

液体火箭发动机稳态运行故障 数据聚类分析研究

张 翔, 徐洪平, 安雪岩, 耿 辉, 张素明
(北京宇航系统工程研究所, 北京 100076)

摘 要: 作为统计学的一个分支, 聚类分析已经被广泛研究了多年, 并形成了系统的方法体系, 利用基于聚类分析方法的数据挖掘在实践中已取得了较好效果。IMS 算法是 NASA 根据智能监控系统以及故障诊断技术的发展而提出的一种利用正常数据库来监控异常数据的可实时监控算法, 已经被 NASA 使用在各个方面, 收获了令人满意的效果。在分析美国航天飞机主发动机 (SSME) IMS 算法的基础上, 提出一种基于液体火箭发动机稳态运行故障数据的夹角余弦 IMS 聚类分析研究方法, 并对其进行仿真验证。

关键词: 液体火箭发动机; IMS; 聚类分析

中图分类号: V434-34 **文献标识码:** A **文章编号:** 1672-9374 (2015) 02-0118-05

Clustering analysis for fault data in steady process of liquid propellant rocket engine

ZHANG Xiang, XU Hong-ping, AN Xue-yan, GENG Hui, ZHANG Su-ming
(Beijing Institute of Space System Engineering, Beijing 100076, China)

Abstract: As a branch of statistics, clustering analysis has been widely studied for many years, and formed a system of the systematic method. The data mining based on clustering analysis method has achieved a good effect in practice. IMS algorithm presented by NASA according to the development of the intelligent monitoring system and fault diagnosis technology is a real-time monitoring algorithm to use normal database to monitor the abnormal data, which has been used by NASA in many aspects and gained a satisfactory effect. The included angle cosine IMS clustering analysis method based on steady process fault data of LRE is proposed in this article on the basis of analysis of the IMS algorithm of SSME. It was verified by simulation analysis.

Keywords: liquid propellant rocket engine; IMS; cluster analysis

收稿日期: 2014-07-17; 修回日期: 2014-12-19

作者简介: 张翔 (1990—), 男, 硕士研究生, 研究领域为航天运输系统故障诊断及健康管理

0 引言

液体运载火箭发动机的故障诊断与评估是确保火箭可靠性与安全性的的重要手段^[1-5], 而聚类分析研究是数据挖掘、故障诊断等技术的重要支撑内容^[6-7]。聚类是对数据对象的集合无指导地进行分组, 聚类算法的好坏直接影响聚类与故障诊断的效果^[8]。感应监视系统 (Inductive Monitoring System, IMS)^[9-10]纯粹从数据出发, 不必知道复杂的系统模型, 适合分析处在复杂太空环境的航天器。它主要用于深空探测的自主控制和状态监控, 目前已经成功应用于 NASA 的一些复杂系统, 如航天飞机健康监控、空间站控制力矩陀螺故障检测、运载火箭推进系统状态监控、新一代可重复使用空间运载器状态监控等。哥伦比亚航天飞机的 IMS 系统在起飞前就检测到了故障机翼温度的异常变化, 明显优于传统的模型推理健康监控方法。另外, 在故障前近 15 h, IMS 就检测到了国际空间站控制力矩陀螺 (ISS-CMG1) 的异常特征。

我国航天事业经过几十年蓬勃发展, 积累了

丰富的发动机飞行数据、试车以及仿真数据。这为开展基于数据分析的发动机故障诊断方法提供了基础^[11]。同时, IMS 良好的工程应用效果为开展基于故障数据聚类方法上提供了一定借鉴作用。本文在分析 IMS 对于正常数据聚类分析的基础上, 针对液体火箭发动机稳态运行故障数据射线状分布的特点, 提出一种基于故障方向夹角余弦的聚类分析算法, 并对其进行仿真实验。

1 基于稳态故障数据的聚类方法

IMS 是利用正常运行积累的历史数据或仿真数据^[12], 建立系统特征的正常知识库, 当实时数据偏离正常库一定范围时, 则认为出现了异常, 或者将会出现异常, 发出异常警报。本文针对我国液体发动机稳态运行的故障数据的特点, 对 IMS 算法提出一定的修改使其对不同发动机稳态运行故障模式数据类也同样具有良好的区分性。

本文利用某型液体火箭发动机仿真模型, 生成发动机在正常状况和常见故障状况下的仿真数据, 具体包括发动机的 12 个监测参数数据, 各参数名称与代号如表 1 所示。

表 1 某型液体火箭发动机仿真数据参数表

Tab. 1 Simulation data parameters of a liquid propellant rocket engine

序号	参数名称	参数代号	序号	参数名称	参数代号
1	氢涡轮转速	nh	7	燃烧室室压	Pc
2	氧涡轮转速	no	8	氢泵后压力	Phpe
3	燃烧室氢路喷前压力	Phc	9	氧泵后压力	Pope
4	燃烧室氧路喷前压力	Poc	10	发生器出口温度	Tg
5	发生器氢路喷前压力	Phg	11	氢泵流量	qh
6	发生器氧路喷前压力	Pog	12	氧泵流量	Qo

通过仿真发动机稳态过程运行得到的正常状况和常见故障状况下的监测参数数据点在参数空间中具有不同的分布位置, 建立故障方向, 并以新数据点到原点的方向与故障方向的夹角余弦 \cos 大小为依据进行聚类, \cos 定义为:

$$\cos = \frac{P_N - O \cdot \overrightarrow{center(i) - O}}{|P_N - O| \cdot |\overrightarrow{center(i) - O}|} \quad (1)$$

其聚类思路如下:

1) 算法初始化。选取正常工作过程中的一个样本点作为基点, 用来表征故障数据群的故障

方向原点。根据仿真数据，对于待聚类的一个新数据点 P_N 含有 12 个分量 $P_N(1), P_N(1) \dots P_N(12)$ 。如果没有建立过聚类，则建立聚类群 C_1 ，该聚类群是以原点 O 到 P_N 为故障方向，范围为与故障方向夹角余弦值大于 c 的区域，其中 c 为设定 $[0,1]$ 内一定值。

2) 判断是否落入某一聚类群范围 c 内。如果该新数据点落入多个聚类群范围内，则利用公式 (1) 分别计算该数据点 P_N 到原点方向与各聚类中心到原点方向的夹角余弦值，找出最大的夹角余弦值及其对应的聚类中心，并将该新数据点归为该数据类。如果由该数据点 P_N 计算得到的 \cos_{\max} 未落入任何一个聚类群的 $[0,c]$ 内，则建立一个新的聚类群 C_{i+1} ，该数据群是以原点 O 到 P_N 为故障方向，范围为与故障方向夹角余弦值大于 c 的区域。

3) 判断是否更新聚类中心。如果这个夹角余弦值 $\cos_{\max} \geq \delta > c$ (δ 为设定 $[0,1]$ 内一定值，且大于 c)，则不更新该数据中心；如果 $\delta > \cos_{\max} > c$ 则将 P_N 并入到该聚类群中，同时更新该聚类群的中心 $\text{center}(i)$ ，其更新公式为：

$$\text{center}(i)_{\text{new}} = \frac{\text{center}(i) + P_N}{2} \tag{2}$$

4) 判断是否到数据结尾。如果对整个数据完成聚类则结束算法，如果还未到达数据末尾，返回第 1) 步，进行下一帧数据迭代聚类。

2 仿真实验

为验证所提出的基于聚类分析的液体火箭发动机稳态过程故障程度评估方法，仍利用上文提到的某型液体火箭发动机仿真模型，分别生成：

- 1) 正常状态仿真数据：发动机在正常状况下的稳态过程仿真数据。
- 2) 单一故障理想状况仿真数据：发动机在稳态过程中发生处于 6 级故障严重程度（气蚀/泄漏率分别为 5%，10%，15%，20%，25%，30%）下的 6 种常见故障状况（故障 A：氢泵气蚀；故障 B：氢泵泄漏；故障 C：氧泵气蚀；故障 D：氧泵泄漏；故障 E：氢主阀泄漏；故障 F：氢副

控阀泄漏）的运行仿真数据。

3) 单一故障加干扰状况仿真数据：发动机在稳态过程中发生处于 4 级故障严重程度（气蚀/泄漏率分别为 5%，10%，15%，20%）下的 3 种常见故障（故障 A：氢泵气蚀；故障 B：氢泵泄漏；故障 C：氧泵气蚀）注入扰动干扰的运行仿真数据。为了保证数据的合理性，未添加随机噪声，而是注入其他故障模式的极小故障程度（故障模式 A-F，故障程度 0.1%~0.5%）作为干扰。

选取故障发生时刻之后 0.5 s 时刻的数据，并对上述三大类数据分别进行归一化处理，得到所需的仿真数据。并定义聚类准确度 ρ 作为衡量算法精确度指标，其定义式如下：

$$\rho = \frac{\text{正确分类的样本数}}{\text{总样本数}} \times 100\% \tag{3}$$

2.1 单一故障理想状况聚类仿真测试验证

将 2) 中得到的数据集进行乱序排列，选取不同的 c 值和 δ 值并进行聚类，结果见表 2。

表 2 单一故障理想状况聚类仿真测试结果
Tab. 2 Clustering simulation testing result for ideal state of single fault

序号	c	δ	聚类准确度 $\rho/\%$
1	0.4	0.5	83.33
2	0.4	0.7	83.33
3	0.6	0.7	100
4	0.6	0.9	100
5	0.8	0.9	100

从实验结果看，当 c 选取较小时聚类准确度 ρ 依然能超过 80%，而当选取 $c=0.6$ 时已能将所有的样本正确分类。这说明不同故障模式在空间分布差别是非常大的，也就是各个故障类的故障方向相互远离，易于分类。同时，可以注意到在 c 相同的情况下，改变 δ 的取值从上表中并未看到变化，经分析认为这是由于该组测试数据为理想数据。并且经过对所有经过归一化数据点集的 PCA 分析^[13-14]（如图 1 所示）可以得出各聚类间的空间分布（图 1 仅显示在前两维主元下的空间

分布), 也就是类内各数据在空间分布近似呈线性分布, 数据没有噪点, 没有波动。所以由 δ 决定的聚类中心变化范围非常小, 甚至可以忽略不计, 并未对整个聚类结果造成大的影响。

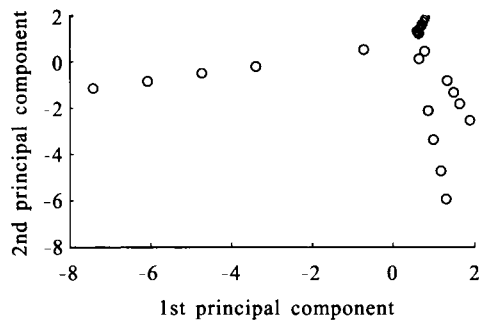


图 1 各聚类群在前二维主元坐标(PCA)下的空间分布
Fig. 1 Space distribution of clustering groups in first two principal components coordinate system

此处对图 1 仅能清晰显示出三组 (共 6 组聚类群) 聚类做出解释: 由于各种故障的发生位置不一致, 造成影响的测量参数 (表 1 所示) 也不一致, 所以其他聚类群可能在其他维数主元坐标下能够明显的区分出来。

2.2 单一故障加干扰状况聚类仿真测试验证

对 3) 得到的数据集进行乱序排列, 选取不同的 c 值和 δ 值并进行聚类, 结果如表 3 所示。

表 3 单一故障注入干扰状况聚类仿真测试结果
Tab. 3 Clustering simulation testing result of interference injected by single fault

序号	c	δ	聚类准确度 $\rho/\%$
1	0.3	0.4	74.66
2	0.3	0.5	74.66
3	0.4	0.5	74.66
4	0.4	0.6	91.55
5	0.4	0.8	91.55
6	0.5	0.6	91.55
7	0.7	0.8	83.11
8	0.8	0.9	83.11

从结果分析, 如果选择 c 值过小会造成分类准确度较低, 这是因为聚类群的范围过于宽泛, 可能新的样本点直接落入已有聚类群范围内, 而无法建立新的聚类, 造成错分样本较多, 并且在这种情况下改变 δ 的意义也不大。从第 3 和第 4 组可以看出, 在 c 不变的情况下, 改变 δ 的取值是可以影响聚类准确度的, 当 δ 从 0.5 提升到 0.6 时, 可以从结果看出提高了聚类准确度, 这是由于注入故障后数据点在准确值附近波动造成一定的偏向, 而设定 δ 的作用就是抑制数据的小范围波动, 而对大范围波动作用敏感, 起到了一定抑制噪声/扰动的作用。但从第 4 和第 5 组聚类结果看, δ 的效果是有限的, 一味提高 δ 值并不能更好地提高聚类准确度。此外, 从后两组结果看过分增大 c 值不能增加聚类准确度, 反而使其下降, 这就说明 c 值不能过小或者过大, 其与聚类准确度 ρ 的关系曲线如图 2 所示。

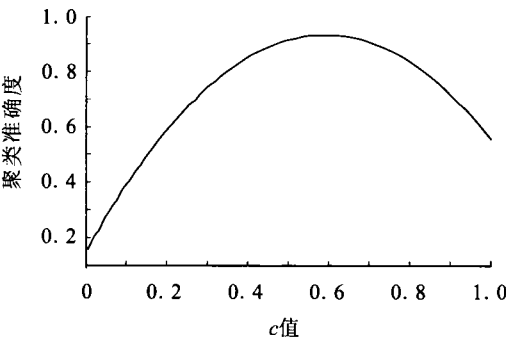


图 2 c 值与聚类准确度 ρ 关系曲线
Fig. 2 Relation curve of c value and clustering accuracy ρ

3 结束语

参考美国航天飞机主发动机 (SSME) 应用的数据分析算法 IMS, 提出一种基于液体火箭发动机稳态运行故障数据的角度余弦 IMS 算法聚类分析研究, 对其进行仿真验证, 为后续运载火箭发动机飞行及试车等数据分析提供了一定的借鉴意义。根据现阶段仿真结果来看, 对于故障模式的定性分析和聚类效果令人满意; 但是目前 c 值和 δ 值的选取仅凭经验, 选取精度较为宽泛, 还需要利用大量的数据和优化算法来选取出最优的 c 值和 δ 值, 这是下一阶段应着重开展的工作。

参考文献:

- [1] 冯辅周, 司爱威. 故障预测与健康管理技术的应用与发展[J]. 装甲兵工程学院学报, 2009, 23(06): 1-6.
- [2] 张育林, 吴建军. 液体火箭发动机健康监控技术[M]. 长沙: 国防科技大学出版社, 1998.
- [3] 代京, 张平, 李行善. 综合运载器健康管理健康评估技术研究[J]. 宇航学报, 2009, 30(4): 1711-1721.
- [4] 张振鹏. 液体火箭发动机故障检测与诊断中的基础研究问题[J]. 推进技术, 2002, 23(5): 353-359.
- [5] 夏鲁瑞. 液体火箭发动机涡轮泵健康监控关键技术及系统研究[D]. 长沙: 国防科学技术大学, 2010.
- [6] 杨小兵. 聚类分析中若干关键技术的研究[D]. 杭州: 浙江大学, 2005.
- [7] 章兢, 张小刚. 数据挖掘算法及其工程应用[M]. 北京: 机械工业出版社, 2006.
- [8] 刘红岩, 陈剑. 数据挖掘中的数据分类算法综述[J]. 清华大学学报: 自然科学版, 2002, 42(6): 727-730.
- [9] MARTIN RA, SCHWABACHER MA, MATTHEWS BL. Data-driven anomaly detection performance for the Ares I-X ground diagnostic prototype[R]. USA: NASA, 2010.
- [10] DAVID L I. Inductive monitoring system constructed from nominal system data and its use in real-time system monitoring, AIAA 2004-8062[R]. USA: AIAA, 2004.
- [11] 王莉. 数据挖掘中聚类方法的研究[D]. 天津: 天津大学, 2003.
- [12] IVERSON D L, MARTIN Rodney, SCHWABACHER Mark, et al. General purpose data-driven system monitoring for space operations[J]. Journal of Aerospace Computing, Information, and Communication, 2012, 9 (2): 26-44.
- [13] 袁志发, 宋世德. 多元统计分析[M]. 北京: 科学出版社, 2009.
- [14] 杨沛武. PCA 多元统计方法在过程监控中的应用研究[D]. 苏州: 江南大学, 2008.

(编辑: 王建喜)

(上接第 117 页)

5 结论

根据发动机设计部门对试验控制精度的要求, 进行了某型号姿轨控发动机控制系统的设计, 其控制模块用于操作人员对试验系统上的各路电磁阀进行手动与自动控制, 其复记模块可以实时记录系统状态以便于指挥人员了解系统工作状态并能够进行数据分析判断。该控制系统基于 FPGA 和高速固态继电器, 采用 PXI 平台以及上下位机模式, 控制精度为 0.1 ms, 能够满足试验需要。

参考文献:

- [1] 郭霄峰. 液体火箭发动机试验[M]. 北京: 中国宇航出版社, 1990.
- [2] 谢龙, 李海涛, 沈赤兵. 火箭发动机试验通用自动测试系统研究[J]. 计算机测量与控制, 2007, 15(4): 491-495.
- [3] 马爱华. 发动机试验数据自动采集处理系统研究与开发[J]. 计算机测量与控制, 2007, 15(7): 978-980.
- [4] 张志红. 可编程控制器(PLC)在发动机冷调试验中的应用[J]. 火箭推进, 2007, 33(1): 59-61.
- [5] 班超. FPGA 高精度时间测量[D]. 北京: 北京邮电大学, 2013.
- [6] 张朗. 基于 FPGA 的高精度测时仪研制[D]. 南京: 南京理工大学, 2012.
- [7] 戴彦, 申彦春, 王金红. FPGA 在小卫星姿态控制系统中的应用[J]. 电光与控制, 2009, 16(7): 76-79.
- [8] 李锦清, 刘文定, 赵鹏. FPGA 在卫星 AOCS 测试中的应用[J]. 测控技术, 2013, 32(12): 41-44.
- [9] 钱金川, 金林升. 固态继电器简介[J]. 机床电器, 2004, 31(4): 58-59.
- [10] 朱英明, 周长刚. 固态继电器(SSR)原理及应用中一些问题的探讨[J]. 机床电器, 2006, 33(1): 9-10.
- [11] 王昌华. 固态继电器及其应用[J]. 西北电力技术, 2004, 32(5): 9-52.
- [12] 邹进和, 李守明. 新型大电流直流固态继电器的设计[J]. 沈阳师范大学学报: 自然科学版, 2004, 22(4): 268-270.

(编辑: 陈红霞)