

Baseline Modeling of Engine Variable Bleed Valve Based on Gene Expression Programming Algorithm

Tao Sun¹, Runze Zhang², Nanxiang Li³

¹Basic College, Naval Aviation Academy, Yantai Shandong

²The Unit 91899 of Army, Huludao Liaoning

³Naval Staff, Beijing

Email: happyli.dong@163.com

Received: Feb. 22nd, 2019; accepted: Mar. 6th, 2019; published: Mar. 13th, 2019

Abstract

Aiming at the problem that it is difficult to obtain the characteristic index of engine operation valves, adopting gene expression programming algorithm (GEP), using the quick access recorder (QAR) data, this paper mines and analyzes the engine variable bleed valves (VBV) operation baseline. The results show that, during VBV nonlinear regression analysis, GEP has the advantages of little deviation and high precision. GEP can also get specific and convenient function relations, which can improve the accuracy of VBV monitoring and fault diagnosis effectively.

Keywords

Variable Bleed Valve, Gene Expression Programming Algorithm, Quick Access Recorder, Multi-Parameter Regression

基于基因表达式算法的某型发动机可调放气阀门基线建模

孙 涛¹, 张润泽², 李南翔³

¹海军航空大学基础学院, 山东 烟台

²91899部队, 辽宁 葫芦岛

³海军参谋部, 北京

Email: happyli.dong@163.com

收稿日期: 2019年2月22日; 录用日期: 2019年3月6日; 发布日期: 2019年3月13日

摘要

本文针对航空发动机工作阀门的特性指标难以获取的问题,采用基因表达式编程算法(gene expression programming, GEP),利用飞机快速存储记录器(quick access recorder, QAR)数据对发动机的可调放气阀门(Variable bleed valve, VBV)工作基线进行挖掘分析。结果表明,基因表达式编程在处理阀门的非线性回归分析时,构建的模型非常接近运行规律,具有偏差小、精度高的特点,并且能得到具体、方便实用的函数关系,能够有效地提高发动机阀门监控和故障诊断的准确性。

关键词

可调放气阀门, 基因表达式编程算法, 快速存储记录器, 多参数回归

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

对于航空发动机来讲,压气机的喘振会导致强烈的机械振动和热端超温,在极短的时间内就可能会造成发动机部件的严重损坏。现役的涡扇发动机多采用可调放气阀门(variable bleed valve, VBV)放出部分压气机空气以达到防喘的目的,因此研究可调放气活门对于发动机状态监控,及时发现不良变化,预防飞行事故具有重大的意义。

目前国内所用的涡扇发动机基本由国外公司生产,其部件特性、结构参数与控制特性等均无从得知,无法建立发动机进行视情监控和故障诊断的数学模型。各航空公司在使用发动机的过程中,积累大量可用数据,基本上涵盖了发动机气路、控制、振动、油路等多方面的性能参数。因此通过对这些数据的挖掘,利用系统的输入输出参数对发动机的系统进行辨识,构造一个相对真实的反映对象的数学模型,以大数据平台为数据基础,结合先进的智能分析方法,让客观数据说话,尽量避免人为干预,建立客观标准的大数据模型,是人工智能的一个重要应用。

由于航空发动机的各个系统均为复杂的非线性系统,传统的辨识系统难以取得较好的结果,基因表达式编程、神经网络、支持向量机等智能算法能在有限样本条件下,在解决非线性高维模式识别问题方面,表现出较高的准确性与泛化性[1] [2]。相对于神经网络、支持向量机等智能算法,使用基因表达式编程进行函数挖掘可以得到比较明确的函数模型,具有误差小、精度高等特点,适于涡扇发动机可调放气阀门工作基线建模[3] [4]。文献[5]采用支持向量机对模拟线路的故障进行诊断。对原始数据进行非线性变换,将非线性不可分的原始数据变换到高维线性可分的空间中,通过在高维空间上的距离构造真实度和虚假度函数,并以这两个函数为基础进行分类,取得较好的分类结果,但运算时间较长。文献[6]采用粗糙集和支持向量机的方法研究了发动机视情维修问题。利用粗糙集对原始数据进行处理,去除异常样本,进一步用支持向量机方法确定发动机不同维修等级。但支持向量机方法只根据一类数据的信息,没有充分融合不同时间的数据。

本文利用基因表达式编程算法,建立发动机可调放气活门基线模型。并用所建模型进行计算,得到的均方差为 0.1208,满足精度要求,本文所用的方法具有较强的工程实践价值。

2. 基因表达式编程算法

基因表达式编程(GEP)是基于生物基因结构和功能发明的新型自适应演化算法。GEP 是从遗传算法和遗传程序设计中发展而来, 具有遗传算法的优点, 编码方式较为简单, 易于遗传修饰的操作, 并且具有结构和功能上的多样性。

基因表达式编程的基本步骤如图 1 所示。该过程从随机产生一定数量的染色体个体(初始种群)开始, 然后对这些染色体进行表达, 依靠一个适应度样本集(也被称为选择环境, 即问题的输入)计算出每个个体的适应度(算法中适应度采用相关系数 R 的平方)。然后个体按照其适应度被选中, 进行有修饰的复制, 留下具有新特性的后代。接下来, 这些新的个体也要经历相同的发展过程: 基因组的表达, 面临选择环境, 选择和有修饰的复制, 该过程重复若干代, 直到发现一个优良解为止[7] [8] [9]。

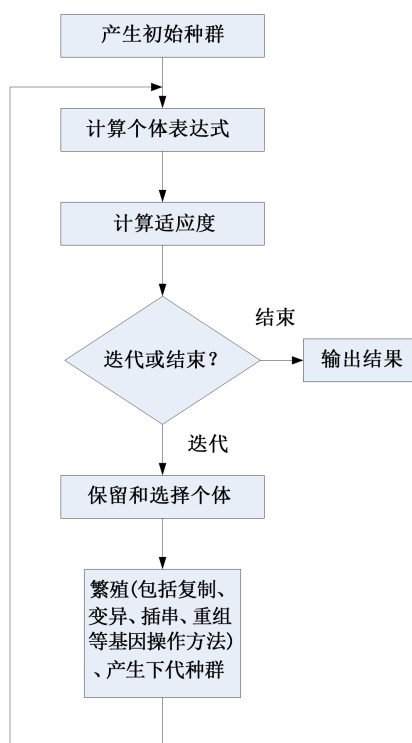


Figure 1. GEP algorithm flowchart

图 1. GEP 算法流程

3. VBV 工作基线函数的 GEP 算法设计

所谓 VBV 工作基线是指处于良好工作状态的发动机在标准状态下, VBV 与其控制参数之间的函数关系。本文用于 VBV 基线挖掘的数据平台来源于飞机的快速存储记录器(QAR) [10]。QAR 是一个包含大量航班数据的数据库, 它基本涵盖了发动机控制、气路、滑油以及振动等多方面的性能参数。

根据航空发动机的压气机工作原理, 排除影响 VBV 开度等次要参数, 确定与 VBV 密切相关参数为 N1K12 (低压转子修正转速与额定转速的百分比); N2K25 (高压转子修正转速与额定转速的百分比); VSV (可变静子叶片的角度, 单位为度); TAT (外界大气温度, 单位为摄氏度); T2.5 (高压压气机进口温度, 单位为摄氏度), 其工作基线模型简化为:

$$VBV = f(N1K12, N2K25, VSV, TAT, T2.5) \quad (1)$$

在建立 VBV 模型样本时, 为更好的研究 VBV 的工作基线, 依据以下原则进行数据筛选: 选用发动机循环数较少的数据; 选用发动机性能良好的数据; 剔除有显著差异的数据。基于这些原则, 本文选择了一架机左、右两台发动机(两台发动机的装机时间均为 2016 年 5 月, CSN 均为 4200), 收集在 2018 年 1 月的 25 个航班, 共 50 组 QAR 解码以后的数据。每次航班的数据由于飞行时间不同, 共有 7500 到 12000 条数据不等。在使用 GEP 算法进行基线函数挖掘之前, 数据经过了清洗、筛选、填充和压缩等预处理操作, 保证数据可靠和统一。

将 GEP 算法用于工作基线函数挖掘具体工作, 就是将 5 个 VBV 控制参数和可能的函数关系构造到初始染色体中, 编码形成表达式树, 再由实际控制参数进行适应度评判。通过对染色体进行变异、插串、根插串、单点重组、两点重组、基因重组等操作方式, 使染色体不断改变, 越来越适应 VBV 数据所呈现的规律, 直到最终匹配。将适应度最高的染色体译码出来, 即为工作基线函数。

本文编写 GEP 算法的工具使用的是 MATLAB 语言, 算法中具体设置如表 1 所示。

Table 1. Parameters of GEP algorithm

表 1. GEP 算法参数设置

迭代次数	300
种群大小	1000
函数集	+, -, *, /, d/dt, sqrt(), ln(), abs(), cos(), exp()
尾部集	a, b, c, d, e
头部长度	8
染色体上基因个数	9
基因连接方法	+
染色体长度	153
变异概率	0.04
单点重组概率	0.3
两点重组概率	0.3
基因重组概率	0.1
插串概率	0.1
插串长度	1,2,3
根插串概率	0.1
根插串长度	1,2,3
适应度计算方法	确定系数计算
精度要求	不小于 0.95

算法中适应度的判别采用确定系数 R^2 , 即相关系数 R 的平方。当 R^2 的值越接近 1, 则函数拟合优度越高, 相反则不适用。拟合优度越高, 自变量对因变量的解释程度越高, 即自变量引起的变动占总变动的百分比越高。当迭代次数超过 1000 或者 R^2 大于等于 0.95 时, 算法终止, 程序跳出, 此时的结果即为最后的拟合式。

以确定系数为适应度判别, 让进化过程朝着拟合优度越来越大的方向进化, 迭代计算的适应度变化曲线如图 2 所示。

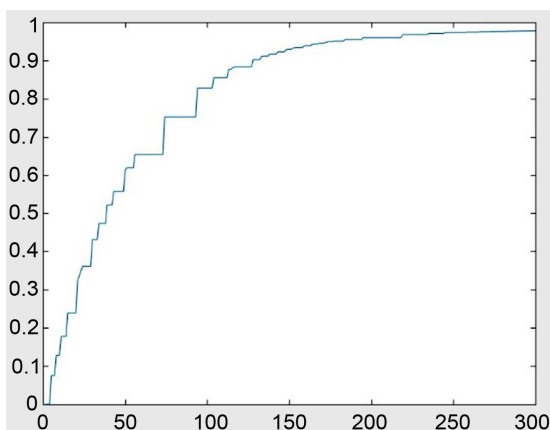


Figure 2. Variation curve of fitness

图 2. 适应度变化曲线

当适应度为 0.98 时，原 QAR 实际数据和利用 GEP 算法模拟出的曲线及其误差如图 3 所示。

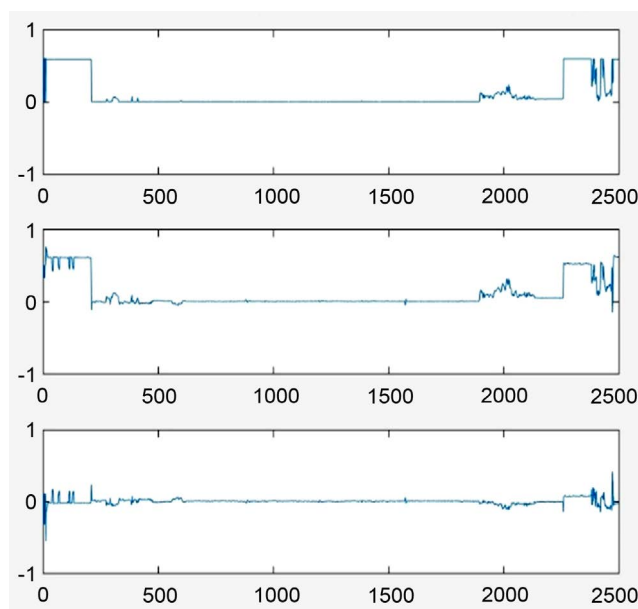


Figure 3. The original measured curve, fitting curve and their errors

图 3. 原实测曲线、拟合曲线及两者误差

其中，图中数据为 10000 条数据的 4 倍压缩。由图 3 可知，GEP 算法模拟出的曲线和 QAR 实际数据形状很接近，其误差很小。

4. VBV 工作基线函数曲线拟合结果分析

GEP 算法以一台发动机的数据经过 300 代的迭代计算后，确定的 VBV 工作基线为：

$$VBV = 0.2957 + 1.15 \left| d^{\frac{4}{3}} \right| + 0.2ad^2 + 0.6714d + 0.1126a + 0.0043b + 0.0049c + 0.0047e \quad (2)$$

其中： a, b, c, d, e 分别代表 N1K12, N2K25, TAT, VSV, T2.5 这 5 个控制参数。从式(2)可以看出，VBV 与 VSV 关系最为密切，同时也受 N1K, N2K, TAT, T2.5 四个参数的线性影响。

为了深入讨论算法拟合程度, 将基线方程应用到其它发动机数据中进行验证, 验证时采用均方根误差(RMSE)。均方根误差用来反映所得的函数数据和实际数据之间的离散程度。

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3)$$

式(3)中, \hat{y} 为 GEP 公式计算得到的值, y 为真实值。

使用 50 组发动机数据, 同时为了增加对比, 利用 GEP 公式和神经网络方法(NN, 初始值设置同 GEP)得到的检验结果如下表 2。

Table 2. RMS error obtained GEP method

表 2. GEP 方法得到的均方根误差

Data	NN	GEP
20180107-1-1	0.2039	0.1059
20180107-1-2	0.1579	0.1079
20180107-2-1	0.1921	0.1312
20180107-2-2	0.2341	0.1311
20180108-1-1	0.2050	0.1050
20180108-1-2	0.1902	0.1102
20180108-2-1	0.1329	0.1089
20180108-2-2	0.2184	0.1104
20180108-3-1	0.3168	0.1178
20180108-3-2	0.2154	0.1164
20180109-1-1	0.3685	0.1285
20180109-1-2	0.2289	0.1289
20180109-2-1	0.3882	0.1082
20180109-2-2	0.3088	0.1088
20180109-3-1	0.3238	0.1288
20180109-3-2	0.3360	0.1260
20180109-4-1	0.3399	0.1099
20180109-4-2	0.4415	0.1125
20180109-5-1	0.5274	0.1274
20180109-5-2	0.3443	0.1443
20180110-1-1	0.6173	0.1173
20180110-1-2	0.5125	0.1175
20180110-2-1	0.6897	0.1287
20180110-2-2	0.4252	0.1252
20180111-1-1	0.5230	0.1150
20180111-1-2	0.4341	0.1141
20180111-2-1	0.4661	0.1521

Continued

20180111-2-2	0.4567	0.1517
20180112-1-1	0.2969	0.1069
20180112-1-2	0.5415	0.1115
20180113-1-1	0.5377	0.1077
20180113-1-2	0.5360	0.1060
20180113-2-1	0.5237	0.1057
20180113-2-2	0.4258	0.1078
20180113-3-1	0.5200	0.1200
20180113-3-2	0.4235	0.1235
20180113-4-1	0.3273	0.1273
20180113-4-2	0.4502	0.1202
20180114-1-1	0.6719	0.1009
20180114-1-2	0.4501	0.1031
20180114-2-1	0.3652	0.1050
20180114-2-2	0.4091	0.1090
20180114-3-1	0.3021	0.1038
20180114-3-2	0.2301	0.1036
20180114-4-1	0.2887	0.1077
20180114-4-2	0.3000	0.1060
20180115-1-2	0.3407	0.1070
20180115-2-1	0.4124	0.1094
20180115-2-2	0.3401	0.1095
20180115-3-1	0.3287	0.1045

由表 1 中的结果可知, 50 组数据利用 GEP 函数表达式进行验证得到的结果相差不大, 说明拟合函数总体可靠。RMSE 在 0.1050 到 0.1443 之间, 平均为 0.1208, 表明拟合函数数据分散在实际曲线附近, 离散度很小, 与实际曲线拟合很高。而神经网络方法误差明显高于 GEP 算法, 误差基本在 0.2 以上, 个别达到了 0.6, 可见 GEP 算法优于神经网络方法。

5. 结论

本文通过对某型发动机 QAR 中的数据进行挖掘, 利用 GEP 算法得到了一个相对准确的发动机可调放气阀门(VBV)的工作基线数学函数。计算结果表明, GEP 算法挖掘得到的均方根误差为 0.1208。在有限样本条件下, 对于民航涡扇发动机 VBV 工作基线建模具有较高的准确性, 构建的模型可以满足对可调放气活门进行监控和故障诊断的需要。

参考文献

- [1] Ferreira, C. (2017) Gene Expression Programming: A New Adaptive Algorithm for Solving Problems. *Computer Science*, 21, 87-129.

-
- [2] Ferreira, C. (2002) Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence. *Engineering Applications of Artificial Intelligence*, **1**, 223-225.
- [3] 朱耀春. 基于基因表达式编程技术的非线性系统辨识研究[D]: [博士学位论文]. 北京: 华北电力大学, 2008.
- [4] 元昌安. 基于 GEP 函数发现的智能模型库关键技术研究[D]: [博士学位论文]. 成都: 四川大学, 2006.
- [5] Luo, H., Wang, Y.R. and Cui, J. (2011) A Vague Description Method for Analog Circuit Fault Diagnosis Based on Description Sphere. *Chinese Journal of Aeronautics*, **24**, 768-776. [https://doi.org/10.1016/S1000-9361\(11\)60091-5](https://doi.org/10.1016/S1000-9361(11)60091-5)
- [6] 李冬, 宋岩, 马力, 等. 基于粗糙集和支持向量数据描述的发动机视情维修研究[J]. 燃气轮机技术, 2013, 26(1): 46-50, 58.
- [7] 王欢. 涡扇发动机执行机构建模与控制方法研究[D]: [硕士学位论文]. 南京: 南京航空航天大学, 2012.
- [8] 钟诗胜, 崔智全, 付旭云. Rolls-Royce 发动机基线挖掘方法[J]. 计算机集成制造系统, 2010, 16(10): 2265-2270.
- [9] 林兆福, 范作民. 发动机基线方程的建立和应用[J]. 中国民航大学学报, 1992(4): 20-32.
- [10] 孔成安, 李文华, 尹湛. 利用 QAR 数据实施飞机性能监控[J]. 中国民用航空, 2008(10): 54.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2161-8801, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: csa@hanspub.org