

文章编号: 1000-8055(2008)06-1127-04

基于模糊粗集的航空发动机特征参数提取算法

宋云雪, 张传超, 史永胜

(中国民航大学 航空工程学院, 天津 300300)

摘 要: 提取发动机性能状态的特征参数, 是提高发动机故障识别正确率和可靠性的必要条件. 针对实际航空发动机故障参数所具有的模糊和连续性特点, 提出了一种基于模糊粗糙集的特征参数提取算法, 并应用到某型航空发动机故障识别. 研究表明: 属性约简的核则为导致发动机故障的特征参数, 以此特征参数进行故障诊断, 可保证较高的诊断精度; 同时, 该算法的抗干扰性提高了整个系统故障识别的正确率. 该算法可用于航空发动机故障分类、故障诊断以及状态监控.

关 键 词: 航空、航天推进系统; 模糊集; 粗糙集; 特征参数; 故障识别

中图分类号: V231

文献标识码: A

Extracting algorithm for diagnostic parameters of aero-engine based on fuzzy-rough sets

SONG Yun-xue, ZHANG Chuan-chao, SHI Yong-sheng

(College of Aeronautical Engineering,

Civil Aviation University of China, Tianjin 300300, China)

Abstract: Extracting diagnostic parameters of aero-engine performance is an essential approach to enhance the validity and the reliability of aero-engine fault recognition. In view of fuzzy and consecutive characteristics of aero-engine fault parameters, an extracting algorithm based on fuzzy-rough sets was put forward and the algorithm was used for fault recognition of one aero-engine. The research result shows that the core attributes of reductions are diagnostic parameters of aero-engine failure, whereby the algorithm ensures high precision of fault diagnosis. The anti-interference ability of the algorithm enhances the validity of fault recognition. The algorithm can be also used for fault classification, fault diagnosis and performance monitoring of aero-engine.

Key words: aerospace propulsion system; fuzzy sets; rough sets; diagnostic parameters; fault recognition

由于航空发动机结构复杂、工作环境恶劣, 时刻都可能发生故障. 据国际民航组织的统计, 1988 - 1993 年的 6 年间, 由于发动机起火、叶片故障、发动机脱离机翼等原因而发生的飞行事故多达 34 起. 但是, 由于诸多非线性因素的影响, 现有

发动机状态监测中使用的参数只能部分地反映发动机性能状态, 无法进行故障识别、准确定位和预测. 如何能够从所获得的发动机工况参数、气动热力参数和机械性能等大量信息中找出能全面反映发动机性能状态的特征参数和新规律, 是提高发

收稿日期: 2007-06-08; 修订日期: 2007-09-12

基金项目: 天津自然科学基金(06 YFJ MJ C01700)

作者简介: 宋云雪(1968 -), 女, 辽宁鞍山人, 高工, 硕士, 主要从事结构智能化设计、故障诊断与结构修理的研究.

动机故障识别正确率和可靠性的必要条件. 本文将采用模糊集和粗糙集相结合的方法, 建立发动机故障识别的特征参数提取算法.

在航空发动机故障识别实践中, 决策表中用于识别的各参数通常是连续的, 而且很多情况下是模糊的. 而常规的粗糙集只能解决离散型属性的约简问题, 因此, 有必要研究如何将具有模糊性的连续属性值进行合理离散.

目前离散化方法较多, 如等距离划分算法、等频率划分算法等^[1], 这些算法使用起来较繁琐. 由于模糊集是处理模糊、不确定问题的有效工具, 具有聚类功能, 适合对连续属性值离散. 因此, 一些学者将粗糙集和模糊集结合来进行属性约简^[2], 该方法能够有效地克服人工离散连续性属性值所造成的主观性影响, 但是阈值的选取仍然具有较大主观性.

本文在用模糊粗糙集方法对具有连续条件属性值的信息系统进行约简时, 通过采用阈值优化策略, 克服属性值模糊聚类阈值选取的主观性, 从而得到最佳属性值分类, 为进一步确定导致航空发动机故障的关键特征参数奠定基础.

1 基于模糊粗糙集的发动机故障特征参数提取算法

1.1 航空发动机故障诊断模型

在构建航空发动机故障诊断模型时, 根据实际问题, 应选择具有物理意义明确、有较强分辨力和代表性的性能参数作为发动机故障识别的特征指标. 这些指标选择的优劣对发动机故障诊断结果会产生重要影响. 以某种飞机地面定检状态的故障诊断问题为例, 航空发动机故障诊断模型可表示成下列非线性方程

$$y = f(X_1, X_2, \dots, X_n) \quad (1)$$

式中 y 为输出值, 该值显示发动机状态; $X_1, X_2, \dots, X_n$ 为导致发动机故障的影响因素, 例如: 进口总温 T_{1^*} , 低压转子转速 N_1 , 高压转子转速 N_2 , 尾喷口指示值 Φ_{pc} , 滑油压力 P_M , 发动机机匣振动值 B 等.

假设 $T = (U, S, C, D)$ 是一个发动机故障信息系统, 其中 $U = \{u_1, \dots, u_n\}$ 是发动机故障数据记录的集合, 简称论域; $C = \{\text{进口总温, 低压转子转速, 高压转子转速, 尾喷口指示值, 滑油压力, 发动机机匣振动值}\}$ 是连续的条件属性集; $D = \{\text{发动机状态}\}$ 是离散的决策属性集; $S = C \cup D$ 为属

性集.

1.2 算法描述

基于模糊粗糙集的特征参数提取算法基本思想是首先采用模糊聚类方法对诊断模型中各因素指标进行分类与离散, 然后通过粗糙集约简算法进行特征参数的提取. 其中对事物进行分类模糊聚类方法由影响因素指标的选择、数据标准化、标定和聚类四部分组成. 具体算法描述如下:

1) 选择发动机故障因素指标 $\{X_n\}$ 和收集整理发动机故障样本数据; 从标准化方法中, 例如最大最小法、正态标准化法和归一化法等方法, 选择合适的方法对各指标数据进行标准化处理.

2) 在数量积法、夹角余弦法、相关系数法、最大最小法、算术平均最小法、几何平均最小法、绝对值指数法、指数相似系数法、绝对值倒数法、绝对值减数法、非参数法、贴近度法以及主观评定法等标定方法中, 选择一种方法进行各指标数据标定. 将论域中的元素两两之间都赋以 $[0, 1]$ 内的相似系数, 然后求模糊相似矩阵.

3) 从传递闭包法、最大树法和编网法等模糊聚类方法中, 选择一种方法将各指标参数分别进行聚类.

4) 用 F -统计量法或 λ 变化率法^[3] 分别确定各指标分类的最佳阈值 λ , 进而确定最佳分类.

5) 根据各因素指标的最佳分类, 分别将各模糊连续的因素指标进行离散, 提取适合粗糙集解决的决策表.

6) 从粗糙集的约简方法, 如基于可辨识矩阵和逻辑运算的约简算法、归纳属性约简算法、MFBARK 算法、基于特征选择的约简算法, 以及基于区分矩阵的约简算法中, 选择一种方法对决策表进行属性约简, 则约简的核就是航空发动机故障识别的特征参数.

2 航空发动机故障特征参数提取实例

2.1 样本的选取和标准化

选取某型发动机地面定检加力状态的 2 组正常数据和 7 组故障数据^[4] 为样本数据, 采用归一法进行标准化, 处理后的数据见表 1.

2.2 各属性进行标定和聚类

本文以“进口总温”的标定和聚类为例, 其它属性可以参照进行. 采用最大最小法计算相似系数, 求出模糊相似矩阵, 绘制最大树和动态聚类图, 见图 1 所示. 得出该属性的所有分类.

表 1 标准化后数据表
Table 1 Table of standardized data

U	T_1^*	N_1	N_2	Φ_{pc}	P_M	B
1	0.10	0.84	0.86	0.92	0.74	0.20
2	0.38	1.00	1.00	1.00	1.00	0.20
3	0.15	0.98	0.99	0.86	0.71	0.19
4	0.63	0.73	0.73	0.93	0.68	0.44
5	0.12	0.98	0.99	0.87	0.73	0.13
6	0.45	0.79	0.79	0.94	0.74	0.46
7	0.35	0.94	0.97	0.90	0.76	0.50
8	1.00	0.90	0.88	1.00	0.94	0.50
9	0.15	0.85	0.87	0.85	0.68	1.00

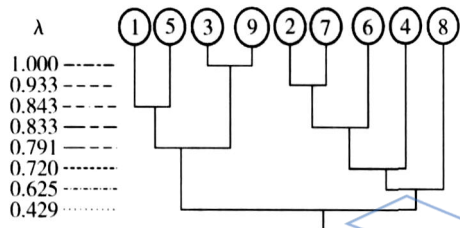


图 1 “进口总温”动态聚类图

Fig.1 Dynamic clustering figure of inlet temperature

2.3 最佳阈值λ的确定

根据属性动态聚类图,计算λ的变化率,确定最佳λ,从而确定各属性的最佳分类.根据表 2 λ变化率的最大值 $c_{max} = \infty$,对应的分类数为 4,对应的 $\lambda = 0.791$.按此方法可以得到其余属性的最优分类,见表 3.属性“进口总温”、“低压转子转速”、“高压转子转速”、“尾喷口指示值”、“滑油压力”、“发动机机匣振动值”分别以 $a \sim f$ 来表示.

表 2 “进口总温”分类数 c 及 λ 的变化率

Table 2 c and λ of inlet temperature

c	2	3	4	5	6	7	8
λ	0.625	0.720	0.791	0.833	0.843	0.933	1
c_{max}	0.058	0.013	∞	0.01	0.045	0.003 5	0

2.4 属性离散化和提取决策表

根据表 3,属性 $a \sim f$ 分别被分为 4,2,2,2,4,4 类.将所有的属性分别进行离散化,离散后的决策表见表 4.在表 4 中 $a = 0,1,2,3$ 分别表示“偏低、低、中等、高”; $b = 0,1$ 分别表示“低、高”; $c = 0,1$ 分别表示“低、高”; $d = 0,1$ 分别表示“小、

大”; $e = 0,1,2,3$ 分别表示“偏低、低、中等、高”; $f = 0,1,2,3$ 分别表示“偏低、低、中等、高”; $g = 0,1$ 分别表示发动机状态为“故障”或“正常”.

表 3 所有属性最优分类、最佳阈值 λ 和分类数 c

Table 3 λ, c and optimization sort of all attributes

属性	λ	c	分 类
a	0.791	4	$\{u_1, u_3, u_5, u_9\}\{u_2, u_6, u_7\}\{u_4\}\{u_8\}$
b	0.939	2	$\{u_1, u_2, u_3, u_5, u_6, u_7, u_8, u_9\}\{u_4\}$
c	0.908	2	$\{u_1, u_4, u_6, u_8, u_9\}\{u_2, u_3, u_5, u_7\}$
d	0.970	2	$\{u_1, u_3, u_4, u_5, u_6, u_7, u_9\}\{u_2, u_8\}$
e	0.970	4	$\{u_1, u_3, u_4, u_5, u_6, u_9\}\{u_2\}\{u_7\}\{u_8\}$
f	0.920	4	$\{u_1, u_2, u_3\}\{u_4, u_6, u_7, u_8\}\{u_5\}\{u_9\}$

2.5 基于粗糙集的属性约简

约简前的粗糙集决策表见表 4,论域为 $U = \{u_1, \dots, u_9\}$,条件属性 $C = \{a, b, c, d, e, f\}$,决策属性 $D = \{g\}$.采用基于区分矩阵的方法对决策表进行约简,区分函数 $\Delta^* = \{a, c, f\} \setminus \{b, d, c, f\} \setminus \{b, c, e, f\}$,因此,决策表有 3 个约简 $\{a, c, f\}, \{b, c, d, f\}$ 和 $\{b, c, e, f\}$,核为 $\{c, f\}$.

表 4 离散化后的数据表

Table 4 Table of dispersed data

U	a	b	c	d	e	f	g
1	0	1	0	0	0	1	0
2	1	1	1	1	3	1	0
3	0	1	1	0	0	1	1
4	2	0	0	0	0	2	0
5	0	1	1	0	0	0	1
6	1	1	0	0	0	2	0
7	1	1	1	0	1	2	0
8	3	1	0	1	2	2	0
9	0	1	0	0	0	3	0

2.6 算法对比验证

采用单纯粗集的约简算法对表 1 数据进行约简,约简为 $\{a, b, d, f\}, \{a, b, e, f\}, \{b, d, e, f\}$ 和 $\{a, c, e, f\}$,核为 $\{f\}$.通过对比,发现本文所提出的算法能够得到较小的约简,并且约简个数较少、冗余信息少.分别以第 7,8,9 个样本为测试样本,其余样本作为训练样本,以约简或核属性集中各属性值为神经网络的输入,发动机状态作为输出

表 5 测试精度和结果
Table 5 Precision and results of tests

约简和核	样本 7	样本 8	样本 9	<i>g</i>
{ <i>a, c, f</i> }	1.740 ×10 ⁻⁵	3.086 ×10 ⁻⁶	1.154 ×10 ⁻⁸	0
{ <i>b, c, d, f</i> }	4.945 ×10 ⁻¹	2.7 ×10 ⁻³	2.12 ×10 ⁻²	0
{ <i>b, c, e, f</i> }	3.400 ×10 ⁻³	1.1 ×10 ⁻³	3.878 ×10 ⁻⁵	0
{ <i>b, c, f</i> }	1.361 ×10 ⁻⁵	2.111 ×10 ⁻¹⁴	7.994 ×10 ⁻¹⁴	0
{ <i>c, f</i> }	7.780 ×10 ⁻⁹	1.765 ×10 ⁻⁴	6.263 ×10 ⁻¹⁰	0

进行对比计算,结果见表 5. 结果表明,测试结果均正确. 除含冗余属性较多的{ *b, c, d, f* }和{ *b, c, e, f* }外,其余的约简属性集和核属性集的测试精度都很高,进一步分析可以发现当 *c, f* 异常时,发动机出现故障,而 *a, b, d, e* 不是关键属性,因此可认为约简的核{ *c, f* }就是导致发动机故障的特征参数. 而采用单纯粗糙集的约简算法,核属性集漏掉了关键属性 *c*,这充分说明本文采用阈值优选的策略进行属性离散的合理性和正确性.

将该算法用于航空发动机故障状态识别,选取了 5 个待测样本,采用本文算法提取特征参数,然后进行模式识别. 对比单纯的模式识别方法,识别的正确率由 60 % 提高到 80 %. 结果表明,采用本文算法提取特征参数后,去除了干扰信息的影响,在保证诊断精度的同时,提高了故障识别的正确率,这充分证明了该算法具有较强的抗干扰性. 当典型模式足够多时,识别精度会进一步提高.

3 结 论

1) 在航空发动机故障识别过程中,使用该算法所提取的参数集合为整个故障参数集合中较小的约简;约简的核就是导致发动机故障的特征参数,以此特征参数进行故障诊断,可保证较高的诊断精度.

2) 该算法的抗干扰性,在保证诊断精度的同

时,提高了故障识别的正确率;当典型模式足够多时,识别的精度会进一步提高.

3) 该算法采用阈值优选策略,以阈值λ变化率的最大值作为依据,得到各具有连续性的属性最优分类,进一步克服了模糊聚类中选取阈值和单纯粗糙集离散化的主观性,增强了约简的客观性、科学性和合理性.

4) 该算法也适用于具有连续决策属性的航空发动机故障识别中特征参数的提取;在航空发动机故障识别与分类、健康状态监控、故障诊断等方面有广阔的应用前景.

参考文献:

[1] 王国胤. Rough 集理论与知识获取[M]. 西安:西安交通大学出版社,2001:99-116.

[2] 刘靖,陈福生,张勤. 基于粗糙集和模糊集的属性约简算法[J]. 计算机工程与科学,2005,27(2):42-44.
LIU Jing, CHEN Fusheng, ZHANG Qin. A reduction algorithm based on rough and fuzzy sets[J]. Computer Engineering & Science, 2005, 27(2):42-44.

[3] 谢季坚,刘承平. 模糊数学方法及其应用[M]. 武汉:华中科技大学出版社,2005:85-98.

[4] 王广,李军. 基于粗糙集理论的航空发动机故障诊断[J]. 航空发动机,2005,31(4):51-53.
WANG Guang, LI Jun. Aeroengine fault diagnosis based on rough sets theory[J]. Aeroengine, 2005, 31(4):51-53.