

数据挖掘技术在航空发动机PHM中的应用^{*}

张叔农, 康 锐

(北京航空航天大学工程系统工程系, 北京 100083)

摘 要:总结了数据挖掘工具的方法和模式以及数据挖掘的流程, 对其在航空发动机故障预测和健康管理 (PHM) 中的常用方法和模式, 如神经网络、时间序列分析、模糊逻辑、遗传算法、证据理论、支持向量机等及其应用案例进行了阐述。

关键词:数据挖掘; PHM; 故障预测; 健康管理; 航空发动机

中图分类号:V23 **文献标志码:**A

Application of Data Mining Technology to Aeroengine Prognostic and Health Management

ZHANG Shunong, KANG Rui

(Department of Project Systems Engineering, BUAA, Beijing 100083, China)

Abstract The methods, modes and processes of data mining technology were summarized. Several methods and modes commonly used on aero engine prognostic and health management (PHM) such as neural network, time series analysis, fuzzy logic, genetic algorithm, evidence theories and were introduced, practical examples were also given.

Keywords data mining; PHM; prognostic; health management; aeroengine

1 引言

随着 EHM(发动机健康监控和管理) 和 PHM 概念的产生, 航空发动机 PHM 技术又有了迅猛的发展, 主要体现在对发动机整机和部件的异常检测、故障诊断、故障预测软硬件能力的增强以及对发动机健康管理功能的完善上。在航空发动机 PHM 方法中, 数据挖掘技术便是重要的一部分。文中简介了数据挖掘技术的方法/模式和流程, 并对目前在航空发动机 PHM 中几种常用的方法/模式, 如神经网络(NN)、时间序列分析、模糊逻辑、遗传算法、证据理论、支持向量机(SVM) 等及其应用案例进行了阐述。

2 数据挖掘技术在航空发动机 PHM 中的应用

所谓数据挖掘, 就是从数据库中抽取隐含的、以前未知的、具有潜在应用价值的信息过程。数据挖掘流程一般包括三部分: 数据准备、数据建模与数据挖掘、结果解释与评价。数据准备主

要对各类数据(结构化、半结构化、异构型、非结构化数据) 进行预处理, 包括数据清理、数据集成、数据选择和数据转换。数据建模与数据挖掘是 KDD(知识发现) 过程的核心组成部分, 其工作是根据挖掘任务选择适当的数据挖掘任务表达与算法(主要包括模糊方法、粗糙集理论、云理论、证据理论、神经网络、遗传算法、归纳学习、支持向量机等), 并利用其来寻找有趣的模式(主要包括分类、聚类、关联、回归、时序、偏差等), 然后用某种方式来表示模式。结果解释与评价: 主要对挖掘出的模式进行评估, 识别表示知识的真正有趣的模式, 并利用可视化和知识表达技术, 向用户展示所挖掘出的相关知识。在实际应用中, 数据挖掘流程是一个反复循环的过程, 从数据预处理、数据挖掘、模型建立到结果评价, 各部分之间并没有明显的界限, 可以从任何一个步骤返回前面的环节^[1]。

用于航空发动机 PHM 数据挖掘的信息源主要包括气路传感器、滑油/燃油系统传感器、振

^{*} 收稿日期: 2006-12-04

作者简介: 张叔农(1968-), 女, 江西万载人, 博士研究生, 研究方向: 故障预测和健康管理, 失效物理, 可靠性系统工程等。

动传感器、结构评估传感器等采集的数据,以及 FADEC(全权限数字电子控制系统)代码信息、同一飞机的发动机数据、机载发动机模型、发动机历史飞行数据、物理状况、共模失效、发动机使用因素、发动机和机群的维修历史记录、推理信息^[3]等。在诸多报导中,神经网络、时间序列分析、模糊逻辑、遗传算法、证据理论、支持向量机在航空发动机PHM中应用较多。

2.1 神经网络

神经网络由多个神经元按照某种方式相互连接形成,依靠网络状态对外部输入信息的动态响应来处理信息。它常用于分类、聚类、预测和模式识别。神经网络的核心是结构和算法,它通过模仿生物神经网络对训练数据进行挖掘,逐步计算网络连接的权值,具有强大的非线性映射能力、并行处理能力和学习能力。

在诸多数据挖掘方法中,神经网络在发动机PHM中的应用探索是最为丰富的,主要体现在对发动机的起动、关机与改变工况等动态过程的状态监控;对发动机气路故障的监测与诊断;基于油液分析的发动机故障诊断和预测;液体火箭发动机泄漏故障的识别;发动机可靠性分析与故障预测等方面。发动机PHM中的神经网络模型主要有逆向传输(BP)网络、径向基函数(RBF)网络、概率神经网络(PNN)、自组织映照模型、模糊神经网络等^[3]。

具体将神经网络应用于发动机PHM中,需要重点考虑的问题有:1)对于复杂系统,可以考虑采用多重神经网络模型结构,以便减少分类中的模式匹配搜索量,有效地解决分类空间的组合爆炸问题;2)提高网络训练速度方面的研究,可以考虑从求解算法、简化网络结构的角度入手;3)训练样本的选取以及不同来源样本的综合运用,以及对模型的依赖及测量噪声引起的误诊和漏诊^[3]。

2.2 时间序列

序列模式是描述基于时间或其它序列的经常发生的规律或趋势,并对其建模。时间序列模式是其中的重要分支。它通常采用趋势分析技术(即发现时序数据的发展趋势)和相似搜索技术(查找与给定查询序列最接近的序列)对时序数据库进行挖掘。序列模式将关联模式和时间序列模式结合起来,重点考虑数据之间在时间维

上的关联性。时间序列分析已有不少用于发动机滑油监控、诊断和预测的案例。如文献[4]利用AR模型对某型航空发动机滑油监控系统中铁元素含量变化趋势进行了预测分析,并对模型的预测效果进行了检验和比较。也可对滑油监控系统中其他需重点监控的金属元素镁、铝、铜等分别建立AR模型用于预测。又如文献[5]利用波克斯-詹金斯方法对航空发动机滑油光谱的数据进行分析,对滑油铁的含量建模和预测等等。

2.3 模糊逻辑

模糊逻辑是经典集合理论的扩展,它以模糊集合和模糊推理为数学基础,以不确定性的事物作为研究对象。模糊集合的核心概念是隶属度函数。模糊推理就是由前提(A'),依规则(if A then B)推理,得到结论(B')。规则是一种蕴含关系。如何选择蕴含关系和合成算法是模糊推理的关键。模糊推理注重的是把握结论的趋势,即近似的而不是精确的结果。模糊逻辑系统已用于许多特别是基于规则的分类领域。在发动机PHM中,模糊逻辑常用于数据融合和决策支持。如表1显示了航空发动机部分监控参数与导出参数及可能说明的故障原因等的相关关系规则^[9]。这些规则构成一些状态,这些状态一般不呈现,但在某种故障情节存在时的假设下将/应该会意识到。

表 1 部分监控参数与导出参数及可能的故障原因

座舱指示				导出参数				可能的原因
高压转速	排气温度	发动机压比	燃油流量	压气机压比	压气机程升比	涡轮落压比	涡轮落温比	
+	+	+	+	+	+	+	+	低压转速指示不正常
+	+	+	+	+	+	o	o	低压涡轮性能下降
o	o	o	+	o	o	o	o	燃油流量指示不正常
+	o	o	o	o	o	o	o	高压转速指示不正常
+	+	-	+	-/o	-/o	+/o	+	放气量过大
-	+	-	+	-	+	-/o	+	压气机故障
+	+	-	-	+	+	+	+	静叶可调位置全开
+/o	-/o	-	-	-	-	-/o	-/o	风扇故障
-	+	+/o	-	-	-	o	+	高压涡轮不正常

注: + : 增加; - : 下降; o : 不变

2.4 证据理论

证据理论是经典概率论的扩充。该理论的重要贡献在于划清了不确定和未知的界限。在证据

理论中,一个样本空间成为一个识别框架,用 Ω 表示。 Ω 由一系列对象构成,对象之间两两互斥,且包含当前要识别的全体对象。 Ω 的所有子集的集合记为 $2^{\wedge} \Omega$,若 Ω 中有 n 个对象,则有 $2^{\wedge} n$ 个子集。每个子集对应一个命题(证据或结论),证据理论的基本问题是:已知识别框架 Ω ,判明 Ω 中一个先验的未定位对象属于 Ω 的某个子集 A 的程度。Anand 等于 1994 年提出了一个基于证据理论的通用数据挖掘框架 EDM。该框架的算法在处理并行、分布及异构数据库时有显著的优越性,因而基于证据理论的方法在数据挖掘中具有潜在的应用性。美国 Impact 公司对齿轮箱的故障预测采用了证据理论,该方法被应用于 JSF(联合攻击战斗机)PHM 等系统中,可以对由于低周疲劳断裂或齿裂引起的齿轮齿牙故障进行实时的预测,并在美国宾州 ARL 实验室对结果进行了验证。其过程是首先采集齿轮箱的速度、扭矩和加速度,利用物理模型实时计算齿轮齿的裂纹初始和增长,振动特征用于确定现场或试验时实际的故障概率,然后利用 D-S 证据理论将基于振动信号的结果和基于物理模型的结果进行融合,得到故障的联合概率^[7-9]。

2.5 遗传算法

遗传算法是模拟生物的进化和遗传,借助 3 个过程——选择、交叉和变异的操作,使要解决的问题从初始解逐步逼近最优解,解决了许多全局优化问题。遗传算法通过编码将优化等问题从问题空间映射到操作空间,再通过译码将操作结果从操作空间映射回问题空间。对于具体的问题,常常有限制条件,即存在一个可行解空间。为了保证最后的解是可行的,可以采取两种方法:一是将可行解空间与操作空间一一对应;二是将可行解空间包含于问题空间中。遗传算法是依据随机技术来保证其寻优方向的确定算法。已有理论证明,不加改造的遗传算法不能达到全局最优,只能寻到全局最优的邻域。在达到该邻域后可采用局部寻优来最终达到全局最优点。文献[9]通过使用真实样本实验的方法,在 BP 神经网络、GA 遗传算法与改进的 GA-2BP 复合算法中,找出能迅速精确地预测航空发动机拆换期望值的最佳方法。文献[10]利用遗传算法确定各参数的权系数,使所得到的综合指数值比采用专家调查法所得综合指数值更能灵敏、准

确地反映发动机整体性能的变化情况,提高了发动机性能正常与否的识别率,并以此来发现发动机潜在早期故障,防止故障的扩大。

2.6 支持向量机

支持向量机是由 Vapnik 在上世纪 90 年代中期提出的一种机器学习算法。采用核函数方法将输入空间映射到一个高维内积空间中,通过解一个线性约束的二次规划问题得到全局最优解。支持向量机调整参数少,且模型结构由最能提供信息特征的支持向量点反映,可以大大降低建模的复杂性。支持向量机以其良好的理论背景,从结构风险最小化原则为机器学习提供了一个崭新的方向。支持向量机在最小化经验风险的同时,最小化置信风险,有效地保证了较强的推广特性且推广能力可以得到方便的控制^[11-13]。文献[11]提出了一种基于支持向量机的航空发动机滑油金属含量预测方法,并将其应用于某型发动机滑油的铁金属含量预测。又如文献[12]为了提高对航空发动机性能趋势预测的精度,利用支持向量机方法来预测表征发动机整体性能的参数-性能综合指数,建立了基于支持向量回归的一步及多步预测模型预测精度比 AR 模型更高,适合中、长期预测。

3 结束语

利用数据挖掘技术进行航空发动机的故障预测和健康管理,美国 NASA 航空安全大纲的蓝图已规划到了 2025 年^[13]。国内也已展开了诸多研究,以神经网络及其混合模型、时间序列分析等的应用案例居多,由于各类方法各有其优点及局限性,各种案例研究正在不断尝试、扩展和深入。

参考文献:

[1] 张叔农,黄明晟.数据挖掘技术综述[C]//中国航空学会可靠性工程委员会第十届学术年会论文集.北京:国防工业出版社,2006,7.

[2] Allan J Volponi ,Tom Brotherton ,Robert Luppold . Development of an information fusion system for engine diagnostics and health management [R] . AIAA 2004—6461.

[3] 韩庆田,杨兴根,张振家.神经网络在发动机状态监控与故障诊断中的应用综述[J].导弹与航天运载技术,2004,5:23—29.

[4] 胡金海,谢寿生.基于 AR 模型对滑油中金属元素含

量的预测[J]·燃气涡轮试验与研究,2003,16(1) .

[5] 江涌.时序分析在航空发动机故障诊断中的应用[J]·中国民航学院学报,2003,14(4) :26—29.

[6] 陈大光.燃气涡轮发动机状态监视与故障诊断[Z] .1988.

[7] Roemer M J, Kacprzynski G J . Advanced diagnostic and prognostic for gas turbine engine risk assessment[C] // Aerospace Conference Proceedings . 2000IEEE, Volume 6, Issue , 2000.

[8] 孙博,康锐,谢劲松.国外故障预测与健康管理系统的研究和应用现状综述[J]·系统工程与电子技术,2007,29(10) :1762—1767.

[9] 宋金光,许春生.基于遗传算法和BP 神经网络的发动机拆换期望值预测[J]·航空动力学报,2003,18(5) :676—680.

[10] 胡金海,谢寿生.基于遗传算法的发动机性能监控和故障诊断[J]·推进技术,2003,24(3) :198—200.

[11] 尉询楷,李应红,王硕,路建明,汪诚.基于支持向量机的航空发动机滑油监控分析[J]·航空动力学报,2004,19(3) :392—397.

[12] 胡金海,谢寿生,骆广琦,等.基于支持向量机方法的发动机性能趋势预测[J]·推进技术,2005,26(3) :260—264.

[13] Jonathan Litt , Donald L Simon , Claudia Meyer , et al . NASA aviation safety program aircraft engine health management data mining tools roadmap[R] . NASA/T M — 2000 — 210030. April 2000.

(上接第 163 页)

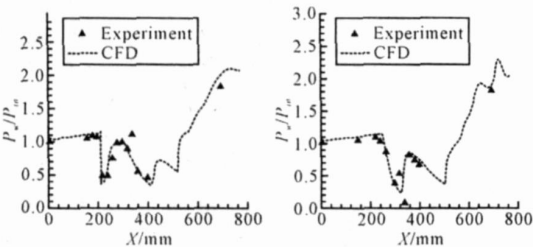


图 10 实验和数值模拟 上壁面压力对比

图 11 实验和数值模拟 下壁面压力对比



图 12 压力等值线



图 13 密度等值线

通过对实验状态 1 的数值模拟,对计算的实验模型壁面的静压值和实验数据进行了对比。图 10、图 11 分别给出了第一种状态的燃烧室上壁面与下壁面压力值与数值模拟的结果对比:隔离段内上下壁面压力均逐渐增高,在隔离段出口,由于上壁面有向后的台阶存在形成了膨胀波,故上壁面压力下降,而下壁面的压力下降位置要略为靠后一些,在膨胀波后又在壁面形成了压缩波,故上、下壁面压力均重新上升,在图 12、图 13 中可以明显地看到燃烧室内波系的分布和燃烧室尾部激波串的形成。在燃烧室的扩张段由于测点数量的限制,无法看到扩张段壁面压力的波动情况,但总体上计算结果和实验数据能够较好的符合。

5 结论

1) 相同出口反压(小反压值)条件下,不同的来流流量对燃烧室未受扰动部分的流场压力分布情况影响较小;

2) 燃烧室出口反压的增加,燃烧室内压力扰动逐渐向上游传播,扰动范围增加,隔离段可以在一定的程度上防止燃烧室出口反压升高对设备喷管流场的影响;

3) 相同反压下,相应地增加隔离段的长高比有助于减小反压对燃烧室流场引起的扰动;

4) 数值模拟和实验结果能够较好的符合。

参考文献:

[1] 刘陵.超音速燃烧与超音速燃烧冲压发动机[M]·西安:西北工业大学出版社,1993.

[2] 约翰·霍普金斯大学应用物理实验所.冲压发动机技术[M]·李存杰译.西安:国防工业出版社,1980.

[3] T O Mohieldin and S N Tiwari . Numerical study of 2D dual mode scramjet combustion part 1: Cold flow analysis[R] . AI AA 2003—7036.

[4] J Philip Drummond and Glenn S Diskin . Fuel air mixing and combustion in scramjets[R] . AI AA — 2002—3878.

[5] 余勇,丁猛,刘卫东,梁剑寒,等.煤油超音速燃烧的试验研究[J]·国防科技大学学报,2004,26(1) :1—4.

[6] C T Smith , C P Goynes , R H Krauss and J C McDaniel . Development of a three component particle image velocimeter technique for a scramjet combustor[R] . AI AA 2005—4104.