DOI: 10. 19650/j. cnki. cjsi. J2311763

# ALC-PFL: 基于个性化联邦学习的轴承寿命预测方法\*

陈 曦<sup>1</sup>,王 辉<sup>2</sup>,陆思良<sup>3</sup>,严如强<sup>1,4</sup>

(1. 东南大学仪器科学与工程学院 南京 210096; 2. 上海航天电子技术研究所 上海 201109; 3. 安徽大学电气工程与自动化学院 合肥 230601; 4. 西安交通大学高端装备研究院 西安 710049)

**摘 要:**现有数据驱动下的轴承寿命预测方法往往使用特定工况的数据训练相应工况的模型,忽略了其他工况数据所蕴含的有益 退化特征,导致模型预测精度受限。为了充分挖掘和利用不同工况下的轴承退化特征,本文提出基于个性化联邦学习的轴承寿命 预测方法(ALC-PFL)。在该方法中,不同工况轴承的监测数据被存储于多个客户端,一个中心服务器与多个客户端协同工作,以 模型传输、融合和本地更新的方式,为客户端建立个性化预测模型。提出自适应本地融合算法,将中心服务器聚合的全局模型与 客户端本地模型有效融合,保留有助于客户端初始化模型的退化特征,以提升预测性能。用两个轴承数据集对所提方法进行验 证,结果表明其能为不同工况的轴承搭建高性能寿命预测模型,与本地训练方法相比,该方法所得均方根误差降低了至少13%。 关键词:联邦学习;轴承;剩余寿命预测;卷积神经网络

中图分类号: TH17 TN911.7 文献标识码: A 国家标准学科分类代码: 510. 4030

# ALC-PFL: Bearing remaining useful life prediction method based on personalized federated learning

Chen Xi<sup>1</sup>, Wang Hui<sup>2</sup>, Lu Siliang<sup>3</sup>, Yan Ruqiang<sup>1, 4</sup>

(1. School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China; 2. Shanghai Aerospace Electronic Technology Institute, Shanghai 201109, China; 3. School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China; 4. iHarbour Academy of Frontier Equipment, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Existing data-driven methods for predicting the remaining useful life of bearings often rely on data from a specific operating condition to train the corresponding prediction model. The valuable degradation features present in data from other conditions are disregarded. To effectively capture and utilize degradation features across diverse operating conditions, this article proposes a personalized federated learning-based method for bearing remaining useful life prediction. In this method, monitoring data from bearings under different conditions are distributed among multiple clients, while a central server collaborates with these clients to develop personalized prediction models by model transfer, combination, and local updates. To integrate the global model aggregated by the central server with the local model, an adaptive local combination algorithm is introduced, which preserves useful degradation features that aid in initializing the client's model and enhancing prediction performance. The proposed method is evaluated by using two datasets of bearings. The results show its ability to construct high-performance prediction models for bearings operating under different operating conditions. In comparison to local training method, this method manifests a minimum decrease of 13% in root mean square error.

Keywords: federated learning; bearing; remaining useful life prediction; convolutional neural network

0 引 言

轴承是旋转机械设备的重要零部件之一,用于支撑

收稿日期:2023-08-04 Received Date: 2023-08-04

机体并实现高精度回转。由于其常在高温、高压、高负载 等恶劣环境下工作,易出现磨损及退化现象。对轴承运 行状态进行监测,并开展剩余使用寿命(remaining useful life, RUL)预测研究,有助防"失效"于未然,确保设备整 体的安全运行<sup>[1]</sup>。

<sup>\*</sup>基金项目:国家自然科学基金(52275130)项目资助

随着工业大数据时代的到来,智能故障诊断技术迭 代创新<sup>[2]</sup>,基于数据驱动的 RUL 预测方法也成为研究热 点<sup>[3]</sup>。特别是近年来,算力赋能逐步深入,深度学习 (deep learning, DL)理论空前发展,由深层网络搭建的 RUL 预测模型取得了更令人瞩目的预测效果<sup>[4]</sup>。例如, Zhu 等<sup>[5]</sup> 将时频域信号作为多尺度卷积神经网络 (convolutional neural network, CNN)输入实现了轴承的 RUL 预测: Guo 等<sup>[6]</sup> 提出一种基于循环神经网络 (recurrent neural network, RNN)的健康指标构建方法,有 效提高了模型的 RUL 预测精度。借助堆积的网络层,基 于 DL 的方法能直接从状态监测数据中自动发现并学习 层次化退化特征,具备重要的应用价值与现实意义。然 而,在实际工业中,机械设备在不同工况下工作是很常见 的,但现有方法多使用特定工况下有限的退化样本训练 适用于该工况的 RUL 预测模型,忽略了其他工况数据可 能蕴含的有利退化特征,因此所得模型预测精度受限。 已有基于迁移学习的 RUL 预测研究证明,跨越工况的退 化知识有望受益于不同工况的退化模型<sup>[7-8]</sup>。

作为一种新型学习范式,联邦学习(federated learning, FL)<sup>[9]</sup>能以模型迁移的方式捕捉多方数据的共 享特征[10],为解决上述问题提供了新的思路。与集中式 学习(centralized learning, CL)需要将所有数据集中存储 在同一位置进行模型训练的方法不同,FL 允许分布式参 与者将数据存储于各自客户端。在这种架构下,一个中 心服务器协调多个客户端共同训练,生成一个可泛化至 所有客户端的全局模型。近年来,在故障预测与健康管 理领域,基于 FL 的故障诊断成果卓著<sup>[11-13]</sup>, RUL 预测研 究也逐渐成为热点。例如, Kamei 等<sup>[14]</sup>将 FL 与 CL 下不 同结构的模型的性能进行了比较,总结出可在 FL 下表现 更好的 RUL 预测模型; Guo 等<sup>[15]</sup>提出了一种 FL 下的 RUL 预测框架,协同以"孤岛"形式存在的多个全寿命样 本,为各客户端搭建 RUL 预测模型;Chen 等<sup>[16]</sup>提出一种 在 FL 下使用模型剪枝的轴承 RUL 预测方法, 高效地完 成了为分布式数据建模的任务。

尽管上述方法取得了出色的效果,但全局模型的性能易受各客户端数据分布的影响<sup>[17]</sup>,传统 FL 下的 RUL 预测方案尚且无法为分布式、不同工况数据搭建可泛化 至所有端的全局模型。为了应对这一挑战,个性化联邦 学习(personalized federated learning, PFL)应运而生。在 PFL 中,全局模型参数由所有客户端共享,客户端模型参 数则在全局模型的基础上根据其本地数据进行个性化微 调,由此所得模型受其他端数据分布影响较小,因此可适 应客户端的差异化需求。

鉴于此,本文提出了一种 PFL 下的轴承 RUL 预测方法(adaptive local combination based personalized federated learning, ALC-PFL)。在该方法中,不同工况的状态监测

数据分别存储于不同客户端。一个中心服务器与各个客 户端进行协同,通过模型传输、融合以及本地更新的方 式,为各客户端建立个性化的 RUL 预测模型。提出针对 各客户端模型初始化的自适应本地融合(adaptive local combination, ALC)算法,保留全局模型浅层退化特征,并 根据本地数据自适应地训练更新参数,实现全局模型与 本地模型的有效融合。具体来说,在每轮的训练过程中, 多个客户端先用各自单一工况的数据训练本地模型并将 其上传至中心服务器;中心服务器将接收到的本地模型 进行聚合,得到富含多种工况数据退化特征的全局模型, 并将其返回至各客户端;各客户端利用 ALC 算法得到本 地初始化模型,并在此基础上进行新一轮的更新与上传。 经过多轮训练后,客户端将获得适用于其特定工况数据 的个性化 RUL 预测模型。

# 1 理论背景及问题描述

#### 1.1 FL 方法简述

FL于 2016年由谷歌提出,是一种灵活且高效的机器学习方法。通过去中心化的方式,它有效地打破了数据孤岛,推动了多方协作的实现,为多行业与人工智能的结合及落地带来了新可能。其主要步骤为:首先,各客户端确定一个结构及参数完全相同的初始化模型,并利用各自数据进行模型更新。之后,这些本地模型参数会被上传至中心服务器。接着,中心服务器使用联邦平均(federated averaging, FedAvg)算法,将接收到的多个模型聚合,形成一个全局模型,并将此下发给各客户端。最后,各客户端再利用本地数据对接收到的全局模型进行更新,并将其再次上传至中心服务器。以上过程会重复多次,直到达到预设的最大训练轮次或全局模型实现收敛。

## 1.2 问题描述

在经典 FL 架构中, 假设有 K 个客户端 {1, 2, …, K} 和一个中心服务器, 它们总体的优化任务为:

$$\min_{\boldsymbol{\beta}} \mathcal{L}_{FL} = \sum_{k=1}^{K} \frac{N_k}{N} \mathcal{L}_k(\boldsymbol{\beta}_k; \boldsymbol{R}_k)$$
(1)

式中: 
$$\mathcal{L}_k(\boldsymbol{\beta}_k; \boldsymbol{R}_k) = \frac{1}{N_k} \sum_{i=1}^{N_k} l(\boldsymbol{\beta}_k; \boldsymbol{x}_i^k, \boldsymbol{y}_i), N = \sum_{k=1}^{K} N_k, l(\cdot)$$

为各客户端的损失函数, $\beta_k$ 为本地模型参数, $R_k$ 为本地 数据, $x_i^k$ 为输入数据, $y_i$ 为输出标签,各客户端样本数为  $N_k$ ,N为总样本数, $\beta$ 为全局模型。在每轮的训练中,客 户端使用式(2)所示随机梯度下降(stochastic gradient descent, SGD)算法对预测模型进行本地训练(local training, LT):

$$\boldsymbol{\beta}_{k} \leftarrow \boldsymbol{\beta}_{k} - \delta \nabla l(\boldsymbol{\beta}_{k}, \boldsymbol{B}_{k})$$
(2)

式中: $\delta$ 为模型学习率,"←"代表更新操作, $B_k$ 为批样本 集合。之后,中心服务器使用式(3)所示 FedAvg 算法将 来自所有客户端的模型聚合为全局模型:

$$\boldsymbol{\beta} = \sum_{k=1}^{K} \frac{N_k}{N} \boldsymbol{\beta}_k \tag{3}$$

历经多轮聚合后,**β** 会趋于收敛,并在各客户端取得 良好表现。

但在实际应用中,若各客户端数据分布差异较大, $\beta$ 会难以泛化至所有客户端。例如在本研究场景下,各客 户端拥有不同工况的状态监测数据,找到适宜于所有数 据预测的 $\beta$ 是非常困难的。然而,考虑到 $\beta$ 包含有益于 所有工况数据的退化特征<sup>[18]</sup>,因此本研究将基于分布式 本地数据,在 $\beta$ 辅助下,找到适用于各客户端的个性化模 型,如式(4)所示。

$$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_K\} = \arg\min\left(\sum_{k=1}^K \frac{N_k}{N} \mathcal{L}_k\right)$$
 (4)

其中,  $\mathcal{L}_{k} = \frac{1}{N_{k}} \sum_{i=1}^{N_{k}} l(\boldsymbol{\beta}_{k}; \mathbf{x}_{i}^{k}, y_{i}; \boldsymbol{\beta})_{\circ} \mathbf{x}_{i}^{k}$ 为第 k个客户

端存储的特定工况数据, $\boldsymbol{\beta}_k(k=1, 2, \dots, K)$ 为最终所得个性化 RUL 预测模型。

# 2 基于个性化 FL 的轴承寿命预测方法

#### 2.1 网络结构描述

考虑到非平稳条件下振动信号的复杂特性<sup>[5,19-20]</sup>,本 文使用多尺度核网络结构捕捉轴承运行过程中的浅层及 深层的退化特征。对于每个客户端,所设计的 RUL 预测 模型结构如图 1 所示,主要包含两个多尺度特征提取模 块以及末端回归主路。输入层为一个时间步的一维时频 域特征,先经维度变形后成为通道数为 21 的一维短向 量,再经一个卷积层进行局部退化特征的识别。选用具 有 4 个大小不一卷积核的多尺度特征提取模块 A,在局 部特征基础上进行多尺度浅层退化特征的提取,同理,模 块 B 在浅层特征基础上进行深层退化特征的提取,同理,模 均最化特征。最后,利用最大池化层及卷积层保留重要 特征,为后续全连接层(神经元个数为 160)的回归提供 有利支持。各层所用激活函数均为 Tanh,全连接层 Dropout 率<sup>[21]</sup>为 0.5。

# 2.2 总体框架

所提轴承寿命预测方法 ALC-PFL 的总体框架如图 2 所示,由 K 个客户端和一个中心服务器组成,模型的训练 主要包含准备阶段以及多轮"ALC+LT"阶段:

1)准备阶段

各客户端收集并存储运行在特定工况的轴承振动加



图 1 预测模型的网络结构

Fig. 1 Network architecture of the RUL prediction model

速度信号,并对所得数据执行预处理操作:基于原始振动 信号的峰度及  $3\sigma$  准则<sup>[22]</sup>确定退化开始时间  $t_f$ ,对退化 阶段 信号进行短时傅里叶变换(short-time Fourier transform, STFT)及最小-最大归一化(min-max normalization, MMN)处理,得到可作为模型输入的时频域特征 信号。

以线性递减函数制定归一化输出标签 y<sub>label</sub>,即真实 RUL,其表达式为:

$$y_{label} = \frac{t_w - t_n}{t_w - t_f} \tag{5}$$

式中:t<sub>n</sub>为当前时刻,t<sub>w</sub>为全寿命周期。

另外,将退化阶段分为低风险退化区及高风险退化 区,两区分界点为 t<sub>ent</sub>,其定义如式(6)所示。

$$t_{cut} = 0.8 \times (t_w - t_f) \tag{6}$$

根据此点位置制定超前标签  $y_{adv-label}$ , 如式(7)所示。

$$y_{adv-label} = \begin{cases} \frac{t_{cut} - t_n}{t_{cut} - t_f}, & t_n \in [t_f, t_{cut}) \\ 0, & t_n \in [t_{cut}, t_w] \end{cases}$$
(7)

处于低风险退化区的超前标签值线性递减,而在高风险区,轴承濒临失效,发生生产中断的几率较高,因此 设定此区超前标签值为0。

在进行模型的监督学习训练之前,客户端需要将样本划分为训练集与测试集,并准备好训练集的输入数据 及相应输出标签。具体地说,对于某个客户端 k,将训练



图 2 所提 ALC-PFL 轴承寿命预测框架 Fig. 2 The framework of the proposed ALC-PFL method for bearing RUL prediction

集的输入数据和线性标签组合成训练集合  $D_k$ ,同时,将 训练集的输入数据与超前标签组成另一个训练集合  $D'_k$ 。随后,各客户端使用  $D_k$  对图 1 所示模型进行如 式(2) 所示的监督学习训练,进而得到每个客户端的初 代本地模型 $\beta_k$ 。

2) 多轮"ALC+LT" 过程

准备阶段完成后,中心服务器联合多个客户端为分 布式的不同工况轴承数据搭建个性化 RUL 预测模型,期 间涉及多轮"ALC+LT"过程。以第1轮"ALC+LT"过程 为例,主要包含5个步骤,分别是:

(1)模型上传。多个客户端将各自的本地模型参数 打包为一个消息,再将消息通过网络传输到中心服务器端,实现模型的上传。

(2)模型聚合。中心服务器在接收多个模型参数 后,将各模型对应位置参数使用式(3)所示 FedAvg 算法 进行相加与平均,形成一个全局模型。

(3)模型下发。中心服务器将聚合而来的全局模型 下发给各个客户端。

(4)模型初始化。各客户端在接收全局模型参数 后,使用 ALC 算法基于 **D**<sup>'</sup><sub>4</sub> 将全局模型与上一代本地模 型有效结合,形成新的本地初始化模型。

(5)本地模型训练。各客户端在本地初始化模型基础上,使用 **D**<sub>k</sub> 对模型参数进行更新。

在多轮学习过程中,步骤(1)~(5)依次迭代并循环 进行,直至达到预设最大轮次。其中,除步骤(4)外,其 他步骤与常规 FL 一致。

具体地,在步骤(4)中,对于客户端k,使用 ALC 算法 将接收到的全局模型 $\boldsymbol{\beta}^{i-1}$ 与其上一轮本地模型 $\boldsymbol{\beta}_{k}^{i-1}$ 有效 结合,生成新初始化模型 $\hat{\boldsymbol{\beta}}_{k}^{i}$ ,如式(8)所示。  $\hat{\boldsymbol{\beta}}_{k}^{t} \leftarrow \boldsymbol{\beta}_{k}^{t-1} \circ \boldsymbol{W}_{kl} + \boldsymbol{\beta}^{t-1} \circ \boldsymbol{W}_{km}$ (8)

式中:  $W_{k,l}$ 和  $W_{k,g}$ 为本地模型及全局模型的聚合权重矩阵, "。"表示哈达玛积运算。若 $w_{k,l}^{p}$ 和 $w_{k,g}^{p}$ 代表各模型中第p个 元素的聚合权重,有 $w_{k,l}^{p}$ , $w_{k,g}^{p} > 0$ ,且 $w_{k,g}^{r}$ + $w_{k,l}^{p} = 1$ 。然而, 在本研究中,仅依据模型参数值,很难判断两模型中退化特 征的留与否,即聚合权重值不能同时确定。受文献[23] 启 发,将两个模型的聚合权重合并为局部更新权重 $W_{k}$ ,式(8) 可进一步用式(9) 来表示:

 $\hat{\boldsymbol{\beta}}_{k}^{\iota} \leftarrow \boldsymbol{\beta}_{k}^{\iota-1} + (\boldsymbol{\beta}^{\iota-1} - \boldsymbol{\beta}_{k}^{\iota-1}) \circ \boldsymbol{W}_{k}$ (9)

其中,认为 $\beta^{\iota-1}$ 与 $\beta_k^{\iota-1}$ 的差值为"更新值"。

通常情况下,对于由深层神经网络搭建的 RUL 预测 模型,其浅层网络往往用于提取普适的退化特征,而深层 网络则负责提取更为抽象且具有特定特点的高级退化特 征<sup>[18]</sup>。在本研究中,考虑到全局模型是由多个客户端的 本地模型聚合而成,因此包含了多种工况数据的退化特 征。因此,对于本地初始化模型的浅层来说,可将 $\beta^{(-1)}$ 的 浅层参数全部保留,即此时  $W_k = 1$ 。而对深层来说,为了 捕捉各客户端的深层个性化退化特征,其更新权重  $W_k^n$ 则 通过 SGD,根据本地数据进行自适应地学习调整,如 式(10)所示。

$$\boldsymbol{W}_{k}^{h} \leftarrow \boldsymbol{W}_{k}^{h} - \gamma \, \nabla_{\boldsymbol{w}_{k}^{h}} \mathcal{L}(\hat{\boldsymbol{\beta}}_{k}^{t}, \boldsymbol{D}_{k}^{\prime}, \boldsymbol{\beta}^{t-1}) \tag{10}$$

其中, $\gamma$ 为权重学习率。此时,各客户端使用超前标签所在集合 $D'_{k}$ 训练 $W^{h}_{k}$ 。该操作不仅能让模型富有本地数据的退化特征,还富有失效预警信息,有效改善最终模型的早期预测能力,进而提高预测的准确性和可靠性。之后使用式(11)将学到的 $W^{h}_{k}$ 归一化至[0,1]范围内:

$$\boldsymbol{W}_{k}^{h} = \max(\boldsymbol{0}, \min(\boldsymbol{1}, \boldsymbol{W}_{k}^{h}))$$
(11)

基于上述分析,式(9)可进一步表示为:

 $\hat{\boldsymbol{\beta}}_{k}^{t} \leftarrow \boldsymbol{\beta}_{k}^{t-1} + (\boldsymbol{\beta}^{t-1} - \boldsymbol{\beta}_{k}^{t-1}) \circ [1^{l-h}; \boldsymbol{W}_{k}^{h}]$ (12)

其中,*l* 为预测模型的总层数,*h* 为模型中所定义的深 层结构层数。待步骤(4)完成后,各客户端得到既嵌有全 局浅层退化特征、又与本地数据特征相关的初始化模型。

接下来进入步骤(5),各客户端使用包含线性标签的集合 $D_k$ ,在初始化模型基础上进行如式(13)所示的本地更新:

$$\boldsymbol{\beta}_{k}^{t} \leftarrow \boldsymbol{\hat{\beta}}_{k}^{t} - \delta \nabla_{\boldsymbol{\hat{\beta}}_{k}^{t}} \mathcal{L}(\boldsymbol{\hat{\beta}}_{k}^{t}, \boldsymbol{D}_{k}, \boldsymbol{\beta}^{t-1})$$
(13)

其中,  $\boldsymbol{\beta}_{k}^{t}$  为第 k 个客户端在 t 时刻训练好的个性化 RUL 预测模型。在经历预设轮次的"ALC+LT"后,每个 客户端都将获得一个具备高精度 RUL 预测能力的个性 化模型。

# 3 试验验证

#### 3.1 试验数据

使用 XJTU-SY<sup>[24]</sup>及 PHM-2012<sup>[25]</sup>两个数据集的部分轴承退化数据对所提方法进行验证,所用的轴承参数如表 1 所示。XJTU-SY 数据集来自机械装备监测联合实验室的轴承加速寿命实验台,由水平及竖直方向

上的 PCB 352C33 型加速度传感器采集得到,采样频率 为 25 600 Hz,采样间隔为 60 s,每次采样时长为 1.28 s,共记录 3 个工况下 15 个全寿命周期的样本。 PHM-2012 数据集来自法国 FEMTO-ST 研究所的 PRONOSTIA 实验台,由 DYTRAN 3035B 型加速度传感 器采集得到,采样频率为 25 000 Hz,采样间隔为 10 s, 采样时长为 0.1 s,共包含 3 个工况下 17 个全寿命周期 的样本。在本研究中,选用 XJTU-SY 数据集中 3 个工 况及 PHM-2012 数据集中 1 个工况下的样本进行试验 验证,相关工况及样本信息如表 2 所示。

表1 两个数据集中的轴承参数

#### Table 1 Characteristics of tested bearings in two datasets

参数名称	XJTU-SY 轴承数值	PHM-2012 轴承数值
内圈滚道直径/mm	29.30	20.00
外圈滚道直径/mm	39.80	32.00
基本额定动载荷/N	12 820	4 000
基本额定静载荷/N	6 650	2 470
滚珠直径/mm	7.92	3. 50
滚珠个数	8	13

表 2 两个数据集中 4 个工况下样本的详细信息 Table 2 Detailed information of two datasets

数据集		工况	栏木「 立际 寿命 / t ( x60 ₅) ]	
	负载/N 转速/(r·s <sup>-</sup>			
	12 000	35.0	XB1_1[123/76], XB1_2[161/44], XB1_3[158/60], XB1_4 [122/120], XB1_5 [52/40]	
XJTU-SY	11 000	37.5	XB2_1 [491/454], XB2_2 [161/47], XB2_3 [533/330], XB2_4 [42/31], XB2_5 [339/141]	
	10 000	40.0	XB3_1 [2538/2361], XB3_2 [2496/0], XB3_3 [371/339], XB3_4 [1515/1418], XB3_5 [114/6]	
РНМ-2012	4 000	30.0	PB1_1 [467/218], PB1_2 [145/137], PB1_3 [395/288], PB1_4 [238/180], PB1_5 [410/401], PB1_6 [408/401], PB1_7 [376/368]	

基于上述样本,将不同工况的样本划分至不同客户 端,即4个客户端分别存储4个工况下的样本,每个客户 端样本被细分为训练样本及测试样本,以模拟FL下多客 户端协同搭建个性化寿命预测模型的场景。各客户端样 本划分如表3所示,其中客户端1、2以及3的样本来自 XJTU-SY数据集,客户端4的样本来自PHM-2012数据 集。在本研究的准备阶段,使用大小为0.02 s,步长为 0.01 s 的汉明窗<sup>[26]</sup>对各样本的振动加速度信号进行时 频域变换处理,得到模型的输入特征。

# 3.2 评价指标

选用以下 3 个指标来对所得模型的性能进行定量评估,分别是均方根误差(root mean square error, RMSE)、 平均绝对误差(mean absolute error, MAE)、得分函数

表 3 各客户端样本划分情况

Table 3	Division of	f training :	set and	testing	set in	each	client
---------	-------------	--------------	---------	---------	--------	------	--------

客户端	训练样本	测试样本
1	XB1_1, XB1_2, XB1_3, XB1_4	XB1_5
2	XB2_1, XB2_2, XB2_3, XB2_4	XB2_5
3	XB3_1, XB3_2, XB3_3, XB3_5	XB3_4
4	PB1_1, PB1_2, PB1_3, PB1_4, PB1_5, PB1_6	PB1_7

(score function, SF),其计算公式分别为:

$$RMSE = \sqrt{\frac{1}{q} \sum_{i=1}^{q} (\hat{y}_i - y_i)^2}$$
(14)

$$MAE = \frac{1}{q} \sum_{i=1}^{q} |\hat{y}_i - y_i|$$
(15)  
$$y_i E = \frac{y_i - \hat{y}_i}{q} \times 100$$

$$A_{i} = \begin{cases} \exp(-\ln(0.5) \cdot (E_{i}/5)), & E_{i} \leq 0\\ \exp(\ln(0.5) \cdot (E_{i}/20)), & E_{i} > 0 \end{cases}$$
(16)  
$$SF = \frac{1}{q} \sum_{i=1}^{q} (A_{i})$$

其中, $y_i$  为实际寿命值, $\hat{y}_i$  为预测寿命值,q 为各个 样本所能划分的输入特征总个数。对于 RMSE 和 MAE, 它们都可直观表现出预测值与实际值的偏离程度,偏离 越大,两种指标数值越大。然而,它们无法区分预测值是 高于还是低于真实值,即对高估及低估不敏感。而在实 际工程中,提前预知轴承失效对于确保旋转机械的安全 运行至关重要,特别是在濒临失效的高风险区内。因此, 使用 SF 来对高估情况实施更大的惩罚,使得分更少。综 上,模型的 RMSE 及 MAE 越低, SF 越高,模型的表现 越好。

#### 3.3 所提方法与常用方法的对比分析

在本节中,将本文所提 ALC-PFL 方法与以下常用方法进行对比。

Central:集中式训练方法,将所有客户端的非同工况数据聚合起来,之后进行模型训练,最后在所得模型上对不同工况样本进行测试。

Local:本地训练方法,各客户端使用自有训练样本进行模型训练,此过程不涉及知识传输,所得模型个数与客户端个数相同,最后在本地模型上进行测试。

FedAvg:在每轮学习中,各客户端会迭代训练本地模型,并将其参数上传至中心服务器。后者采用加权平均的方式聚合来自多个客户端的模型,以生成全局模型。最后,在经过多轮学习的全局模型上进行测试。

各方法所用模型结构如图 1 所示,均采用均方误 差为损失函数。所提 ALC-PFL 方法的试验参数如 表 4 所示。Central 方法总训练次数为 250。Local 方 法各客户端的本地训练次数为 250。FedAvg 方法训 练轮数及每轮本地训练次数与所提方法整体一致。 对比方法均用线性标签所在集合 **D**<sub>k</sub> 进行监督学习下 的模型训练。

表 4 试验参数 Table 4 Experiment parameters

参数	选用值	参数	选用值
客户端个数	4	模型学习率 $\delta$	0.001
训练轮数	5	每轮本地训练次数	50
批样本数量	128	每轮权重训练次数	1
权 重 受 习 索 🗸	1		

图 3 (a)~(c)分别展示了各方法在测试样本上的 3个评价指标,均为10次有效试验的平均值。总体来 看,ALC-PFL 方法的 RMSE 和 MAE 值均小于对比方法, 而 SF 值均大于对比方法,表现最优。将其与 Local 方法 进行比较,4个测试样本的 RMSE 指标分别降低了 20.91%、13.55%、14.72%、13.17%, MAE 指标分别降低 了 18.85%、16.82%、16.45%、11.12%,SF 指标分别提升 了 9.91%、11.90%、6.32%、5.56%。这表明对于各客户 端而言,使用 ALC 算法有效保留模型底层参数,使各客 户端初始模型富含更多浅层退化特征将有助于各预测模 型性能的提升。然而,由于 Local 方法可用数据有限,模 型可提退化特征较少,最终表现就相对逊色。另外, Central 和 FedAvg 方法表现较差,这是由于各客户端样本 工况不同,退化模式多变,直接通过数据训练或参数聚合 得到的单一或全局模型难以胜任不同工况下的寿命预测 任务。通过与 FedAvg 方法效果对比,可看出各客户端使 用 ALC 算法进行模型初始化的必要性,也进一步强调了 为各客户端建立个性化预测模型的必要性。图4(a)~(d) 则分别展示了各客户端测试样本退化阶段 RUL 预测曲 线,相比于常用方法,ALC-PFL方法能更好地拟合真实退 化值,提供更精准的预测结果。



图 3 不同方法在测试样本上的评价指标值 Fig. 3 Values of evaluation criteria of different methods on testing samples





#### stage of each client testing sample

## 3.4 多客户端建模结果及分析

在实际轴承 RUL 预测任务中,不排除为多组非同工 况数据建立模型的情况。为了探究客户端个数对模型性 能的影响,本小节基于表 3 中的客户端,考虑 2、3 和 4 个 客户端,即使用(1, 2)、(1, 2, 3)以及(1, 2, 3, 4)客户 端进行建模的场景。

表5列举了以上3种场景的建模效果。在2个客户 端的协同建模任务中,客户端3和4训练样本全程不参 与模型训练,因此不对测试样本 XB3 4 及 PB1 7 进行测 试:在3个客户端(1.2.3)的协同建模任务中,客户端4 训练样本全程不参与模型训练,因此不对测试样本 PB1 7 进行测试。从 XB1 5、XB2 5 和 XB3 4 的测试结 果来看,随着所用客户端个数的增加,所得模型的 RUL 预测性能均有小幅提升。这是因为在该场景下,使用更 多的客户端可以为全局模型提供更多的退化信息,使得 每轮学习中由 ALC 算法得到的初始化模型的退化特征 更丰富,从而提升了模型的预测性能。另外,与单一客户 端建模相比(即 Local 方法),多客户端建模效果更好,以 XB1 5 为例,分别使用 2、3、4 个客户端进行协同建模,所 得模型的 RMSE 值分别降低了 18.42%、20.40%、 20.91%, MAE 值降低了 16.70%、18.11%、18.85%, SF 值 则提升了 5.52%、6.67%、9.91%。同样,在 XB2\_5 上,所 得模型的 RMSE 值分别降低了 5.73%、8.20%、13.55%, MAE 值降低了 10.11%、10.94%、16.82%, SF 值则提升 了 11.12%、11.76%、11.90%。综上,该方法能够灵活地 为分布于多个客户端的不同工况数据建立个性化 RUL 预测模型,并取得良好的预测效果。

#### 表 5 多客户端下的建模效果

 
 Table 5
 Prediction performance of the ALC-PFL method in different number of clients

样本	评价指标	2个客户端	3个客户端	4个客户端
	均方根误差	0. 143 9	0. 140 4	0. 139 5
XB1_5	平均绝对误差	0.1122	0.1103	0.1093
	得分函数	0.4108	0.415 3	0.427 9
	均方根误差	0.095 3	0.092 8	0.0874
XB2_5	平均绝对误差	0.0764	0.0757	0.0707
	得分函数	0.4014	0.4037	0.404 2
	均方根误差	-	0.095 2	0.093 8
XB3_4	平均绝对误差	-	0.076 5	0.072 6
	得分函数	-	0.469 0	0.4791
PB1_7	均方根误差	-	-	0.094 9
	平均绝对误差	-	-	0.077 5
	得分函数	-	-	0.441 8

#### 3.5 超前标签的影响

在轴承濒临失效的高风险退化阶段,提前预知其 RUL 至为关键。为进一步验证"ALC-PFL"环节使用超 前标签对模型进行初始化,进而改善其后期预警表现的 作用,本节将模型初始化所用超前标签替换为线性标签, 并在所提框架下进行模型训练。

在4个客户端下,使用两种标签所得模型的预测效 果如表6所示。总体来看,使用两种标签所得到的模型 在测试样本上的 RMSE 和 MAE 值非常接近,表明两种方 法下的模型预测能力相当,都能够准确预测不同工况下 轴承的剩余寿命。然而,使用超前标签模型的 SF 值略高 于使用线性标签的模型,尤其在测试样本 XB3\_4 和 PB1 7上表现更为明显。

表 6 不同标签进行模型初始化的影响 Table 6 Influence of different labels for model initialization

测试样本	评价指标	超前标签	线性标签	
	均方根误差	0. 139 5	0. 138 9	
XB1_5	平均绝对误差	0.1093	0.1103	
	得分函数	0.427 9	0.427 1	
	均方根误差	0.087 4	0.0879	
XB2_5	平均绝对误差	0.0707	0.0714	
	得分函数	0.404 2	0.3997	
XB3_4	均方根误差	0.093 8	0.094 1	
	平均绝对误差	0.072 6	0.072 0	
	得分函数	0.4791	0.454 3	
PB1_7	均方根误差	0.094 9	0.094 3	
	平均绝对误差	0.077 5	0.077 2	
	得分函数	0.441 8	0. 421 1	

图 5 (a)~(d)分别展示了所得模型在各客户端测 试样本退化区间后 20%时间段内的预测曲线。每个子图 下方的柱状图表示了真实值与预测值之间的差值。其 中,线性误差为真实值与使用线性标签所得模型的预测 值之差,而超前误差为真实值与使用超前标签所得模型 的预测值之差。在轴承进入退化后期后,两种方法的预 测结果都能逼近真实值。然而,使用超前标签所得模型 预警效果更优,其误差为正值概率更高,即"欠预测"情 况更多。同时,预测值不会过于小于真实值。对于各客 户端测试样本 XB1\_5、XB2\_5、XB3\_4 与 PB1\_7,其超前误 差和线性误差在各自退化后期为正值的概率分别为 (52.96% & 50.00%)、(73.52% & 35.29%)、(76.47% & 32.35%)与(81.81% & 18.18%)。这进一步证明了使 用超前标签进行模型初始化的有效性,所得模型能为真 实场景下不同工况的轴承提供可靠 RUL 预测。



in the later stage of degradation

# 4 结 论

为了充分挖掘和有效利用不同工况下的轴承退化特征,本文提出一种基于 PFL 的轴承 RUL 预测方法 ALC-PFL,联合分布于多个客户端的不同工况的轴承状态监 测数据,为其建立个性化预测模型,实现不同工况下轴承 RUL 的高精度预测。主要结论如下:

1)该方法除了使用特定工况数据进行该工况下预测 模型的训练外,还利用了 FL 下其他工况数据所包含的全 局浅层退化特征,进而提升个性化模型预测性能。

2)各客户端在初始化模型的学习过程中,选用超前标签训练模型深层的更新参数,为初始化模型注入预警信息,进而改善模型在轴承退化后期阶段的预警能力。

3)多个工况下轴承的试验证明了所提方法的可行性 和有效性,相比于常用方法,所提方法预测精度更高,在 轴承退化阶段后期的预测结果更为合理可靠。

#### 参考文献

- [1] 陈雪峰, 訾艳阳. 智能运维与健康管理[M]. 北京:机械工业出版社,2018.
   CHEN X F, ZI Y Y. Intelligent maintenance and health management[M]. Beijing: China Machine Press, 2018.
- [2] 雷亚国,贾峰,孔德同,等. 大数据下机械智能故障诊断的机遇与挑战[J]. 机械工程学报,2018,54(5):94-104.

LEI Y G, JIA F, KONG D T, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era[J]. Journal of Mechanical Engineering, 2018, 54(5): 94-104.

- [3] 彭宇,刘大同.数据驱动故障预测和健康管理综述[J].仪器仪表学报,2014,35(3):481-495. PENGY,LIUDT. Data-driven prognostics and health management: A review of recent advances[J]. Chinese Journal of Scientific Instrument,2014,35(3):481-495.
- [4] 陈志强,陈旭东,OLIVIRAJVD,等.深度学习在设备故障预测与健康管理中的应用[J]. 仪器仪表学报,2019,40(9):206-226.

CHEN ZH Q, CHEN X D, OLIVIRA J V D, et al. Application of deep learning in equipment prognostics and health management [J]. Chinese Journal of Scientific Instrument, 2019, 40(9): 206-216.

- [5] ZHU J, CHEN N, PENG W. Estimation of bearing remaining useful life based on multiscale convolutional neural network [J]. IEEE Transactions on Industrial Electronics, 2019: 66(4): 3208-3216.
- [6] GUO L, LI N P, JIA F, et al. A recurrent neural network based health indicator for remaining useful life

prediction of bearings[J]. Neurocomputing, 2017, 240: 98-109.

- ZHANG W, LI X, MA H, et al. Transfer learning using deep representation regularization in remaining useful life prediction across operating conditions [J]. Reliability Engineering and System Safety, 2021, 211: 107556.
- [8] CHEN J, LI D, HUANG R, et al. Aero-engine remaining useful life prediction method with self-adaptive multimodal data fusion and cluster-ensemble transfer regression [J]. Reliability Engineering and System Safety, 2023, 234: 109151.
- [9] MCMAHAN B, MOORE E, RAMAGE D, et al. Communication-efficient learning of learning of deep networks from decentralized data [C]. Artificial Intelligence and Statistics, 2017: 1273-1282.
- [10] COLLINS L, HASSANI H, MOKHTARI A, et al. Exploiting shared representations for personalized federated learning[C]. International Conference on Machine Learning, 2021: 2089-2099.
- [11] ZHANG W, LI X, MA H, et al. Federated learning for machinery fault diagnosis with dynamic validation and self-supervision [J]. Knowledge-Based Systems, 2021, 213:106679.
- [12] MA X, WEN C, WEN T. An asynchronous and real-time update paradigm of federated learning for fault diagnosis[J]. IEEE Transactions on Industrial Informatics, 2021, 17(12): 8531-8540.
- [13] 李响,付春霖,雷亚国,等.保证数据隐私的装备协同智能故障诊断 联邦迁移学习方法[J].机械工程学报,2023,59(6):1-9.
  LIX,FUCHL,LEIYG, et al. Federated transfer learning method for privacy-preserving collaborative intelligent machinery fault diagnostics [J]. Journal of Mechanical Engineering, 2023, 59(6): 1-9.
- [14] KAMEI S, TAGHIPOUR S. A comparison study of centralized and decentralized federated learning approaches utilizing the transformer architecture for estimating remaining useful life [J]. Reliability Engineering and System Safety, 2023, 233: 109130.
- [15] GUO L, YU Y, QIAN M, et al. FedRUL: A new federated learning method for edge-cloud collaboration based remaining useful life prediction of machines [J].
   IEEE/ASME Transactions on Mechatronics, 2023, 28(1): 350-359.
- [16] CHEN X, WANG H, LU S, et al. Bearing remaining useful life prediction using federated learning with Taylorexpansion network pruning [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 3512610.

- [17] LUO J, WU S. Adapt to adaptation: Learning personalization for cross-silo federated learning [C]. IJCAI: Proceedings of the Conference, 2022: 2166-2173.
- [18] CHEN X, WANG H, LU S, et al. Remaining useful life prediction of turbofan engine using global health degradation representation in federated learning [J]. Reliability Engineering and System Safety, 2023, 239: 109511.
- [19] LIU R, WANG F, YANG B, et al. Multiscale kernel based residual convolutional neural network for motor fault diagnosis under nonstationary conditions [J]. IEEE Transactions on Industrial Informatics, 2019, 16 (6): 37973806.
- [20] 王辉,徐佳文,严如强.基于多尺度注意力深度强化 学习网络的行星齿轮箱智能诊断方法[J].机械工程 学报,2022,58(11):133-142.

WANG H, XU J W, YAN R Q. Multi-Scale attention based deep reinforcement learning for intelligent fault diagnosis of planetary gearbox[J]. Journal of Mechanical Engineering, 2022, 58(11): 133-142.

- [21] GAL Y, GHAHRAMANI Z. Dropout as a bayesian approximation: Representing model uncertainty in deep learning [C]. International Conference on Machine Learning, 2016: 1050-1059.
- [22] LI N, LEI Y, LIN J, et al. An improved exponential model for predicting remaining useful life of rolling element bearings [J]. IEEE Transactions on Industrial Electronics, 2015, 62(12): 7762-7773.
- [23] ZHANG J, HUA Y, WANG H, et al. FedALA: Adaptive local aggregation for personalized federated learning[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(9): 11237-11244.
- [24] 雷亚国,韩天宇,王彪,等. XJTU-SY 滚动轴承加速 寿命试验数据集解读[J]. 机械工程学报, 2019, 55(16):1-9.
  LEI Y G, HAN T Y, WANG B, et al. XJTU-SY rolling element bearing accelerated life test datasets: A tutorial[J]. Journal of Mechanical Engineering, 2019,

55(16): 1-9.

- [25] NECTOUX P, GOURIVEAU R, MEDJAHER K, et al. PRONOSTIA: An experimental platform for bearings accelerated degradation tests [C]. IEEE International Conference on Prognostics & Health Management, 2012: 23-25.
- [26] LI X, XU Y, LI N, et al. Remaining useful life prediction with partial sensor malfunctions using deep adversarial networks [J]. IEEE/CAA Journal of Automatica Sinica, 2022, 10(1): 121-134.

#### 作者简介



陈曦,2017年于西安理工大学获得学士 学位,2020年于上海大学获得硕士学位,现 为东南大学博士研究生,主要研究方向为信 号处理、机械设备的智能健康管理。 E-mail:230208849@ seu.edu.cn

Chen Xi received her B. Sc. degree from

Xi'an University of Technology in 2017, and M. Sc. degree from Shanghai University in 2020. She is currently a Ph. D. candidate at Southeast University. Her current research interests include mechanical signal processing, federated learning and prognostics and health management for machine components.



严如强(通信作者),分别在 1997 年和 2002 年于中国科学技术大学获得学士学位 和硕士学位,2007 年于美国马萨诸塞大学 阿默斯特分校获得博士学位,现为西安交通 大学教授、博士生导师,主要研究方向为机 械系统状态监测与故障诊断、信号处理、无

线传感网络。

E-mail: yanruqiang@ xjtu. edu. cn

**Yan Ruqiang** (Corresponding author) received his B. Sc. and M. Sc. degrees both from University of Science and Technology of China in 1997 and 2002, and Ph. D. degree from University of Massachusetts, Amherst, USA in 2007, respectively. He is currently a professor and a Ph. D. advisor at Xi' an Jiaotong University. His main research interests include mechanical system condition monitoring and fault diagnosis, signal processing, and wireless sensor network.