

DOI:10.19651/j.cnki.emt.2518392

基于特征优选及 Informer 的航空发动机剩余寿命预测*

卢翔¹ 高新越¹ 王杜² 康千灼¹ 何晟²

(1. 中国民航大学交通科学与工程学院 天津 300300; 2. 中国民航大学航空工程学院 天津 300300)

摘要: 航空发动机健康监测传感器众多,传感器选择是否得当,将直接影响发动机剩余寿命预测效果。提出一种基于特征优选的传感器选择方法,结合 Informer 预测剩余寿命,提高了预测的精度。首先利用差分聚类算法对真实的飞行工况进行分类,接着由航空发动机退化机理构建健康因子,利用巡航阶段的数据建立回归树模型,选出重要的传感器,最后基于 Informer 进行航空发动机剩余寿命的预测。利用 NASA 最新发布的真实飞行条件下航空发动机退化数据库进行了实验,实验结果表明:与不选择传感器相比,所提方法的预测结果均方根误差下降了 14%,平均评分函数下降了 29%;与传统的依据传感器退化趋势或传感器变化差异二种选择方法相比,均方根误差分别下降了 10%、8%,平均评分函数分别下降了 48%、27%;将提出的剩余寿命预测方法与 CatBoost、LightGBM、XGBoost、BiLSTM 和 Transformer 算法相比,精度分别提升了 36%、24%、14%、6%和 5%。

关键词: 剩余寿命预测;特征优选;Informer;回归树;航空发动机

中图分类号: TP391.5;TN807 **文献标识码:** A **国家标准学科分类代码:** 520.2

Civil aviation engine remaining useful life prediction on optimal feature selection and Informer

Lu Xiang¹ Gao Xinyue¹ Wang Du² Kang Qianzhuo¹ He Sheng²

(1. College of Transport Science and Engineering, Civil Aviation University of China, Tianjin 300300, China;

2. College of Aeronautical Engineering, Civil Aviation University of China, Tianjin 300300, China)

Abstract: There are many sensors for civil aviation engine health monitoring. The proper choice of sensors will directly affect the prediction effect of engine remaining useful life. A sensor selection method based on optimal feature selection is proposed and Informer algorithm is used to predict the remaining useful life, which improves the prediction accuracy. Firstly, the differential clustering algorithm is used to classify the real flight conditions, and the health factors are constructed from the degradation mechanism of civil aviation engine, and the regression tree model is established with the data of cruise stage to select important sensors. Finally, the remaining useful life of civil aviation engine is predicted based on Informer algorithm. Using NASA's newly released civil aviation engine degradation database under real flight conditions, the experimental results show that the root mean square error of prediction results decreases by 14% and the average scoring function decreases by 29% compared with no sensor selection. Compared with the traditional selection method based on sensor degradation trend or sensor data difference, the root mean square error decreases by 10% and 8% respectively, and the average scoring function decreases by 48% and 27% respectively. Compared with CatBoost, LightGBM, XGBoost, BiLSTM and Transformer algorithms, the accuracy of the proposed remaining life prediction method is improved by 36%, 24%, 14%, 6% and 5%, respectively.

Keywords: remaining useful life prediction;optimal feature selection;Informer;regression tree;civil aviation engine

0 引言

航空发动机剩余寿命预测是发动机预测与健康管理 (prognostics and health management, PHM) 中的一项核心

技术^[1],为发动机运行安全性、经济性和预测性维修提供重要的理论支持。

基于神经网络的剩余寿命预测技术具有建模简单、费用投入低且预测效果好的特点,并且能够适应民航发动机

收稿日期:2025-03-19

* 基金项目:国家自然科学基金(U2433213)、中国民航大学研究生科研创新项目(2023YJSKC08009)资助

集成化、系统化和精细化的发展方向,引起了国内外研究者高度重视。当前,国内外学者以卷积神经网络(convolutional neural network, CNN)^[2-7]、长短时记忆神经网络(long short term memory, LSTM)^[8-13]、门控神经单元(gated recurrent unit, GRU)^[14]、生成对抗网络(generative adversarial network, GAN)^[15]为基础,对剩余寿命预测的神经网络模型开展研究。CNN 通过卷积核捕捉时间序列的局部特征,擅长处理多维数据,但是网络受限于卷积核的大小,捕获远程信息的能力有限。LSTM 虽然可以高效地学习时间序列数据,但是易受噪声影响且对于高维数据的非线性拟合效果不佳。GRU 计算效率较高,但更依赖近期信息,不擅长处理长时间序列退化数据。GAN 适合航空发动机故障数据稀缺的场景,然而,其训练不稳定,对多传感器数据的特征融合能力有限,剩余寿命预测精度和稳定性难以保证。

2017 年,产生了一种完全基于注意力机制的神经网络模型,即 Transformer 模型^[16],此模型在自然语言处理领域与当时最好的翻译模型相比提高了 2 个双语互译质量分数并且具有更强的并行速度,Transformer 也逐渐应用到航空发动机剩余寿命预测领域^[17-19]。2021 年,Zhou 等^[20]基于 Transformer 提出 Informer 模型,对电网变压器的负载进行了预测,该模型进一步降低了计算复杂度、减少了维度和网络参数量并且能够一步获得所有预测结果,极大地提高了计算效率与预测精度。Informer 模型在回归类任务中具有显著的预测潜力,但是基于 Informer 的航空发动机剩余寿命预测的研究仍然较少。

特征的选择对于减少计算的复杂度,提高模型的预测效果,降低发动机运营成本具有关键性的影响。在航空发动机故障诊断领域,学者从传感器之间的线性相关性、传感器的敏感度和传感器间的相互作用等方面研究传感器的选择对于预测模型的影响。Hou 等^[21]考虑传感器间的线性相关性,通过计算皮尔逊相关系数,从 44 个传感器中选择 6 个重要的传感器,构建了航空发动机健康监测模型,提高了燃气轮机健康诊断模型的适用性;Giorgi 等^[22]对传感器进行了灵敏度分析,剔除了低敏感度的传感器,使模型的计算复杂度下降了一半,同时提高了模型的适应度;Urbanowicz^[23]考虑传感器之间的交互作用对计算特征的重要影响,使预测模型可以灵活处理各类不同的数据,提高了计算的效率。

相比于在航空发动机故障诊断中的传感器选择方法,目前在剩余寿命预测领域中的传感器选择方法仍是传统的依据传感器退化趋势选择传感器^[24-25],或者从适航的角度出发,考虑发动机监测数据的变化差异选择传感器^[26]。这两种方法或是基于主观的判断,或是只考虑到单一的数据变化信息,缺少特征选择,不能保证传感器选择的有效性。而在使用深度学习模型进行寿命预测的过程中,由于部分传感器的低敏感度、传感器之间的线性相关性或者传感器

之间的相互作用,会出现信息冗余、计算量大或数据冲突的问题,进而导致模型的预测精度低。因此,传感器的选择对航空发动机剩余寿命预测来说至关重要。

基于此,本文提出一种基于特征优选及 Informer 的航空发动机剩余寿命预测方法,考虑民用航空发动机性能退化原理^[27],由发动机的气路部件性能参数得到健康因子,建立测量参数与健康因子树模型,筛选重要的传感器,完成特征优选,然后基于 Informer 对航空发动机进行剩余寿命预测。

1 剩余寿命预测方法

1.1 基于健康退化数据的剩余寿命预测流程

基于健康退化数据的剩余寿命预测流程如图 1 所示。整个预测流程分为 3 个步骤:第 1 步,对航空发动机历史退化数据进行数据预处理;第 2 步,进行特征优选,选择对航空发动机健康状态最重要的传感器;第 3 步,建立航空发动机剩余寿命预测模型。

在数据预处理阶段,将训练数据用于建立传感器选择模型,测试数据直接应用于寿命预测。考虑到在真实的飞行条件下,传感器数据中会包含噪声,因此首先需要对数据进行降噪处理。又由于航空发动机工况复杂,为了减少计算量以及方便选择传感器,利用差分聚类算法对航空发动机的工况进行分类。

在特征优选阶段,主要考虑航空发动机健康退化的原理,构建航空发动机健康因子,建立回归树来计算航空发动机传感器重要度,同时对回归树进行超参数优化,选择最优的超参数即最小叶子数、最大分裂次数与构建回归树的变量个数,最后选取最重要的 5 个传感器应用于寿命预测。

在剩余寿命预测阶段,主要基于 Informer 构建寿命预测模型,若是寿命预测的精度不符合要求,则调整编码器输入序列的长度,以及解码器中输入序列的长度(包含先验序列长度,预测序列长度),确定最优预测模型参数。

1.2 工况聚类

在数据预处理阶段,利用差分聚类算法对航空发动机工况进行分类,划分发动机在爬升、巡航、下降 3 个飞行阶段下的工况,减少回归树的计算量,差分聚类算法具体如式(1)~(5)所示。

$$Deta = W_{alt}(i+1) - W_{alt}(i) \quad (1)$$

$$\{alt_i\} = find(\{Deta \mid \leq a\}) \quad (2)$$

$$climb = [alt_s, \min(\{alt_i\})] \quad (3)$$

$$cruise = [\min(\{alt_i\}), \max(\{alt_i\})] \quad (4)$$

$$decend = [\max(\{alt_i\}), alt_e] \quad (5)$$

式中: $W_{alt}(i+1)$ 为时间序列下第 $i+1$ s 的飞行高度, $W_{alt}(i)$ 为时间序列下第 i 秒的飞行高度, $\{alt_i\}$ 为民用飞机巡航阶段的飞行数据序列, a 为民用飞机巡航高度的波动阈值, $climb$ 为民用飞机爬升时的飞行数据序列, alt_s 为开始爬升时刻的传感器数据, alt_e 为着陆时刻的传感器数

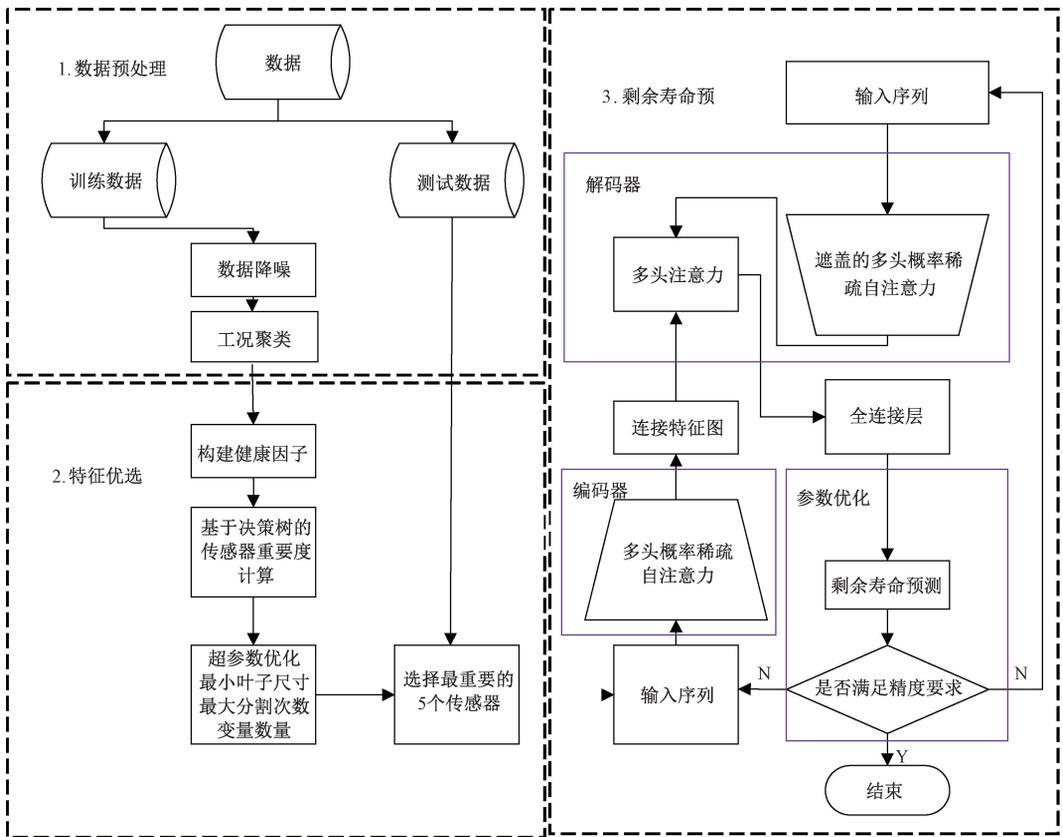


图 1 航空发动机剩余寿命预测流程图

Fig. 1 Flow chart of RUL prediction of civil aircraft engine

据。此算法能够方便、高效地划分民用飞机在真实飞行条件下爬升、巡航、下降的传感器数据。

1.3 特征优选

1) 健康因子的构建

航空发动机健康状态的退化主要是由气路部件故障引起的。对于航空发动机来说,气路部件表面的结垢、灰尘以及水滴等侵蚀气路部件表面的物质、燃烧气体与空气污染物对气路部件的腐蚀、叶尖间隙的增加以及物体的打伤都会造成发动机物理损伤。随着这些物理损伤的累积,会渐渐地造成单个或多个气路部件的性能参数发生变化,主要体现在压气机流量、压气机效率、涡轮流量、涡轮效率等气路部件流量与效率的改变。最后,性能参数的变化使得航空发动机的工作状况改变,从而导致监测参数(转速、温度、压力等)发生偏差。因此,可以通过气路部件流量和效率的变化构建健康因子对发动机的健康状态进行评估。依据上述发动机健康退化机理,考虑由风扇、低压压气机、高压压气机、低压涡轮、高压涡轮的流量与效率的退化量构建航空发动机健康因子。将退化后的发动机与全新发动机相比的各气路部件的流量以及效率的变化值作为退化极限值,接着以发动机当前退化量与退化极限值的比值为健康指标,最后取各个气路部件中健康指标的最小值作为发动机当前的健康状态的评估值,即健康因子。健康因

子的计算过程如式(6)~(8)所示。

$$HI = \min\{hi_F, hi_E\}, hi_F \neq 0, hi_E \neq 0 \quad (6)$$

$$hi_F = 1 - \frac{F_s - F_n}{F_s - F_{end}} \quad (7)$$

$$hi_E = 1 - \frac{E_s - E_n}{E_s - E_{end}} \quad (8)$$

式中: HI 为健康因子, hi_F 为航空发动机气路子系统的流量健康指标, hi_E 为航空发动机气路子系统的效率健康指标, F_n 为当前航空发动机气路子系统的流量, F_s 为航空发动机初始状态下的气路子系统的流量, F_{end} 为航空发动机失效状态下的气路子系统的流量, E_n 为当前航空发动机气路子系统的效率, E_s 为航空发动机初始状态下的气路子系统的效率, E_{end} 为航空发动机失效状态下的气路子系统的效率。

2) 回归树的构建和重要传感器的选择

选择多个监测传感器评估发动机的健康状态会显著提高预测效果,但是用于健康评估的监测传感器并非越多越好。太多的监测传感器会占用发动机的内部空间,增加发动机的重量,带来不必要的维护成本,并且过多的监测传感器会使得评估算法的计算复杂度上升,影响计算效率与预测精度。因此,在评估发动机的健康状态时需要对监测传感器进行合理地选择。关于传感器的选择方法,现在

还缺乏通用的技术,工程上通常是通过观察单个传感器的变化趋势,选择总体呈单调变化趋势并且变化幅度较大的传感器,或者通过经验与不断地尝试来选择合适的传感器评估发动机的健康状态。

本文考虑到航空发动机的健康状态与多监测传感器数据之间呈现复杂的非线性关系,并且缺乏确定的数学模型描述两者的关系。因此,运用非参数化建模的方法,即基于回归树提取与发动机健康状态关系密切的传感器。回归树作为一种非参数建模方法,其最大的优点在于不需要了解监测参数与发动机健康状态背后的数学模型,而是通过使用监测参数对发动机的健康状态进行离散化地分类运算,就可以建立完整的树模型。它最大的特点在于对特征参数的筛选,通过构建好的回归树可以得到特征参数的重要度信息。回归树越顶端的特征参数,其重要度越大,越底部的特征参数,其重要度越小。

以巡航阶段传感器数据建立回归树,回归树特征提取算法分为 3 步,第 1 步为递归二元分割,第 2 步为代价复杂度减枝,第 3 步为重要特征提取。递归二元分割是用特征和分割点将资料分为两个子资料 R_1 和 R_2 ,通过计算真实值与估计值之间的最小误差找到最优的特征 j 和最优的分割点 s ,然后依次往下分割,如式(9)~(11)所示。

$$R_1(j, s) = \{X \mid X_j < s\} \quad (9)$$

$$R_2(j, s) = \{X \mid X_j \geq s\} \quad (10)$$

$$\sum_{X_i \in R_1(j, s)} (HI_i - \hat{HI}_{R_1})^2 + \sum_{X_i \in R_2(j, s)} (HI_i - \hat{HI}_{R_2})^2 \quad (11)$$

式中: R_1 和 R_2 为通过传感器 j 和数值分割点 s 分割成的两个子资料集, X 代表子资料集中的元素, X_j 代表第 j 个传感器的数值, HI_i 代表健康因子真实值, \hat{HI}_{R_1} 代表子资料集 R_1 中健康因子的预估值, \hat{HI}_{R_2} 代表子资料集 R_2 中健康因子的预估值。

当回归树过于复杂并且预测精度不高时,需要对回归树进行修剪。通过给定惩罚因子 α 人为地增大误差进行复杂度的修剪。

$$\sum_{m=1}^{|T|} \sum_{x_i \in R_m} (HI_i - \hat{HI}_{R_m})^2 + \alpha |T| \quad (12)$$

式中: \hat{HI}_{R_m} 代表子资料集 R_m 中健康因子的预估值, T 为终端节点个数。

构建回归树模型后,利用贝叶斯优化算法对回归树的最小叶子数、最大分裂次数进行超参数优化,其步骤为:

步骤 1) 设定待优化参数范围,并以五折交叉验证的均方误差(mean squared error, MSE)作为目标函数;

步骤 2) 利用贝叶斯算法对回归树进行优化;

步骤 3) 返回优化的最小 MSE 及对应回归树模型的超参数;

步骤 4) 将贝叶斯优化模型的超参数作为最终回归树

模型的参数。

最后可以根据回归树模型得到基于重要度得分的特征重要度信息,并由此提取出对航空发动机健康状态重要的传感器。

1.4 基于 Informer 的寿命预测

相比于 Transformer, Informer 的优势在于使用概率稀疏自注意力降低了时间的复杂度,利用自注意力蒸馏机制降低了计算复杂度与内存使用量,并且使用生成式解码器可以一步获得预测结果,能够高效地完成回归类预测任务。基于 Informer 的剩余寿命预测模型主要包含编码器、解码器与参数优化三部分。将特征优选选出的传感器数据作为输入,取选定长度的时间序列输入到编码器中。编码器中使用概率稀疏自注意力方法,计算方法如式(13)所示。并利用自注意力蒸馏机制将输入序列的长度不断减少,蒸馏机制算法如式(14)所示。获得特征图后输入解码器,解码器中将遮盖的多头自注意力应用于概率稀疏化自注意力机制,避免自回归。通过全连接层一步生成获得所有预测结果,一步生成式计算方法如式(15)所示。

$$A(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (13)$$

式中: \bar{Q} 是稀疏矩阵, K 是键向量矩阵, d 是输入维度, V 是值向量矩阵。

$$X'_{(j+1)} = \text{MaxPool}(ELU(\text{Convld}([X'_j]_{\Lambda B}))) \quad (14)$$

式中: $X'_{(j+1)}$ 表示注意力块,它包含多头概率稀疏自注意和基本操作, Convld 为时间序列的一维卷积算法, ELU 为激活函数。

$$X'_{(feedde)} = \text{Concat}(X'_{token}, X'_0) \in R^{(L_{token}+L_y) \times d_{model}} \quad (15)$$

式中: X'_0 属于 $R^{L_y \times d_{model}}$ 表示目标序列的占位符, X'_{token} 属于 $R^{L_{token} \times d_{model}}$ 表示开始字符。

在参数优化部分,计算所有预测结果的均方根误差,若预测精度不符合要求,则返回调整编码器输入时间序列长度以及解码器输入时间序列长度(包含先验序列长度和预测序列长度),同时调整其他超参数,最后获得最佳的预测模型的参数。

2 实验分析

2.1 真实飞行条件下退化数据集

本文采用美国国家航空航天局 2021 年 6 月发布的基于真实飞行条件下航空发动机运行到失效的时间序列数据(N-CMAPSS 数据库)建立预测模型^[28]。与 CMAPSS 数据库不同的是, N-CMAPSS 数据库考虑了真实的飞行条件下的航空发动机健康退化数据,并且将航空发动机退化过程的开端与实际运营的历史结合,保持了其退化过程高精度性,一共包含 0.904 亿条数据。相比于其他数据,该数据库样本数据更真实,数据量更大,具有更多的退化失效模式,有利于开发基于大数据的预测模型,并且开发的算

法在实际应用中具有可用性与转移性。

为了避免过大的计算量,选取其中的 DS01 数据集,使用其 760 万条数据进行算法开发与验证。DS01 包含 10 台发动机在爬升、巡航、下降 3 个阶段的气路传感器数据。运用 1~6 号发动机退化数据作为训练数据,7~10 号发动机退化数据作为测试数据。数据集中包含的 14 个传感器信息如表 1 所示,包含的 4 个真实飞行环境参数变量如表 2 所示,航空发动机健康参数如表 3 所示。

表 1 传感器数据

Table 1 Sensor data

序号	参数	参数含义	单位
1	T24	低压压气机出口总温	°R
2	T30	高压压气机出口总温	°R
3	T48	高压涡轮出口总温	°R
4	T50	低压涡轮出口总温	°R
5	P15	旁路管道总压	psia
6	P2	风扇入口总压	psia
7	P21	风扇出口总压	psia
8	P24	低压压气机出口总压	psia
9	Ps30	高压压气机出口静压	psia
10	P40	燃烧室出口总压	psia
11	P50	低压涡轮出口总压	psia
12	Nf	风扇物理转速	rpm
13	Nc	核心机物理转速	rpm
14	Wf	燃油流量	pps

表 2 真实飞行环境参数

Table 2 Real flight conditions

参数	参数含义	单位
alt	飞行高度	ft
Mach	飞行马赫数	—
TRA	油门解算器角度	%
T2	风扇入口总温	°R

表 3 健康参数

Table 3 Health parameters

序号	参数	参数含义
1	Fan eff	风扇效率
2	Fan flow	风扇流量
3	LPC eff	低压压气机效率
4	LPC flow	低压压气机流量
5	HPC eff	高压压气机效率
6	HPC flow	高压压气机流量
7	HPT eff	高压涡轮效率
8	HPT flow	高压涡轮流量
9	LPT eff	低压涡轮效率
10	LPT flow	低压涡轮流量

2.2 特征优选

首先采用平滑化方法对飞行条件参数进行降噪处理,采用稳健局部加权回归算法去除异常值,窗宽设置为 10。接着计算所有飞行环境参数的 $Deta$ 值。图 2 展示了利用第 1 台发动机第 1 个飞行循环中 4 498 条数据计算的 $Deta$ 值。图中曲线的首尾段分别为飞机爬升与下降阶段,中间段为飞机的巡航阶段。不难发现在爬升与下降阶段 $Deta$ 值变化幅度大,巡航阶段 $Deta$ 值变化平稳。并且经统计分析,飞机在每次巡航飞行时, $Deta$ 值在以 0 为中心 ± 1 的范围内波动,因此不妨取 a 值为 1 作为巡航高度的波动范围。 $Deta$ 大于 1 为爬升工况下监测传感器数据, $Deta$ 小于 -1 时为下降工况下的监测传感器数据,中间为巡航工况下的监测数据。以取定的 a 划分工况,最终得到所有飞行循环中爬升、巡航、下降工况的数据。

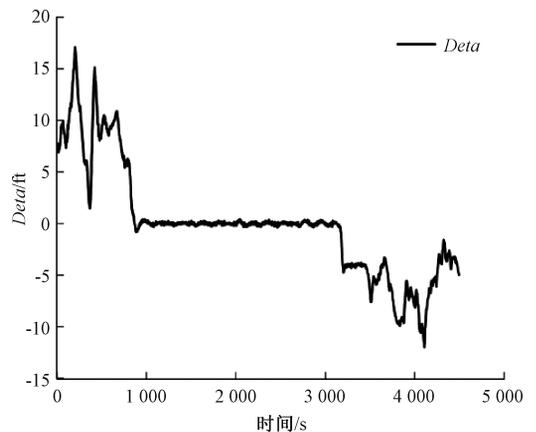


图 2 一个典型飞行循环中的 $Deta$ 值

Fig. 2 $Deta$ value in a typical flight cycle

由于民用飞机绝大部分工作状态处于巡航状态,并且在飞机巡航状态时,发动机处于稳定工作状态,更容易由运行数据预测发动机的剩余寿命,故选择巡航阶段的传感器数据建立回归树模型。

将平滑之后的 14 个传感器参数作为特征参数以预测健康因子建立回归树。采用贝叶斯优化算法对回归树模型的超参数最小叶子数和最大分裂次数进行超参数优化。控制迭代次数为 30 次,最后得到的最小叶子节点数为 3,最大分割次数为 1.3246×10^{10} ,求得回归树的最优特征参数为 T50,最优分割点 1 102.33,五折均方误差值为 0.001 250 3,拟合优度为 0.981 29,可知回归树没有出现拟合现象,故不进行修剪处理,惩罚因子设为 0。回归树最优参数如表 4 所示,回归树的 5 折交叉验证结果如表 5 所示。

表 4 回归树参数及优化值

Table 4 Regression tree parameters and optimized values

参数	参数范围	优化值
最小叶子数	(1,10)	3
最大分割次数	不限制	1.3246×10^{10}

表 5 回归树的 5 折交叉验证结果

Table 5 5-fold cross validation results for regression trees

折数	MSE	拟合优度
折 1	0.001 3	0.981 2
折 2	0.001 3	0.981 3
折 3	0.001 3	0.981 3
折 4	0.001 3	0.981 3
折 5	0.001 3	0.981 2
平均	0.001 3	0.981 3

回归树模型预测的 1~3 号发动机的健康因子的误差范围如图 3 所示,通过预测的误差范围来判断回归树模型的可靠性。图 3 的横坐标表示健康因子的预测值与真实值之间的误差值,纵坐标表示误差值出现的次数。由图 3 可知,对所有发动机健康因子的预测误差都集中在 0 附近,且整体分布呈瘦高型的正态分布,误差分布范围小。由此可知,建立的回归树模型具有高的可靠性,对任意发动机的预测都可以达到较高的预测精度。

由可靠的回归树模型计算得到传感器的重要度大小如图 4 所示。可知对于发动机健康状态最重要的传感器

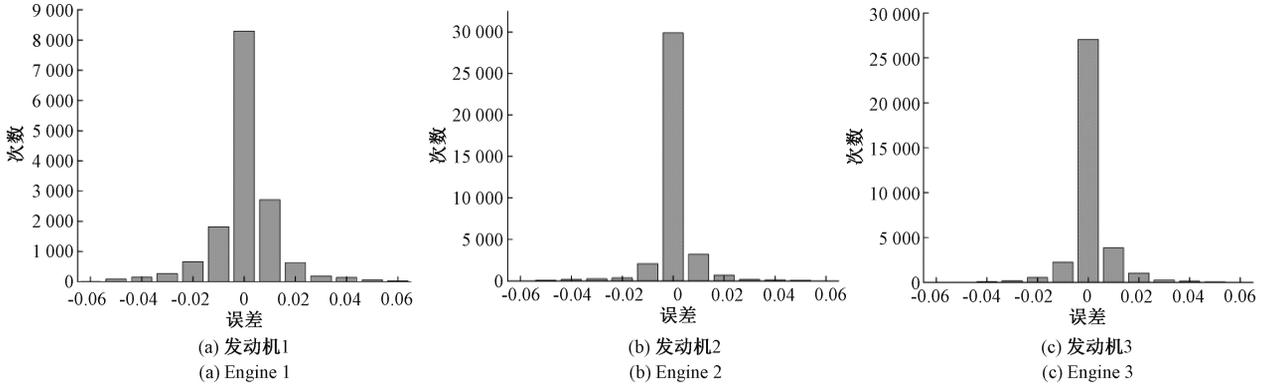


图 3 1~3 号发动机模型预测的健康因子误差范围
Fig. 3 Health index prediction error ranges of 1~3 engines

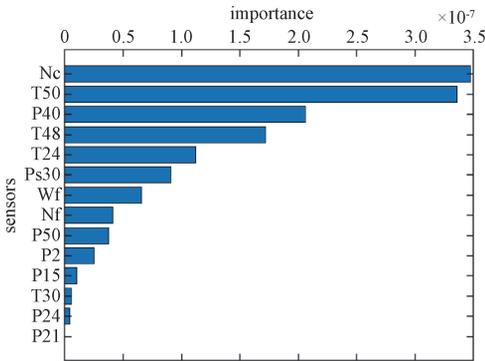


图 4 传感器重要度大小图
Fig. 4 The importance of sensors

按重要度由大到小依次为 Nc、T50、P40、T48 和 T24,其中 T50 为低压涡轮出口总温度,而当前航空公司对发动机健康状态进行监测的指标排气温度(exhaust gas temperature, EGT)也正是指低压涡轮出口总温度,这也反映了回归树计算的传感器重要度的准确性。不妨取重要度大于 10^{-7} 的传感器,即 Nc、T50、P40、T48 和 T24 进行发动机的剩余寿命预测。

2.3 预测结果测试与分析

将 DS01 数据中训练集的 90% 数据进行模型的训练,训练集中的 10% 数据作为验证集,最后用测试集进行模型测试,预测测试集中每台发动机每个飞行循环下的剩余寿

命。由于 Informer 编码器中输入序列的长度、解码器中先验序列长度与预测序列长度对整个预测结果会产生较大的影响,故进行了五组对比实验,实验结果如表 6 所示。其中 seq_len 表示输入序列的长度, label_len 表示先验序列长度, pred_len 表示预测序列长度。最后选择预测精度最高的第 4 组 3 个设置参数,即 seq_len=12, label_len=6, pred_len=3 作为模型的输入参数。除此之外,整个 Informer 模型训练的主要设置参数如表 7 所示。

表 6 参数选择

Table 6 Parameters selection

序号	seq_len	label_len	pred_len	RMSE
1	96	48	24	22.06
2	48	24	12	18.49
3	24	12	6	10.18
4	12	6	3	9.34
5	4	2	1	10.54

最后测试集得到的发动机剩余寿命的预测结果如图 5 所示。图 5 横坐标表示发动机的运行时间,坐标单位为飞行循环数,竖坐标为剩余寿命,单位为飞行循环数。图中呈三角形部分为发动机真实的剩余寿命,其左侧为该发动机剩余寿命的预测值。易知在发动机整个运行周期中,预测算法的预测结果稳定性强且贴近真实值。由于在

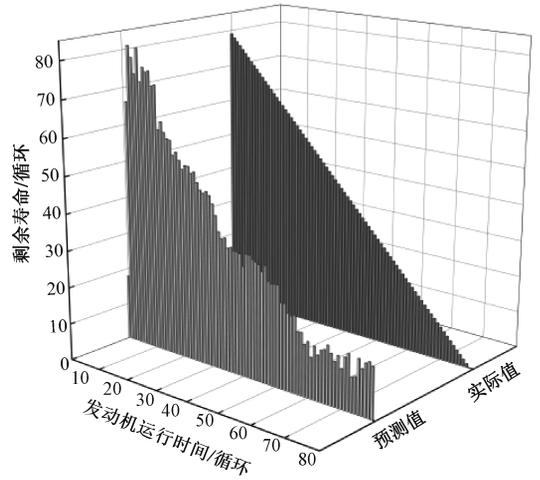
表 7 Informer 模型训练参数

Table 7 Informer model training parameters

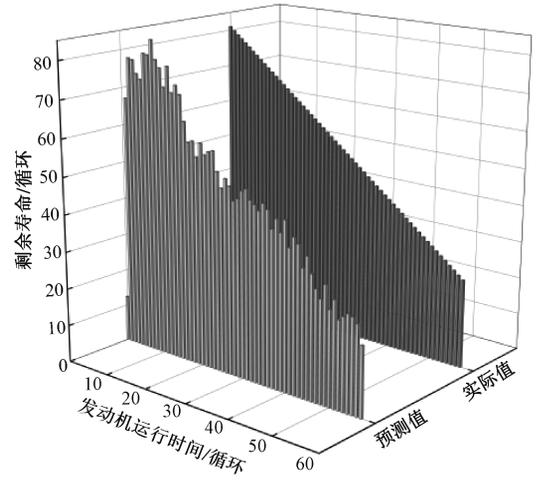
训练参数	参数设置
模型尺寸	512
头的数量	8
译码器的层数	2
解码器层数	1
优化学习率	0.000 1

预测过程中使用了 Drop Last 策略,会丢掉最后不能组成足够学习样本批量的样本。故第 10 台发动机末端因不能组成一个学习批次大小的样本而没有进行预测。

进一步地,为了验证本文提出传感器选择方法的预测效果,选择研究文献[26]里提出的依据传感器变化差异选择方法和经典的依据传感器退化趋势性选择方法进行对比实验。经过计算可知,依据传感器变化差异选择方法计算得出最重要的前 5 个传感器为第 14、9、10、8、12 号传感



(c) 9号发动机
(c) Engine 9



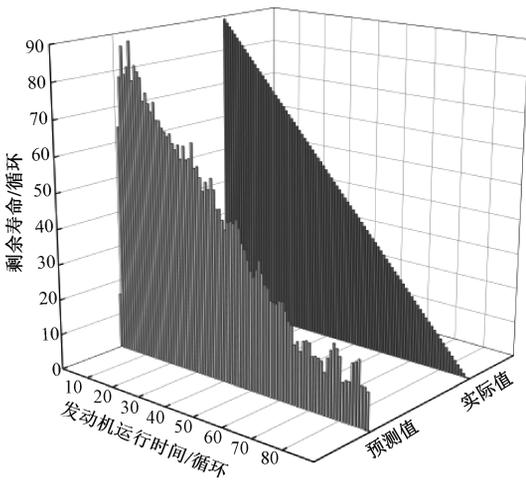
(d) 10号发动机
(d) Engine 10

图 5 测试集发动机剩余寿命预测结果
Fig. 5 RUL prediction results of test engines

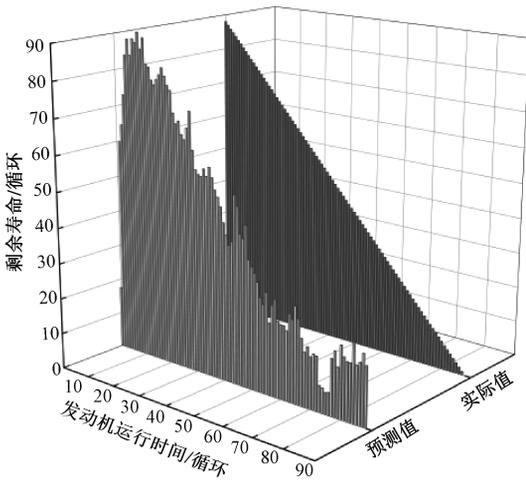
器,经典的依据传感器退化趋势性选择方法计算得出最重要的前 5 个传感器为第 7、5、11、6、14 号传感器,而本文方法所选出最重要的前五个传感器为第 13、4、3、10、1 号传感器,将 3 种方法选出的传感器应用于 Informer 算法进行寿命预测。

预测结果的平均评分函数值如表 8 所示,其中 OR 代表不选择传感器,LR 代表依据传感器退化趋势性选择传感器,SD 代表依据传感器变化差异选择传感器,FO 代表本文的选择传感器方法。可以看到本文所提方法预测结果的评分相比于不选择传感器下降了 29%,与其他 2 种选择方法相比评分分别下降了 48%、27%,预测结果更好。

预测结果的均方根误差 (root mean square error, RMSE)如表 9 所示,可以看出相对于不选择传感器,本文所提方法预测结果的 RMSE 值下降了 14%,与其他两种选择方法相比 RMSE 分别下降了 10%、8%,本文所提方



(a) 7号发动机
(a) Engine 7



(b) 8号发动机
(b) Engine 8

表 8 特征选择的平均分数改进率

Table 8 Mean score improvement ratio of feature selection

预测方法	特征选择	平均分数	精度提升率/%
Informer	OR	5.10	0
	LR	6.08	-19
	SD	5.01	2
	FO	3.60	29

表 9 不同特征选择方法的多次实验 RMSE 统计结果

Table 9 RMSE statistics for multiple experiments with different feature selection methods

预测方法	特征选择	RMSE	标准差	95% 置信区间	精度提升率/%
Informer	OR	10.68	0.33	[10.44,10.92]	0
	LR	10.28	0.41	[9.99,10.57]	4
	SD	9.99	0.29	[9.78,10.20]	6
	FO	9.21	0.26	[9.02,9.40]	14

法具有更好的预测效果。为进一步验证本文提出的特征选择方法(FO)的有效性,对每种方法进行了 10 次独立实验,基于 Informer 模型计算 RMSE、标准差和 95% 置信区间,并对 FO 方法与 OR、LR 和 SD 方法进行了成对 t 检验。结果表明,FO 方法的 RMSE(9.21)显著低于 OR(10.68, $p < 0.001$)、LR(10.28, $p < 0.001$)和 SD(9.99, $p < 0.001$),所有 p 值均小于 0.05,确认 FO 方法在预测精度上具有显著的统计优越性。

为了验证本文所提寿命预测方法的效果,选择当前先进的预测模型 CatBoost、LightGBM、XGBoost、BiLSTM 和 Transformer 进行对比实验分析。CatBoost、LightGBM、XGBoost 均采用贝叶斯优化获得最优超参数,CatBoost 迭代 25 次,树深度为 4,学习率为 0.515 3,正则化系数为 2.65;XGBoost 迭代 25 次,每棵树的特征子采样比例为 0.805 9,树深度为 5,学习率为 0.078 4,树的数量为 214,随机子样本比例为 0.728 0;LightGBM 迭代 25 次,树深度为 8,叶子节点最小样本数为 27,叶子数为 61,L1 正则化为 0.592 4,L2 正则化为 0.046 5,学习率为 0.233 5。BiLSTM 模型设置 2 层 LSTM(分别含 128 和 64 个单元),学习率为 0.000 5,dropout 率为 0.2。Transformer 模型设置 2 层多头注意力机制,学习率为 0.001,dropout 率为 0.2。

通过 10 次重复实验评估每种模型的 RMSE,结果如表 10 所示。由表 10 可知,FO_Informer 模型预测结果的 RMSE 值为 9.21,与 CatBoost、LightGBM、XGBoost、BiLSTM 和 Transformer 算法相比,精度分别提升了 36%、24%、14%、6%和 5%,是四种预测方法中预测精度最高的模型,由此可见本文所提方法具有优越性。

为进一步验证结果的显著性,开展了成对 t 检验和 ANOVA 分析。成对 t 检验表明,FO_Informer 与 CatBoost

表 10 不同预测模型的多次实验 RMSE 统计结果

Table 10 RMSE statistics for multiple experiments with different prediction models

模型	RMSE	标准差	95% 置信区间	精度提升率/%
CatBoost	14.32	0.009 2	[14.31,14.33]	0
LightGBM	12.65	0.000 1	[12.65,12.65]	12
XGBoost	11.17	0.21	[11.01,11.34]	22
BiLSTM	10.02	0.56	[9.59,10.44]	30
Transformer	9.88	0.72	[9.37,10.39]	31
FO_Informer	9.21	0.26	[9.02,9.40]	36

($t = -62.09, p < 0.001$)、LightGBM($t = -41.8, p < 0.001$)、XGBoost($t = -18.54, p < 0.001$)、BiLSTM($t = -4.15, p < 0.01$)和 Transformer($t = -62.667, p < 0.05$)的 RMSE 差异均显著。ANOVA 分析显示 6 组模型间整体差异显著($F = 240.73, p < 0.001$)。这些统计结果证实,FO_Informer 在预测精度和稳定性上具有显著优势,表明本文方法具有稳定性和鲁棒性。

3 结 论

本研究提出了一种基于特征优选及 Informer 的剩余寿命预测方法,并在 N-CMAPSS 数据集上进行了实验验证分析。实验结果表明,与不选择传感器以及传统的依据传感器退化趋势选择传感器方法、依据传感器变化差异性选择传感方法相比,本文提出的选择传感器方法具有优越性,均方根误差分别下降了 14%、10%、8%,平均评分函数值下降了 29%、48%、27%,本文的特征优选方法选出了与航空发动机健康状态更紧密的传感器。本文所提模型与当前的 CatBoost、LightGBM、XGBoost、BiLSTM 和 Transformer 算法相比,精度分别提升了 36%、24%、14%、6%和 5%,本文的剩余寿命预测方法具有良好预测效果。本研究考虑了民机的真实飞行条件,提出的预测模型更加贴近实际应用,为航空发动机剩余寿命预测的应用提供了有力支持。

参考文献

[1] 郭庆,李印龙,郑天翔.基于非线性 Wiener 过程航空发动机性能退化预测[J].推进技术,2021,42(9):1956-1963.
 GUO Q, LI Y L, ZHENG T X. Performance degradation prediction of aero-engine based on nonlinear wiener process[J]. Journal of Propulsion Technology,2021,42(9):1956-1963.

[2] WANG M N, LI Y X, ZHANG Y X, et al. Spatio-temporal graph convolutional neural network for remaining useful life estimation of aircraft engines[J]. Aerospace Systems, 2021, 4(1):29-36.

- [3] 刘斌,许靖,霍美玲,等.基于多尺度自适应注意力网络的剩余寿命预测[J].航空学报,2023,44(5):165-175.
LIU B, XU J, HUO M L, et al. Remaining useful life prediction based on multi-scale adaptive attention network[J]. Acta Aeronautic et Astronautica Sinica, 2023, 44(5): 165-175.
- [4] WANG B, LEI Y G, YAN T, et al. Recurrent convolutional neural network: A new framework for remaining useful life prediction of machinery [J]. Neurocomputing, 2020, 378:117-129.
- [5] LI H, ZHAO W, ZHANG Y X, et al. Remaining useful life prediction using multi-scale deep convolutional neural network [J]. Applied Soft Computing, 2020, 89:106113.
- [6] LIN L, WU J L, FU S, et al. Channel attention & temporal attention based temporal convolutional network: A dual attention framework for remaining useful life prediction of the aircraft engines [J]. Advanced Engineering Informatics, 2024, 60:102372.
- [7] 兰杰,李志宁,李宁,等.基于注意力机制和残差深度分离卷积的 RUL 预测方法[J].电子测量技术,2023,46(15):149-157.
LAN J, LI ZH N, LI N, et al. Prediction method of remaining useful life based on attention mechanism and residual depthwise separation convolution [J]. Electronic Measurement Technology, 2023, 46(15): 149-157.
- [8] PENG ZH Q, WANG Q B, LIU Z R, et al. Remaining useful life prediction for aircraft engines under high-pressure compressor degradation faults based on FC-AMSLSTM [J]. Aerospace, 2024, 11(4):293.
- [9] XU T T, HAN G J, ZHU H B, et al. Multi-resolution LSTM-based prediction model for remaining useful life of aero-engine[J]. IEEE Transactions on Vehicular Technology, 2024, 73(2): 1931-1941.
- [10] MIAO H H, LI B, SUN CH, et al. Joint learning of degradation assessment and RUL prediction for aeroengines via dual-task deep LSTM networks[J]. IEEE Transactions on Industrial Informatics, 2019, 15(9): 5023-5032.
- [11] AN Q L, TAO Z R, XU X W, et al. A data-driven model for milling tool remaining useful life prediction with convolutional and stacked LSTM network [J]. Measurement, 2020, 154:107461.
- [12] 牟含笑,郑建飞,胡昌华,等.基于 CDBN 与 BiLSTM 的多元退化设备剩余寿命预测[J].航空学报,2022,43(7):308-319.
MU H X, ZHENG J F, HU CH H, et al. Remaining useful life prediction of multivariate degradation equipment based on CDBN and BiLSTM [J]. Acta Aeronautic et Astronautica Sinica, 2022, 43(7): 308-319.
- [13] MA M, MAO ZH. Deep-convolution-based LSTM network for remaining useful life prediction[J]. IEEE Transactions on Industrial Informatics, 2021, 17(3): 1658-1667.
- [14] 袁焯,黄虹,程骋,等.基于特征注意力机制的 GRU-GAN 航空发动机剩余寿命预测[J].中国科学:技术科学,2022,52(1):198-212.
YUAN Y, HUANG H, CHENG CH, et al. Remaining useful life prediction of the aircraft engine based on the GRU-GAN network with a feature attention mechanism [J]. Acta Aeronautic et Astronautica Sinica, 2022, 52(1):198-212.
- [15] 张晟斐,李天梅,胡昌华,等.基于深度卷积生成对抗网络的缺失数据生成方法及其在剩余寿命预测中的应用[J].航空学报,2022,43(8):441-455.
ZHANG SH F, LI T M, HU CH H, et al. Missing data generation method and its application in remaining useful life prediction based on deep convolutional generative adversarial network[J]. Acta Aeronautic et Astronautica Sinica, 2022, 43(8): 441-455.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]. Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017:6000-6010.
- [17] GUO J, LEI SH CH, DU B G, et al. MHT: A multiscale hourglass-transformer for remaining useful life prediction of aircraft engine [J]. Engineering Applications of Artificial Intelligence, 2024, 128: 107519.
- [18] LIU Q, ZHANG ZH Y, GUO P, et al. Enhancing aircraft engine remaining useful life prediction via multiscale deep transfer learning with limited data[J]. Journal of Computational Design and Engineering, 2024, 11(1): 343-355.
- [19] ZHOU ZH H, LONG ZH H, WANG R D, et al. An aircraft engine remaining useful life prediction method based on predictive vector angle minimization and feature fusion gate improved transformer model[J]. Journal of Manufacturing Systems, 2024, 76: 567-584.
- [20] ZHOU H Y, ZHANG SH H, PENG J Q, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(12): 11106-11115.

- [21] HOU D L, ZHAO S, WANG Y. Health assessment of gas turbine performance based on SOM-MQE algorithm[C]. 2021 6th International Conference on Intelligent Informatics and Biomedical Sciences(ICIBMS), 2021:30-33.
- [22] GIORGI M G D G, QUARTA M. Hybrid multigene genetic programming-artificial neural networks approach for dynamic performance prediction of an aeroengine[J]. Aerospace Science and Technology, 2020,103:105902.
- [23] URBANOWICZ R J, MEEKER M, CAVA W L, et al. Relief-based feature selection: introduction and review[J]. Journal of Biomedical Informatics, 2018, 85:189-203.
- [24] HAJARIAN N, SOBHANI F M, SADJADI S J. An improved approach for fault detection by simultaneous overcoming of high-dimensionality, autocorrelation, and time-variability[J]. PLOS One, 2020, 15(12): e0243146.
- [25] KIM M, SONG C, LIU K. A generic health index approach for multisensor degradation modeling and sensor selection [J]. IEEE Transactions on Automation Science and Engineering, 2019, 16(3): 1426-1437.
- [26] 赵洪利,陈天铭. 基于双尺度相似性的发动机寿命预测[J]. 推进技术,2022,43(11):355-362.
- ZHAO H L, CHEN T M. Engine life prediction based on two-scale similarity [J]. Journal of Propulsion Technology,2022,43(11):355-362.
- [27] 姜娜. 涡扇发动机性能退化预测与维修策略研究[D]. 哈尔滨:哈尔滨工业大学,2020.
- JIANG N. Research on performance degradation prediction and maintenance strategy of turbofan engines [D]. Harbin: Harbin Institute of Technology,2020.
- [28] CHAO M A, KULKARNI C, GOEBEL K, et al. Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics[J]. Data, 2021, 6(1):5.

作者简介

卢翔(通信作者),博士,教授,主要研究方向为航空器复合材料结构损伤、防护和航空器智能维修。

E-mail: xlu@cauc.edu.cn

高新越,硕士研究生,主要研究方向为深度学习和航空发动机故障诊断。

E-mail:2022082109@cauc.edu.cn

王杜,硕士研究生,主要研究方向为机器学习和航空发动机剩余寿命预测。

E-mail:1771632045@qq.com

康千灼,硕士研究生,主要研究方向为机器学习和航空发动机剩余寿命预测。

E-mail:2023081003@cauc.edu.cn

何晟,硕士研究生,主要研究方向为航空发动机故障诊断。

E-mail:937111628@qq.com