

# 基于运行大数据学习的复杂装备故障诊断技术及其典型应用

## Fault Diagnosis Technology for Complex Equipment Based on the Learning of Big Operation Data and Its Typical Application

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 04-0056-004

**摘要:** 认为通过从复杂装备运行特征大数据中挖掘出故障信息, 实现运行故障的智能诊断, 对保证复杂装备的安全和稳定运行具有重要意义。结合复杂装备运行大数据的特点与机器学习理论, 提出了基于运行大数据学习的复杂装备故障预测诊断方法, 实现了复杂装备运行特征大数据与运行故障的分层关联, 基于大数据分析的复杂装备运行故障特征提取以及基于模糊反向传播(BP)神经网络的复杂装备运行故障诊断。此外, 还将此技术应用到高速电梯的运行监测中, 开发了高速电梯急停故障大数据分析诊断系统, 很好地验证了该方法的有效性。

**关键词:** 大数据; 机器学习; 故障诊断; 复杂装备

**Abstract:** In this paper, we consider that it is significant to dig out the fault information from the big data of the complex equipment operation feature, and diagnose the operation fault intelligently for ensuring the safe and stable operations of complex equipment. By combining the characteristics of complex equipment's big operation data and the machine learning theory, a fault prediction and diagnosis method for complex equipment based on the learning of big operation data is proposed. In this way, some items are realized, including 1) the hierarchical correlation between the complex equipment's big operation data and the operation trouble; 2) the operation trouble feature extraction based on big data analysis; 3) the operation fault diagnosis based on fuzzy back propagation (BP) neural network for complex equipment. Moreover, by applying the proposed method to the high speed elevator's operation monitoring, the big data analysis and diagnosis system for emergency stop fault of high speed elevator is developed, and the effectiveness of the method is well verified.

**Key words:** big data; machine learning; fault diagnosis; complex equipment

刘达新/ LIU Daxin  
裘乐森/ QIU Lemiao  
王志平/ WANG Zhiping

(浙江大学, 浙江 杭州 310027)  
(Zhejiang University, Hangzhou 310027, China)

键智能基础共性技术之一。

随着装备复杂程度的增加, 对于装备运行的监测往往存在装备监测点多、监测点的采样频率高、数据收集时间长等特点, 使得复杂装备故障诊断系统需要处理的运行数据量呈现爆炸性增长, 数百太比特级甚至拍比特级规模的大数据已经屡见不鲜。海量运行数据的产生, 意味着复杂装备故障预测诊断技术迎来了它的大数据时代<sup>[1]</sup>, 也对故障诊断技术的发展提出了新的挑战。由于大数据往往蕴含着很多在小数据量时不具备的深度知识和价值, 只有通过大数据的智能化分析和挖掘, 才能将其价值显露出来, 因此将大数据分析机器学习技术应用于装备运行过程的故障预测诊断, 通过从复杂装备运行特征大数据中挖掘出故障信息, 实现运行故障的快速诊断, 是近年大数据在装备领域的重要应用之一。

国际上在利用大数据技术进行产品性能监测、故障诊断和提升产品性能方面已经有了较为成功的应用

**故**障诊断技术是保证复杂装备安全、稳定运行的重要技术之一。故障诊断技术通过对装备运行状态的监测及其相应数据的分析处

理, 实现对装备运行故障的预测和诊断, 判断装备的状态是否处于异常状态, 或故障状态、劣化状态发生的部位或零部件, 预测状态劣化的发展趋势等, 已广泛应用于大型空分装备、汽轮机组、航空发动机、高速电梯等复杂装备的运行监测控制, 被列为中国智能装备产业重点发展的九大关

收稿时间: 2016-07-22

网络出版时间: 2017-07-05

基金项目: 国家自然科学基金资助项目(51475417, L1524002)

案例。例如:美国GE公司应用2万台飞机引擎上的各种传感器收集飞行数据,精确检测飞机的运行状况,甚至预测故障,并及时进行预防性维保,其对于某些型号的引擎能够提前1个月预测维护需求,预测准确率达到70%;英国Rolls-Royce公司通过对服役于新加坡航空的137台飞机发动机使用大数据分析技术,成功降低了燃油消耗;美国特斯拉汽车公司利用大数据技术,采集和分析用户在驾驶时产生的加速度、刹车、转弯、电池充电和位置信息,以及用户驾驶习惯等大量数据,结合市场数据和企业生产、管理数据,进行产品及生产协同创新。

## 1 大数据与机器学习

狭义的大数据是指不能装载进计算机内存的数据。因此,对于每台电脑来说,当数据量增长到不能装载进内存的时候,此时形成的数据集便成了大数据。学术界对于大数据进行了更为科学的定义<sup>[1]</sup>:大数据是指无法在可接受的时间内用传统软硬件技术和工具对其进行感知、捕捉、管理、处理和服务的数据集合。针对大数据的特点,专家学者们相继提出了3V、4V甚至5V模型来进行描述,包括:Volume(体积大)、Velocity(速度高)、Variety(多样性)、Value(价值)、Veracity(真实性)等。

大数据分析主要分为简单分析和智能化复杂分析两大类<sup>[2]</sup>。简单分析主要采用类似于传统数据库联机分析处理(OLAP)的技术和方法,用结构化查询语言(SQL)完成各种常规的查询统计分析;而大数据的深度知识和价值仅通过简单分析是难以发现的,通常需要使用基于机器学习和数据挖掘的智能化复杂分析技术才能实现。机器学习作为人工智能的核心,是获取大数据所隐含深度知识的关键技术,而且处理的数据规模越大,机器学习模型的效果一般来说会越好<sup>[3]</sup>。

但是,由于大数据的海量、多维、多样、变化快等特性,使得小数据环境下基于内存处理的传统机器学习算法基本已不再适用于大数据下的应用问题。如何适应大数据处理的需求,这对机器学习的研究提出了新的挑战。大数据环境下的机器学习算法研究主要是从以下6个方面开展的<sup>[4]</sup>:

(1)大数据分治策略与抽样。大数据分治与并行处理策略是大数据处理的基本策略,如何学习大数据的分布知识,将其用于优化负载均衡是一个亟待解决的问题。同时,需要依据一定的性能标准对样本空间进行筛选,剔除冗余和噪音数据,在不降低甚至提高某方面性能的基础上,最大限度地降低计算时间和空间的消耗。

(2)大数据特征选择。大数据包含的属性数和记录数巨大,导致处理算法的执行效率低下。如何采用降维和特征选择技术以降低大数据处理难度,是大数据特征选择技术迫切需要解决的问题。

(3)大数据分类。针对不同分类算法研究并行或改进策略成为大数据环境下分类学习算法研究的主要方向,诸如支持向量机分类、决策树分类、神经网络分类等方法。近些年,在人工神经网络基础上发展起来的深度学习技术<sup>[5]</sup>,逐渐展现出其在处理大数据方面的独特优势。由于增加了隐层单元,多层神经网络比感知机具有更灵活且更丰富的表达力,可以用于建立更复杂的数学模型,能够深刻揭示海量数据里所隐含的复杂而丰富的信息,从而做出更为精准的预测。

(4)大数据聚类。聚类学习是最早被用于模式识别及数据挖掘任务的方法之一,并被用来研究各种应用中的大数据库。但是经典聚类算法在大数据环境下面临着数据量大、数据体积大、数据维度高等诸多挑战,改进现有聚类算法,提出新的聚类算法是大数据聚类研究急需解决的关

键问题。

(5)大数据关联分析。大数据的关联分析主要有并行和增量两种途径。并行关联分析算法通过将产生候选项集的过程并行化来提高运行效率,具有良好的加速比和伸缩性;增量方面则主要体现在序列模式挖掘上。

(6)大数据并行算法。将传统机器学习算法运用到大数据环境中的一个典型策略,便是对现有的学习算法并行化。目前大数据并行算法的研究已在一定范围内取得了一些进展,能实现对一定量级大数据的分析处理,如何建立更为高效的并行策略,实现大数据的高效处理仍将是当今的研究热点。

大数据环境下的机器学习,既不是单纯的机器学习,也不是单纯的大数据处理技术所能解决的问题,而是一个同时涉及机器学习和大数据处理的交叉性研究课题。

## 2 基于大数据学习的故障诊断方法

复杂装备运行过程中产生的海量特征数据蕴含了大量的故障信息,如何从复杂装备运行特征大数据中挖掘出故障信息,实现运行故障的快速诊断,对提高复杂装备的安全性,实现稳定运行具有重要意义。基于大数据分析的故障诊断可以在收集到复杂装备运行特征数据的基础上,应用聚类、决策树等机器学习算法对大数据进行知识挖掘,获得与故障有关的诊断规则,从而实现对复杂装备的故障预测和诊断。

目前,复杂装备故障诊断方法主要分为3类:基于数学模型的故障诊断、基于数字信号处理的故障诊断、基于知识的故障诊断<sup>[6]</sup>。基于数学模型的故障诊断方法简单直观且易于理解,但需要深入分析复杂装备的结构、运行原理,对内部结构和运行原理过于复杂的装备难以建立数学模型;基于信号处理的诊断方法不需要

构建数学模型,容易实现,但只在复杂装备有明显的外部特征时才有效,不适用于那些没有明显外部特征的故障;基于知识的故障诊断具有良好的诊断效果和广泛的适用性,对复杂装备的智能化要求较高,是目前实现复杂装备智能故障诊断与健康维护的主要研究方向。

结合复杂装备运行大数据的特点与机器学习理论,我们提出了基于运行大数据学习的复杂装备故障预测诊断方法,包括复杂装备运行特征大数据与运行故障的分层关联,基于大数据分析的复杂装备运行故障特征提取以及基于模糊反向传播(BP)神经网络的复杂装备运行故障诊断3个方面。

(1)复杂装备运行特征大数据与运行故障的分层关联。

通过详细分析复杂装备运行故障出现的原因,建立运行故障原因的层次结构模型,分析和采集复杂装备运行特征大数据,建立复杂装备运行特征参数与运行故障原因的关联映射关系。

(2)基于大数据分析的复杂装备运行故障特征提取。

针对复杂装备运行特征大数据有较多噪声的问题,提出基于小波模

极大值的复杂装备信号降噪方法,对特征大数据进行处理,提高复杂装备运行数据的信噪比。针对复杂装备运行特征大数据维数多、规模大等问题,提出基于粗糙集属性约简的运行故障特征提取方法,筛选运行特征大数据中的冗余属性,获取与复杂装备运行故障相关的特征参数。

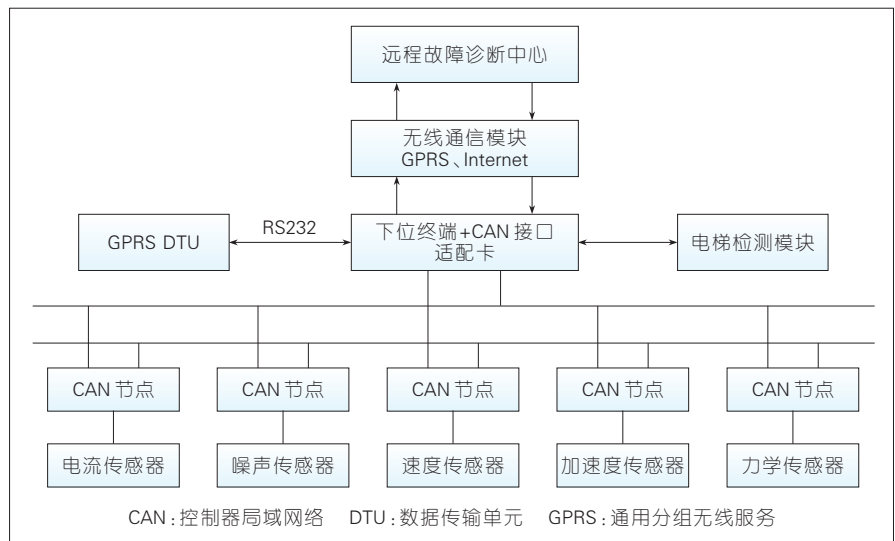
(3)基于模糊BP神经网络的复杂装备运行故障诊断。

基于复杂装备运行故障诊断的模糊BP神经网络构建,基于神经网络的复杂装备运行故障诊断方法和

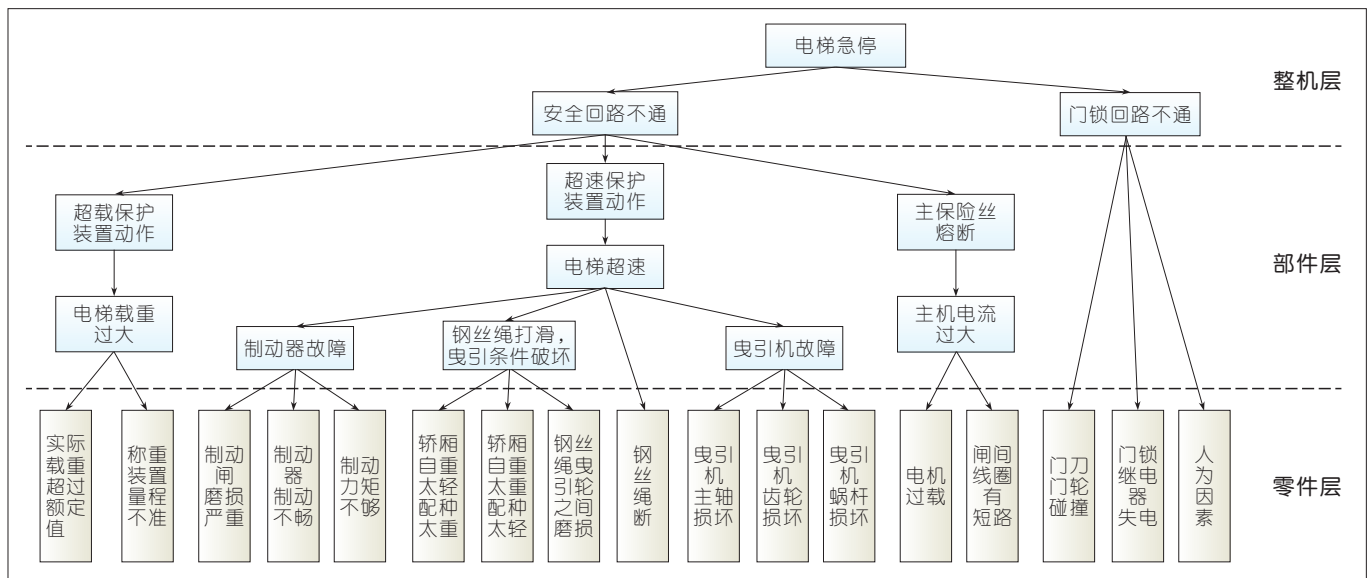
诊断过程的研究,建立复杂装备运行故障诊断模型,实现复杂装备运行特征大数据分析与基于知识的运行故障智能诊断。

### 3 典型应用案例

以高速电梯为研究对象,通过将研究成果应用到高速电梯急停故障的数据分析和智能诊断,结合高速电梯急停故障大数据分析与诊断系统的开发,验证所提方法的有效性。如图1,诊断系统通过收集高速电梯运行特征大数据,建立如图2所示的运



▲图1 高速电梯运行数据实时采集模块



▲图2 高速电梯急停故障的层次性

行特征参数与急停原因的分层关联映射。采用基于小波模极大值的高速电梯运行信号降噪算法去除运行特征大数据中的冗余信息,采用基于属性约简的高速电梯故障特征提取方法对运行大数据进行特征提取,获取高速电梯急停故障的特征参数,基于模糊BP神经网络构建了高速电梯急停故障诊断模型,如图3。

如图4所示,系统在江苏康力电梯企业额定速度7 m/s的KWG型高速电梯急停故障诊断中得到了应用验证,可在系统中查看KWG型高速电梯所有与急停故障相关的信息,有助于维修人员根据诊断结果对故障进行快速修复。

#### 4 结束语

目前,中国在将大数据、人工智能等技术应用于复杂装备故障诊断

方面还处于理论研究阶段,在实际中主要以基于数学模型和基于信号处理的故障诊断方法为主。由于复杂装备的结构复杂,运行工况恶劣多变,难以建立精确的解析模型,运行信号也不能直观地反映装备运行故障的情况,因此传统的基于模型和信号处理的故障诊断方法已经不能适用于复杂装备的故障诊断。而随着物联网技术的发展,越来越多的装备制造商会通过传感器收集装备的运行数据,迅猛增长的装备运行特征数据对于分析其故障具有重要价值。因此,我们结合复杂装备运行大数据的特点,提出了基于运行大数据学习的复杂装备故障预测诊断方法,实现了复杂装备运行特征大数据与运行故障的分层关联,基于大数据分析的复杂装备运行故障特征提取以及基于模糊BP神经网络的复杂装备

运行故障诊断,应用于高速电梯故障预测与诊断,有效提高了复杂装备故障诊断的效率和准确性,取得了很好的应用效果。

#### 参考文献

- [1] GRAHAM R D, GOLDSTON D, DOCTOROW C, et al. Big Data: Science in the Petabyte Era [J]. Nature, 2008, 455(7209):8-9. DOI: 10.1038/455001a
- [2] 李国杰, 程学旗. 大数据研究: 未来科技及经济社会发展的重大战略领域——大数据的研究现状与科学思考[J]. 中国科学院院刊, 2012, 27(6):647-657
- [3] 黄宜华. 大数据机器学习系统研究进展[J]. 大数据, 2015, 1(1): 28-47
- [4] WANG Y, ZHAO X, SUN Z, et al. Peacock: Learning Long-Tail Topic Features for Industrial Applications[J]. Acm Transactions on Intelligent Systems & Technology, 2015, 6(4):1-23. DOI:10.1145/2700497
- [5] 何清, 李宁, 罗文娟, 等. 大数据下的机器学习算法综述[J]. 模式识别与人工智能, 2014, 27(4): 327-336. DOI:10.3969/j.issn.1003-6059.2014.04.007
- [6] HINTON G E, SALAKHUTDINOV R R. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(5786):504-507. DOI:10.1126/science.1127647
- [7] 王宏力, 侯青剑. 故障诊断方法现状与展望[J]. 传感器与微系统, 2008, 27(5):1-4. DOI: 10.3969/j.issn.1000-9787.2008.05.001

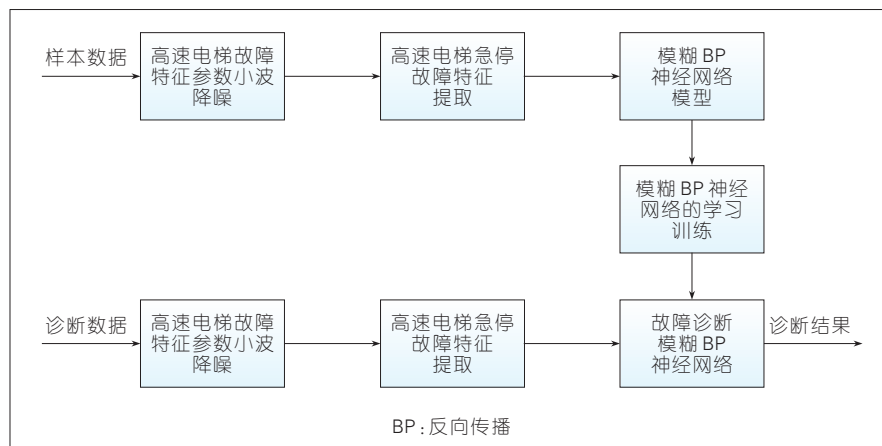


图3 基于模糊BP神经网络的高速电梯急停故障诊断

设备编号	项目名称	客户	地区	是否发生急停	故障原因	安装日期	质保期	维修时间
1 SA00141704	平湖市民文化中心人防工程	杭州康力电梯有限公司	杭州	是	称量装置故障	2014-01-18	2017-01-18	2015-02-05 11:2...
2 SA00141743	上城二期	成都金源电梯有限公司	成都	否		2014-01-16	2017-01-16	2015-02-05 11:2...
3 SA00141056	乌海万达广场	乌海万达广场投资有限公司	乌海	是	制动闸线严重	2014-01-08	2017-01-08	2015-02-05 11:2...
4 SA00141312	杭州绿城大和康力电梯	康力电梯有限公司	康力	是	制动器制动力不足	2014-01-03	2017-01-03	2015-02-05 11:2...
5 SA00141429	康力电梯二期	康力电梯有限公司	康力	否		2014-01-28	2017-01-28	2015-02-05 11:2...
6 SA00141058	湖州康力电梯有限公司	湖州康力电梯有限公司	湖州	是	钢丝绳断裂, 安全	2014-01-31	2017-01-31	2015-02-05 11:2...
7 SA00141089	湖州康力电梯有限公司	湖州康力电梯有限公司	湖州	否		2014-01-31	2017-01-31	2015-02-05 11:2...
8 SA00141835	杭州康力电梯有限公司	杭州康力电梯有限公司	杭州	否		2014-01-19	2017-01-19	2015-02-05 11:2...
9 SA00141090	湖州康力电梯有限公司	湖州康力电梯有限公司	湖州	是	门刀门架碰撞	2014-01-31	2017-01-31	2015-02-05 11:2...
10 SA00141091	湖州康力电梯有限公司	湖州康力电梯有限公司	湖州	否		2014-01-31	2017-01-31	2015-02-05 11:2...
11 SA00141841	甘肃康力电梯有限公司	甘肃康力电梯有限公司	兰州	否		2014-01-19	2017-01-19	2015-02-05 11:2...
12 SA00141853	福州康力电梯有限公司	福州康力电梯有限公司	福州	否		2014-01-13	2017-01-13	2015-02-05 11:2...
13 SA00141854	福州康力电梯有限公司	福州康力电梯有限公司	福州	是	门架碰撞	2014-01-13	2017-01-13	2015-02-05 11:2...
14 SA00141076	康力电梯二期	康力电梯有限公司	康力	否		2014-01-12	2017-01-12	2015-02-05 11:2...
15 SA00141029	康力电梯二期	康力电梯有限公司	康力	否		2014-01-30	2017-01-30	2015-02-05 11:2...
16 SA00141723	康力电梯二期	康力电梯有限公司	康力	是	电机过载	2014-01-10	2017-01-10	2015-02-05 11:2...
17 SA00141787	康力电梯二期	康力电梯有限公司	康力	否		2014-01-19	2017-01-19	2015-02-05 11:2...
18 SA00141782	康力电梯二期	康力电梯有限公司	康力	否		2014-01-05	2017-01-05	2015-02-05 11:2...

图4 高速电梯急停故障诊断结果

#### 作者简介



刘达新, 浙江大学助理研究员; 主要研究方向为复杂产品数字样机与仿真、产品数字化设计与制造、计算机辅助检测等; 曾主持和参加国家自然科学基金项目3项; 已在重要学术期刊发表论文10余篇, 其中SCI/EI检索8篇。



裘乐森, 浙江大学副教授; 主要研究方向为数字化设计制造、制造业信息化、数字化工厂与智能制造等; 曾主持国家自然科学基金2项、浙江省自然科学基金1项, 参与国家重大科技专项、“863”计划项目等10余项, 获国家科技进步二等奖1项、浙江省科技进步奖2项; 已发表SCI/EI论文30余篇。



王志平, 浙江大学硕士研究生; 主要研究方向为产品数字化设计制造、制造业信息化; 曾参与国家自然科学基金等项目2项。