。

考虑到设备故障模式分析存在一定的相似性,为检验诊断效果,对似然函数最大的首选分类以及似然函数值次之的次选分类进行了考察。以首选分类似然函数值 $\ln L_1$ 与次选分类的似然函数 $\ln L_2$ 之差与首选分类的似然函数之比 φ 来评估两种故障模式的相似程度

$$\varphi = \frac{\ln L_1 - \ln L_2}{\ln L_1} \quad (6)$$

式中, φ 为无符号量,显然 φ 值越小,说明两个故障模式越接近。

利用建立的多维高斯贝叶斯模型对机械分析数据库中设备的故障模式进行了检验,检验集中 50 个样本的首选故障模式与次选故障模式分析结果如表 4 所示。

表 4 机械分析数据库的故障模式分类检验结果

样本 编号	实际 故障 模式	首选 故障 模式	次选 故障 模式	φ	样本 编号	实际 故障 模式	首选 故障 模式	次选 故障 模式	φ
115	1	4	1	0.18	140	1	1	4	0.01
116	4	1	4	0.31	141	2	2	6	0.20
117	1	1	5	0.10	142	2	2	1	0.09
118	1	5	1	0.05	143	2	1	5	0.05
119	1	1	4	0.02	144	2	6	2	0.02
120	1	1	4	0.13	145	2	6	2	0.16
121	1	5	1	0.01	146	2	6	2	0.08
122	1	1	4	0.37	147	2	6	2	0.05
123	1	5	1	0.07	148	2	6	2	0.23
124	1	6	1	0.09	149	2	6	2	0.30
125	1	1	5	0.13	150	2	2	1	0.06
126	1	1	5	0.11	151	2	2	6	0.08
127	1	1	5	0.31	152	2	6	2	0.20
128	4	4	1	0.20	153	2	6	2	0.14
129	1	4	1	0.25	154	1	5	1	0.03
130	1	1	4	0.28	155	2	6	2	0.32
131	1	1	5	0.05	156	2	2	6	0.09
132	4	4	1	0.79	157	2	6	2	0.27
133	1	1	4	0.13	158	2	2	1	0.20
134	4	4	1	0.37	159	2	6	2	0.13
135	1	1	2	0.02	160	2	6	2	0.07
136	1	1	5	0.03	161	2	2	1,4	0.03
137	1	1	5	1.00	162	2	5	1	0.13
138	1	4	1	0.17	163	2	6	2	0.12
139	1	1	5	0.04	164	2	2	6	0.01

从表 4 可以看出, 除 137 号样本的首选与次选似然函数之差达到首选似然函数的 1 倍外, 其余的相差绝大多数都在 0.5 以内, 显示首选与次选分类的相似度很高。特别地, 编号为 144~149, 152~155 的样本, 虽然直到次选才正确识别故障模式, 但两者的似然函数差值的比值都小于 0.3, 显示首选与次选分类能有效识别出设备的故障模式。在 50 个检验数据集中, 仅编号 143 和 162 的样本未能在正确识别。这样检验数据集中, 全部 6 个故障模式进行诊断的首选正确识别率为 52%, 首选和次选正确识别率为 96%。

依据影响因素分析方法, 降维后同一故障模式中影响因素最为强烈的特征及其识别情况进行了分析, 如对于编号为 115 的样本, 降维后马氏距离对不同影响参数的偏导数

$$\left(\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \dots, \frac{\partial y}{\partial x_8} \right) = (1.09, 0.36, 0.85, 0.76, -0.54, -0.03, -0.61, -0.19) \quad (7)$$

可见 $\left| \frac{\partial y}{\partial x_i} \right|$ 最大的是第 1 个参量(记为关键特征参量), 其次为第 3 个参量。

把隶属于同一故障模式的样本中关键特征参量相同的样本所占的比例计算出来, 并对不同的故障模式的首选与次选识别率分别进行计算, 结果如表 5 所示。

表 5 不同故障模式的关键特征参量比例及识别率

特征参量	故障模式		
	1	2	4
关键特征参量 1	6	1	2
比例	0.39	0.52	0.50
关键特征参量 2	7	7	7
比例	0.26	0.30	0.25
关键特征参量 3	8	—	8
比例	0.17	—	0.25
关键特征参量 4	5	—	—
比例	0.08	—	—
其他关键特征参量	3	2,3,6,8	—
比例	0.08	0.18	—
样本总数	23	23	4
首选识别率	0.65	0.35	0.75
次选识别率	0.35	0.57	0.25
未识别率	0	0.08	0

注: 依据数据库中数据的实际情况, 后 50 个检验样本中仅存在分类为 1、2、4 的样本, 故表中未列出其他分类的检验结果。

从上表可以看出, 总体而言, 故障模式 1 和故障模式 4 的首选+次选正确识别率达到 100%, 故障模式 2 的首选+次选正确识别率为 92%。分析其原

因, 除基础变形(故障模式 4)外, 连接件故障(故障模式 1)和轴承失效(故障模式 2)的关键特征参量较为分散, 这可能与设备中有多个连接件或轴承有关, 即当存在多个连接件或轴承时, 不明确指出是哪个连接件故障的故障模式分类方法不利于故障的准确诊断。

3.2 钢板缺陷监测数据库诊断结果

对钢板缺陷监测数据库而言, 全部 1 941 个样本中, 583 个为检验样本, 利用本文提出的方法对设备全部 7 个失效模式进行智能分析诊断, 结果显示失效模式的首选正确识别率为 66%, 首选+次选正确识别率为 86%。依据影响因素分析方法, 对不同失效模式的关键特征参量及其比例、识别率进行了分析, 如表 6 所示。

表 6 不同失效模式的关键特征参量比例及识别率

特征参量	失效模式						
	1	2	3	4	5	6	7
关键特征参量 1	7	20	6	6	4	22	22
比例	1.00	0.96	0.99	1.00	1.00	0.97	0.44
关键特征参量 2	—	22	22	—	—	19	18
比例	—	0.04	0.01	—	—	0.03	0.11
关键特征参量 3	—	—	—	—	—	17	16
比例	—	—	—	—	—	0.01	0.10
关键特征参量 4	—	—	—	—	—	—	13
比例	—	—	—	—	—	—	0.08
其他关键特征参量	—	—	—	—	—	—	10,11,14,15,17,19,20,21
比例	—	—	—	—	—	—	0.28
样本总数	50	54	117	19	14	116	213
首选识别率	0.80	0.83	0.85	0.95	0.71	0.84	0.34
次选识别率	0.12	0.11	0.09	0	0.14	0.09	0.40
未识别率	0.08	0.06	0.06	0.05	0.14	0.07	0.26

从表 6 可以看出, 对于失效模式 1, 在全部 50 个样本中, 影响失效的关键特征参量全部指向降维后第 7 个参量, 显示第 7 个参量对该类失效的表征作用最为显著, 该失效的首选和次选识别率之和为 92%。而对于失效模式 2, 关键特征参量为第 20 个参量。从表中也可以发现, 对于没有明确失效特征的失效模式 7(其他缺陷), 由于不存在明确的缺陷特征, 因而对失效模式起决定性作用的关键特征参量分布极为分散, 如关键特征参量 22 的比例为 44%, 关键特征参量 18 的比例为 11%, 还有众多的其他关键特征参量也在其中占较大的比重, 达到 28%。由于缺少明确的判定该缺陷所属分类的关键特征参

量,因而这类缺陷的识别率较低,未识别率高达 26%,这是影响监测结果最重要的因素。

4 结 论

本文利用多维高斯贝叶斯分类模型对机械设备的故障智能诊断进行了研究,提出了基于“马氏距离”估算的影响因素分析方法。文中使用两个数据库对分析方法进行了验证,结果显示采用多维高斯贝叶斯分类方法进行机械设备的故障智能监测是可行的。利用两个数据库中的数据进行训练后,对检验样本中的设备失效/故障模式的首选+次选的正确识别率分别达到 96%和 86%。利用提出的距离估算方法对影响设备失效或故障模式的关键特征参量进行分析表明,对于具有明确缺陷特征的样本,如机械设备的轴承失效或基础变形,或钢板的划伤、锈渍等,多维高斯贝叶斯分类器具有很高的准确性。而对于特征不明确的失效或故障模式,由于无法很好地通过关键特征参量进行设备失效或故障状态的分析,会显著降低诊断的效果。

参 考 文 献

- [1] 陈学东, 范志超, 陈永东, 等. 我国压力容器设计制造与维护的绿色化与智能化[J]. 压力容器, 2017, 34(11): 12-27.
CHEN Xuedong, FAN Zhichao, CHEN Yongdong, et al. Green and intelligent design, manufacturing and maintenance of pressure vessels in China[J]. Pressure Vessel Technology, 2017, 34(11): 12-17.
- [2] DAVID J Smith. Reliability maintainability and risk-practical methods for engineers including reliability centered maintenance and safety related systems[M]. Burlington: Elsevier Butterworth-Heinemann, 2005.
- [3] 朱建新, 陈学东, 艾志斌, 等. 我国石化装置风险管理体系建设设想[J]. 石油化工设备, 2009, 38(6): 45-49.
ZHU Jianxin, CHEN Xuedong, AI Zhibin, et al. Consideration on the construction of risk management system for Chinese petrochemical plants[J]. Petrochemical Equipment, 2009, 38(6): 45-49.
- [4] 陈学东, 范志超, 江慧丰, 等. 复杂加载条件下压力容器典型用钢疲劳蠕变寿命预测方法[J]. 机械工程学报, 2009, 45(2): 81-87.
CHEN Xuedong, FAN Zhichao, JIANG Huifeng, et al. Creep-fatigue life prediction methods of pressure vessel typical steels under complicated loading conditions[J]. Journal of Mechanical Engineering, 2009, 45(2): 81-87.
- [5] 刘浩然, 丁攀, 郭长江, 等. 基于贝叶斯算法的中文垃圾邮件过滤系统研究[J]. 通信学报, 2018, 39(12): 151-159.
LIU Haoran, DING Pan, GUO Changjiang, et al. Study on Chinese spam filtering system based on Bayes algorithm[J]. Journal on Communications, 2018, 39(12): 151-159.
- [6] 赵申坤, 姜潮, 龙湘云. 一种基于数据驱动和贝叶斯理论的机械系统剩余寿命预测方法[J]. 机械工程学报, 2018, 54(12): 115-124.
ZHAO Shenkun, JIANG Chao, LONG Xiangyun. Remaining useful life estimation of mechanical systems based on the data-driven method and Bayesian theory[J]. Journal of Mechanical Engineering, 2018, 54(12): 115-124.
- [7] 刘东, 王昕, 黄建炎, 等. 基于贝叶斯网络的水电机组振动故障诊断研究[J]. 水力发电学报, 2019, 38(2): 112-120.
LIU Dong, WANG Xin, HUANG Jianying, et al. Vibration fault diagnosis for hydro-power units based on Bayesian network[J]. Journal of Hydroelectric Engineering, 2019, 38(2): 112-120.
- [8] 徐宾刚, 屈梁生, 陶肖明. 转子故障贝叶斯诊断网络的研究[J]. 机械工程学报, 2004, 40(1): 66-72.
XU Bingang, QU Liangsheng, TAO Xiaoming. Bayesian network framework for rotor fault diagnosis[J]. Journal of Mechanical Engineering, 2004, 40(1): 66-72.
- [9] MD T, FAISAL K, SYED I. Fault detection and pathway analysis using a dynamic Bayesian network[J]. Chemical Engineering Science, 2019, 195: 777-790.
- [10] 徐小力, 刘秀丽, 蒋章雷, 等. 基于主观贝叶斯推理的多传感器分布式故障检测融合方法[J]. 机械工程学报, 2015, 51(7): 91-98.
XU Xiaoli, LIU Xiuli, JIANG Zhanglei, et al. Multi-sensor distributed fault detection method based on subjective Bayesian reasoning[J]. Journal of Mechanical Engineering, 2015, 51(7): 91-98.
- [11] PAULA R, HOSSAM A, PETRÔNIO V, et al. A new methodology for multiple incipient fault diagnosis in transmission lines using QTA and Naïve Bayes classifier[J]. International Journal of Electrical Power & Energy Systems, 2018, 103: 326-346.
- [12] MURALIDHARAN V, UGUMARAN V. A comparative study of naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis[J]. Appl. Soft Computing, 2012, 12(8): 2023-2029.
- [13] ABDOLRAHMAN P, STEPHEN J, THAHIRAH J, et al. Evolutionary multi-objective fault diagnosis of power transformers[J]. Swarm and Evolutionary Computation,

- 2017, 36: 62-75.
- [14] 陈学东, 范志超, 崔军, 等. 极端条件重大承压设备的设计、制造与维护[J]. 机械工程学报, 2013, 49(22): 66-75.
- CHEN Xuedong, FAN Zhichao, CUI Jun, et al. Risk based design, manufacture and maintenance of extreme pressure equipment[J]. Journal of Mechanical Engineering, 2013, 49(22): 66-75.
- [15] CHANG C, LEE H, LIU H. A review of artificial intelligence algorithms used for smart machine tools[J]. Inventions, 2018, 3(3): 1-28.
- [16] 高金吉. 机器故障诊断与自愈化[M]. 北京: 高等教育出版社, 2012.
- GAO Jinji. Machine failure diagnosis & self-healing[M]. Beijing: High Education Press, 2012.
- [17] DUA D, GRAFF C. Mechanical Analysis Data Set[EB/OL]. Irvine : UCI Machine Learning Repository, 1990[2019-03-11]. <https://archive.ics.uci.edu/ml/datasets/Mechanical+Analysis>.
- [18] Research Center of Sciences of Communication. Steel plates faults data set[EB/OL]. Irvine : UCI Machine Learning Repository , 2010[2019-03-11]. [https : //archive.ics.uci.edu/ml /datasets/Steel+Plates+Faults](https://archive.ics.uci.edu/ml/datasets/Steel+Plates+Faults).
- [19] 盛骤, 谢式千, 潘承毅. 概率论与数理统计[M]. 北京: 高等教育出版社, 1995
- SHENG Zhou, XIE Shiqian, PAN Chengyi. Probability and mathematical statistics[M]. Beijing: High Education Press, 1995
- [20] CHRISTOPHER M. Pattern recognition and machine learning[M]. Singapore: Springer, 2006.
-
- 作者简介: 朱建新, 男, 1976 年出生, 博士, 研究员。主要研究方向为设备可靠性, 工程风险分析及设备智能运维。
E-mail: zhujianxin@hgmri.com