

故障诊断 中的 数据挖掘

胡小平 韩泉东 李京浩/著



GUZHANG ZHENDUAN ZHONGDE
SHUJU WAJUE

国防科技大学出版社

国防科技大学学术
著作专项经费资助出版

故障诊断中的数据挖掘

胡小平 韩泉东 李京浩 著

国防科技大学出版社
湖南·长沙

长沙:国

图书在版编目(CIP)数据

故障诊断中的数据挖掘 胡小平, 韩泉东, 李京浩 著. —长沙:国防科技大学出版社, 2009.11.10

ISBN 978-7-81099-695-2

Ⅰ. 胡… Ⅱ. 数据采集—计算机应用—液体推进

剂火箭发动机—故障诊断—研究 Ⅲ. V434-39

中国版本图书馆 CIP 数据核字(2009)第 133165 号

国防科技大学图书馆 电话:(0731)84572640 网址:www.gfkdbs.com

责任编辑:王 嘉 责任校对:王 嘉

新华书店总店北京发行所经销 印数:1-500 册

开本:850×1168 1/32 印数:8.5 字数:220 千

2009 年 10 月第 1 版第 1 次印刷 定价:26.00 元

ISBN 978-7-81099-695-2

内容简介

数据挖掘技术是新兴的数据分析工具,它是从大量的、不完全的、有噪声的、模糊的、随机的实际应用数据中,提取隐含在其中的、人们事先未知的但又是潜在有用的信息和知识的过程。本书在深入分析和研究故障诊断和数据挖掘技术特点的基础上,以液体火箭发动机的故障检测和诊断为对象,系统地研究了聚类、分类、关联、时间序列分析等数据挖掘技术在故障诊断中的应用,从液体火箭发动机的大量热试车数据和故障仿真数据中提取故障诊断所需的知识,实现了故障诊断中的数据挖掘。

前 言

航天技术作为人类探索宇宙空间的主要手段,在军事、政治、经济和科技等许多领域中正发挥着越来越重要的作用。同时,随着人们对航天器、运载器及其推进系统要求的提高,整个系统也变得更加复杂。“大力神”2 运载火箭有大约 50 万个零部件,“土星”火箭约有 100 万个零部件,“阿波罗”飞船的零部件数目则高达 720 万个。在如此庞大复杂和投资巨大的系统中,人们迫切需要通过提高航天器、运载器及其推进系统的可靠性和安全性;否则,系统的一次故障就可能导致灾难性的后果和巨大的经济损失。

液体火箭发动机系统是运载器及航天器中应用最广泛的动力系统,恶劣的工作环境使其成为航天器系统故障的敏感多发部位。据统计,在国外通信卫星发射失败的事例中,由发动机故障引起的,达 54%。在 1990—1994 年间世界 15 次重大运载火箭发射失败中,由推进系统故障导致的失败次数有 14 次。因此,提高推进系统的运行可靠性是航天器系统可靠性工程中的关键环节。

开展液体火箭发动机故障检测和诊断技术的研究,除了保障航天任务的圆满完成以外,对于减少地面工作人员的工作量、航天员的培训时间以及降低发射与运行成本都具有重要意义。近几十年来,随着计算机、自动控制、人工智能、传感器等技术的飞速发展和广泛应用,液体火箭发动机故障检测和诊断技术的研究已取得了重要进展,主要集中在以下几方面:基于信号分析的故障检测与诊断;基于模型的故障检测与诊断;基于人工智能的故障检测与诊断。

尽管液体火箭发动机的故障检测与诊断技术的研究已经取得了很大的进展,但是面对大量数据,知识的自动获取仍是制约其发展的“瓶颈”。要想构造真正实用、可靠、高效、智能化、集成化的适用多种需求的火箭发动机健康监控产品,仍然需要充分利用现代传感器技术、计算技术、网络通信技术、人工智能技术、信号分析处理技术,并结合液体火箭发动机综合动力学模型的建立,开展更深入系统的研究和开发工作。

数据挖掘技术的出现,为解决这一“瓶颈”问题提供了一种有效的途径。

数据挖掘是对统计学、数据库技术、模式识别、人工智能、可视化技术、数据分析等学科知识的综合运用。它是从大量的、不完全的、有噪声的、模糊的、随机的实际应用数据中,提取隐含在其中的、人们事先不知道的但又是

潜在有用的信息和知识的过程。它把人们对数据的应用从低层次的简单查询和分析处理,提升到从数据中挖掘知识、提供决策支持的高度。一般而言,数据挖掘技术更着眼于设计高效的算法以达到从巨量的数据中发现规则和知识的目的。数据挖掘在没有明确假设的前提下去挖掘信息、发现知识,解决“数据爆炸但知识贫乏”的现象。

在国家自然科学基金项目“数据挖掘和知识发现在火箭发动机故障诊断中的应用研究”(项目编号:50376073)的支持下,从2003年开始,本书作者们开展了数据挖掘在液体火箭发动机故障检测和诊断中应用的有关研究,尝试把数据挖掘技术应用到液体火箭发动机的故障诊断中,以便有效地分析和定位液体火箭发动机发生的故障,及早查出故障发生的原因,从而达到减少试车次数、缩短研制周期、降低研制费用、将由故障引起的损失降低到最低程度的目的。

本书是该项目相关研究成果的一个总结。本书分为12章,第1章论述了液体火箭发动机故障检测和诊断的研究进展;第2章阐述了在故障诊断领域中数据挖掘技术的研究进展;第3章论述了故障检测和诊断中数据挖掘的策略和数据的预处理方法;第4章到第11章依次是基于粗糙集理论的故障检测和诊断、基于关联规则的故障检测和诊断、基于关联规则分类技术的故障检测和诊断、基于数值型关联规则的故障诊断、基于决策树方法的

故障检测和诊断、基于支持向量机的故障检测和诊断、结合动态时间弯曲与决策树方法的故障检测与诊断,以及基于贝叶斯方法的故障检测和诊断。最后,介绍了故障诊断的多种数据挖掘方法的表决系统和软件集成。

感谢国家自然科学基金委员会的项目支持,感谢国防科技大学学术专著出版基金的支持。感谢黄敏超博士、刘洪刚博士提供的部分试车数据,感谢刘冰博士提供的发动机故障仿真程序。感谢参与项目研究的李舟军教授、张丽娟博士、王艳梅、韩泉东、李京浩、何浩、刘枫、刘泽军等研究生的辛勤工作。

由于研究对象的复杂性,由于所采用的数据挖掘方法——相对于航天技术领域而言属于跨学科的计算机技术的范畴,也由于作者的水平有限,书中难免存在一些错误,敬请读者批评指正。

胡小平,韩泉东,李京浩

2009年7月

目 录

第 1 章 液体火箭发动机故障检测和诊断的进展

- 1.1 基于信号分析的故障检测与诊断 (4)
 - 1.1.1 早期的故障检测算法 (4)
 - 1.1.2 振动信号的监测 (5)
 - 1.1.3 羽流光谱技术 (6)
- 1.2 基于模型的故障检测与诊断 (6)
 - 1.2.1 基于信号本身结构模型的故障检测与诊断
..... (7)
 - 1.2.2 基于灰色理论的故障检测与诊断 (7)
 - 1.2.3 基于数学模型的故障检测与诊断 (7)
- 1.3 基于人工智能的故障检测与诊断 (9)
 - 1.3.1 模式识别方法 (9)
 - 1.3.2 专家系统方法 (10)
 - 1.3.3 神经网络方法 (11)
 - 1.3.4 定性推理方法 (13)
 - 1.3.5 信息融合方法 (14)
 - 1.3.6 其他人工智能方法 (14)
- 1.4 小 结 (15)

参考文献.....	(17)
 第 2 章 故障诊断领域中数据挖掘技术的进展	
2.1 引 言	(19)
2.2 数据挖掘技术的应用实例	(22)
2.3 数据挖掘技术在故障诊断和工程数据方面的应用	(23)
2.4 小 结	(30)
参考文献.....	(30)
 第 3 章 故障检测和诊断中数据挖掘策略和数据预处理	
3.1 数据挖掘处理过程模型的选择	(37)
3.2 数据挖掘方法选择和数据分析	(39)
3.3 可用于故障检测和诊断的数据挖掘方法	(46)
3.3.1 分类方法	(47)
3.3.2 孤立点分析(Isolated point)	(50)
3.3.3 关联规则挖掘(Association rules)	(50)
3.3.4 时间序列分析(Time series analysis)	(51)
3.3.5 关于火箭发动机中的数据挖掘方法的一些讨论	(51)
3.4 数据预处理	(52)
3.4.1 数据预处理的基本环节	(53)
3.4.2 液体火箭发动机试车数据预处理实例	(56)

3.5 小 结 (59)

参考文献..... (59)

第 4 章 基于粗糙集理论的故障检测和诊断

4.1 引 言 (61)

4.2 粗糙集的理论基础 (62)

4.2.1 信息表知识表达系统 (63)

4.2.2 决策表和可辨识矩阵 (64)

4.2.3 决策规则的表示 (65)

4.3 决策表属性约简 (66)

4.3.1 数据预处理 (67)

4.3.2 决策表属性约简的必要性 (67)

4.3.3 决策表属性约简算法 (68)

4.4 基于粗糙集的液体火箭发动机故障特征参数选择
..... (69)

4.4.1 利用仿真数据检查属性约简选择故障特征参数的
适用性 (71)

4.4.2 利用属性约简选择发动机试车数据的故障特征
参数 (75)

4.5 决策表值约简 (78)

4.6 基于粗糙集的液体火箭发动机故障诊断实例 ... (81)

4.6.1 利用仿真数据检查值约简诊断故障的适用性
..... (81)

4.6.2 利用值约简诊断发动机试车数据故障
..... (85)

4.7 小 结	(88)
参考文献.....	(88)

第 5 章 基于关联规则的故障检测和诊断

5.1 引 言	(90)
5.2 关联规则的理论基础	(91)
5.2.1 关联规则的基本概念	(92)
5.2.2 关联规则的分类	(93)
5.2.3 关联规则的挖掘过程	(94)
5.3 Apriori 算法	(94)
5.4 利用关联规则检测液体火箭发动机起动/关机过程的故障	(96)
5.5 小 结	(101)
参考文献.....	(101)

第 6 章 基于关联规则分类技术的故障检测和诊断

6.1 引 言	(103)
6.2 基于关联规则的分类技术	(104)
6.2.1 基本概念	(104)
6.2.2 CBA-RG 算法.....	(106)
6.2.3 用于关联规则挖掘的约束	(107)
6.3 利用 CBA 技术检测液体火箭发动机的稳态故障	(108)
6.3.1 属性编码	(108)

6.3.2	数据预处理	(109)
6.3.3	重要的约束	(110)
6.3.4	基于关联规则分类技术的挖掘	(110)
6.3.5	故障检测	(111)
6.3.6	评述	(112)
6.4	小 结	(113)
	参考文献.....	(114)

第 7 章 基于数值型关联规则的故障诊断

7.1	引 言	(115)
7.2	数值型关联规则相关知识	(116)
7.2.1	关联规则介绍	(116)
7.2.2	不产生候选项集的关联规则算法 FP-Growth	(117)
7.2.3	数值型关联规则	(122)
7.3	应用实例	(126)
7.3.1	模糊处理	(126)
7.3.2	连续属性的离散化	(127)
7.3.3	利用 FP-Growth 算法挖掘频繁项集	(130)
7.4	小 结	(135)
	参考文献.....	(135)

第 8 章 基于决策树方法的故障检测和诊断

8.1	引 言	(138)
-----	-----------	-------

8.2	决策树分类算法	(139)
8.2.1	决策树基本算法	(139)
8.2.2	ID3 算法.....	(141)
8.2.3	C4.5 算法	(143)
8.2.4	CART 算法	(145)
8.2.5	SLIQ 算法	(147)
8.2.6	SPRINT 算法	(151)
8.2.7	PUBLIC 算法	(153)
8.2.8	Rain Forest 算法.....	(154)
8.2.9	ID5R 算法	(155)
8.3	故障检测和诊断策略	(156)
8.4	实例分析	(157)
8.4.1	数据预处理	(158)
8.4.2	发动机热试车数据挖掘	(158)
8.4.3	故障仿真数据的挖掘验证	(162)
8.4.4	评述	(164)
8.5	决策树方法的一些研究进展	(164)
8.5.1	结合其他算法提高决策树的精度	(165)
8.5.2	处理大型数据集的并行机制	(167)
8.5.3	多分类器组合决策树分类算法	(168)
8.5.4	形成新的决策树算法	(169)
8.6	决策树方法的研究展望	(172)
8.7	小 结	(173)
	参考文献.....	(174)

第 9 章 基于支持向量机的故障检测和诊断

- 9.1 引 言 (177)
- 9.2 支持向量机分类机理 (178)
 - 9.2.1 线性可分情形 (179)
 - 9.2.2 非线性可分情形 (181)
 - 9.2.3 多类支持向量机 (183)
- 9.3 故障检测和诊断策略 (188)
- 9.4 数据挖掘实例 (191)
 - 9.4.1 数据准备 (191)
 - 9.4.2 热试车数据挖掘与分析 (191)
 - 9.4.3 评述 (193)
 - 9.4.4 仿真数据的挖掘验证与分析 (194)
 - 9.4.5 支持向量机方法的进一步讨论 (195)
- 9.5 支持向量机的一些研究进展 (196)
 - 9.5.1 SVM 的研究进展 (196)
 - 9.5.2 SVM 的发展趋势 (201)
- 9.6 小 结 (203)
- 参考文献..... (203)

第 10 章 结合动态时间弯曲与决策树方法的故障检测
和诊断

- 10.1 引 言 (205)
- 10.2 动态时间弯曲的相关概念 (207)

10.2.1	动态时间弯曲方法的描述	(207)
10.2.2	弯曲路径的求解	(210)
10.2.3	DTW 算法伪代码	(211)
10.3	故障检测和诊断策略	(212)
10.4	故障检测和诊断实例	(213)
10.4.1	数据准备	(213)
10.4.2	基于热试车数据的故障检测和诊断	(214)
10.4.3	基于仿真数据的故障检测和诊断	(217)
10.4.4	评述	(217)
10.5	小 结	(218)
	参考文献	(219)

第 11 章 基于贝叶斯方法的故障检测和诊断

11.1	引 言	(220)
11.2	基本概念	(220)
11.2.1	贝叶斯定理	(221)
11.2.2	贝叶斯假设	(222)
11.3	贝叶斯分类器	(223)
11.3.1	朴素贝叶斯分类器	(223)
11.3.2	贝叶斯网络	(225)
11.3.3	TAN 分类器	(228)
11.4	应用实例	(230)
11.4.1	数据预处理	(230)
11.4.2	利用贝叶斯分类器进行故障分类	(231)
11.4.3	利用仿真数据进行挖掘验证	(236)
11.5	小 结	(238)

参考文献..... (238)

第 12 章 故障诊断表决模块与软件集成

12.1 数据挖掘软件的发展情况 (240)

12.2 系统的软件实现 (242)

12.3 液体火箭发动机故障检测和诊断数据挖掘系统介绍
..... (243)

 12.3.1 系统简介 (243)

 12.3.2 故障诊断表决模块 (249)

12.4 小 结 (251)

参考文献..... (251)