

DOI: 10.19650/j.cnki.cjsi.J2514115

# 知识图谱与时空图神经网络融合驱动的风电机组状态监测\*

金晓航<sup>1,2</sup>, 王奇超<sup>1</sup>, 张元鸣<sup>3</sup>, 孔子迁<sup>4</sup>, 徐正国<sup>5</sup>

(1. 浙江工业大学机械工程学院 杭州 310023; 2. 高端装备机械传动全国重点实验室 重庆 400044;  
3. 浙江工业大学计算机科学与技术学院 杭州 310023; 4. 杭州电子科技大学“一带一路”  
信息技术研究院 杭州 310018; 5. 浙江大学控制科学与工程学院 杭州 310027)

**摘要:**在推动风能产业健康发展的过程中,风电机组状态监测发挥着至关重要的作用。现有数据驱动的状态监测方法主要依赖于风电机组时序类数据(如数据采集与监控系统(SCADA)、状态监测系统数据)的分析,未能有效利用机组文本类数据(如设计手册、操作手册、论文专利、运维记录、故障报告等)中蕴含的信息,在故障传递因果关系分析和监测结果可解释剖析等方面具有一定的局限性。鉴于此,提出了一种知识图谱与时空图神经网络(KG-STGNN)融合驱动的风电机组状态监测方法。首先,利用文本类数据结合机组结构设计等信息构建风电运维知识图谱,形成风电机组有向图结构;然后,将 SCADA 数据嵌入图结构中,生成风电时序图数据;接着,利用高阶图注意力网络(HGAT)和 Transformer 构建状态监测时空图神经网络模型,挖掘出图数据中的空间和时间特征;之后,利用机组历史健康数据训练 KG-STGNN 模型,进行正常行为建模;最后,根据风电机组运行时图结构中节点所表征的信息判断机组的运行状态,构建监测策略以确定故障预警时间并解释状态监测结果。通过两台风电机组案例分析可知:所提方法在机组状态监测中表现优异,具有最低误报率和最早异常预警能力;消融实验验证了引入知识图谱对模型性能提升至关重要;所提监测策略消除了超过 85% 的误报情况,对监测结果也具有较好的可解释性。

**关键词:** 风电机组; 状态监测; 知识图谱; 时空图神经网络; 图注意力网络

**中图分类号:** TH17 **文献标识码:** A **国家标准学科分类代码:** 460.99

## Condition monitoring of wind turbines driven by the integration of knowledge graph and spatio-temporal graph neural network

Jin Xiaohang<sup>1,2</sup>, Wang Qichao<sup>1</sup>, Zhang Yuanming<sup>3</sup>, Kong Ziqian<sup>4</sup>, Xu Zhengguo<sup>5</sup>

(1. College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310023, China;

2. State Key Laboratory of Mechanical Transmission for Advanced Equipment, Chongqing 400044, China;

3. College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China;

4. Belt and Road Information Technology Research Institute, Hangzhou Dianzi University, Hangzhou 310018, China;

5. College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China)

**Abstract:** In promoting the healthy development of the wind energy industry, condition monitoring of wind turbines (WT) plays a crucial role. Existing data-driven condition monitoring methods primarily rely on time-series data, such as supervisory control and data acquisition (SCADA) and condition monitoring system data analysis, failing to effectively utilize the information contained in the WT textual data, such as design specifications, operation manuals, papers, patents, maintenance records, fault reports, etc. They have limitations in the analysis of fault transmission causality and the interpretability of analysis results. Therefore, this article proposes a WT condition monitoring method driven by the fusion of a knowledge graph and spatio-temporal graph neural network (KG-STGNN). This method first utilizes textual data combined with WT structural and other information to construct a WT operation and maintenance

收稿日期: 2025-06-04 Received Date: 2025-06-04

\* 基金项目: 国家自然科学基金(62473336)、浙江省自然科学基金(LZ25F030004)、国家重点研发计划(2022YFE0198900)、高端装备机械传动全国重点实验室开放课题(SKLM-T-MSKFKT-202315)项目资助

knowledge graph, forming a directed graph structure for WTs. Then, SCADA data are embedded into the graph structure to generate WT time-series graph data. A high-order graph attention network (HGAT) and a Transformer are used to establish a spatio-temporal graph, the spatial and temporal characteristics in the graph data are mined for condition monitoring. Historical healthy data of the WT are used to train the KG-STGNN model. Finally, the operating status of the WT is determined based on the information represented by nodes in the graph, and a monitoring strategy is constructed to determine fault warning times and interpret the condition monitoring results. Analysis of actual WT cases evaluates the effectiveness of the proposed method, with its performance outperforming traditional graph neural network models. Case studies on two WTs show that the proposed method exhibits excellent performance in condition monitoring, since it has the lowest false alarm rate and the earliest anomaly warning capability. Ablation experiments show that the graph structure constructed from the knowledge graph is critical for improving model performance. The proposed monitoring strategy eliminates more than 85% of false alarms and also provides good interpretability for the monitoring results.

**Keywords:** wind turbine; condition monitoring; knowledge graph (KG); spatio-temporal graph neural networks (STGNNs); graph attention network (GAT)

## 0 引言

随着全球经济的迅速发展,世界各国对环境问题也愈发重视<sup>[1]</sup>。能源需求从传统的化石能源过渡到以风能等为代表的可再生能源。近年来,随着我国“碳达峰”、“碳中和”发展目标的提出,以风能等为代表的清洁能源迎来新的发展机遇<sup>[2]</sup>,其中风力发电技术的发展尤为迅速。然而,由于风资源丰富区多位于环境恶劣、通讯不畅、交通不便的内陆偏远地区以及近海区域,受恶劣的环境及复杂多变的工况影响导致风电机组故障频发。因此,为保障风电机组安全运行,避免因故障导致的长时间停机和巨大经济损失,迫切需要发展状态监测、异常预警等技术,以精准掌握其运行状态并提升运行可靠性<sup>[3-4]</sup>。

风电机组状态监测方法主要有数据驱动与物理模型驱动等方法<sup>[5]</sup>。目前,随着人工智能技术的兴起,数据驱动的方法已成为风电机组状态监测领域的研究热点。该类方法主要围绕数据采集与监视控制(supervisory control and data acquisition, SCADA)<sup>[6-8]</sup>系统为核心进行数据分析与建模。从监测的特征数量和记录时间来看,SCADA数据具有维度高<sup>[9]</sup>、时间跨度大的特点;系统所监测的变量来自机组的不同部件,特征参数之间存在较强的非线性关系<sup>[10]</sup>。鉴于SCADA数据包含了丰富的机组运行状态信息<sup>[11-12]</sup>,许多学者基于SCADA数据构建了风电机组状态监测的深度学习模型。如金晓航等<sup>[13]</sup>将稀疏自编码器和深度神经网络算法相结合,提出了一种基于SCADA数据的风电机组在线运行状态监测方法。Xiang等<sup>[14]</sup>提出了一种结合卷积神经网络(convolutional neural network, CNN)和双向门控递归单元的时空特征模型来提取SCADA数据中包含的机组运行状态信息。Wang等<sup>[15]</sup>提出了一种长短时记忆网络(long short-term memory, LSTM)与Transformer融合的模式,通过多风电

场SCADA数据迁移学习策略,增强了风电机组齿轮箱故障检测建模能力与跨机组故障诊断的泛化能力。上述风电机组状态监测方法在假设输入特征相互独立的基础上提取出监测数据中时间维度上的信息,忽略了输入特征之间潜在的联系<sup>[16-17]</sup>。图神经网络(graph neural networks, GNN)作为一种擅长处理拓扑结构数据的技术,在建模复杂系统、捕捉节点间潜在关系方面有独特优势<sup>[18]</sup>,可被用于解决上述问题。

在GNN框架中,数据通过图结构表示,利用节点之间的信息传播和特征更新,提取出更多与任务相关的信息。因此,GNN为风电机组状态监测提供了新的思路,通过将机组各部件之间的关系和信息流动建模为图,GNN能够捕捉系统性的故障模式,从而进一步提高监测精度并捕捉到传统方法无法识别潜在的故障信号。鉴于信息传播方式、节点信息更新策略的多样性以及图数据的独特性质,GNN衍生出了多种变体,如图卷积网络(graph convolutional networks, GCN)<sup>[19]</sup>、图注意力网络(graph attention networks, GAT)<sup>[20]</sup>、变分图自编码(variational graph auto-encoders, VGAE)<sup>[21]</sup>、高阶图注意力网络(high-order graph attention network, HGAT)<sup>[22]</sup>等,这些方法为风电机组的状态监测提供了更加丰富的建模手段。Feng等<sup>[23]</sup>使用深度图神经网络结合局部-全局最大信息交互实现物理统计特征融合的方法进行风电机组状态监测;Jin等<sup>[24]</sup>使用先验知识构图,建立图时空网络模型(graph spatio-temporal network, GSTN)框架,进行风电机组状态监测,并增强了监测结果的可解释性;Wang等<sup>[25]</sup>为了解决基于密度的聚类算法中的参数敏感性问题,提出了一种使用两层自适应GCN模型用于捕获风电机组中的复杂时空相关性。

然而,现有的基于图神经网络的风电机组状态监测方法在构建图数据时,未能有效利用风电机组海量的文本数据<sup>[26]</sup>。与此同时,尽管借助文本数据的大模型技术(如预训练语言模型)已在工业缺陷检测等领域展现出

强大潜力<sup>[27]</sup>,但风电状态监测领域的方法却未能有效借鉴这一趋势,这导致文本数据的价值被长期忽视。目前,有部分学者尝试使用先验知识来构建图数据,但先验知识依赖个人经验,难以灵活应用于不同机组中。此外,当前监测方法的分析过程和结果在可解释性方面也较为欠缺。

因此,为充分利用风电文本类数据和风电 SCADA 数据,提升监测性能与结果的可解释性,提出了一种知识图谱与时空图神经网络(knowledge graph and spatio-temporal graph neural network, KG-STGNN)融合驱动的风电机组状态监测方法,主要贡献为:

1) 提出了一种知识图谱驱动的 GNN 图数据生成方法,将风电机组中的文本类数据构建为可统一管理的运维知识图谱,根据故障传递因果关系,得到风电机组图结构,结合 SCADA 数据生成输入 GNN 的风电时序图数据。

2) 提出了一种 KG-STGNN 模型,通过将 HGAT 与 Transformer 相结合提取时序图数据中蕴含的时空信息,提高了风电机组状态监测的准确性。

3) 提出了一种节点级监测策略,通过判断故障传递链的建立情况,确定故障预警的时间。该策略充分利用了知识图谱中蕴含的故障模式和传递路径,有效减少了误报现象的发生。

## 1 KG-STGNN 状态监测方法概述

知识图谱与时空图神经网络融合驱动的风电机组状态监测方法主要包括风电机组图结构构建、时序图数据生成、正常行为建模和状态监测分析 4 个部分,具体流程如图 1 所示。

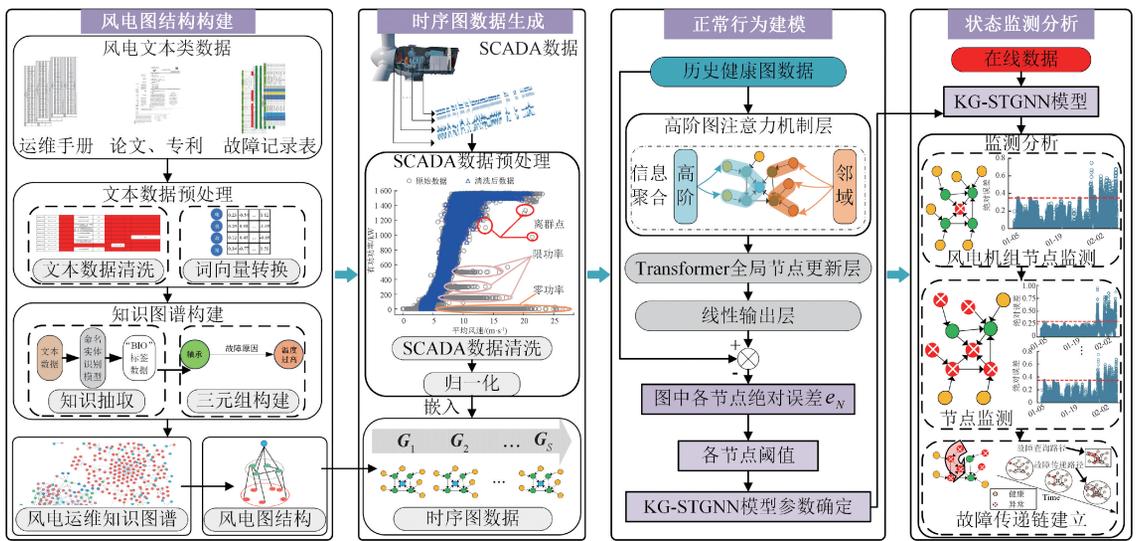


图 1 知识图谱与图神经网络融合驱动的风电机组状态监测方法

Fig. 1 A condition monitoring method for wind turbines driven by the fusion of knowledge graph and graph neural network

1) 在风电机组图结构构建部分,为突破传统方法中输入特征相互独立、忽视特征间潜在关联的局限,该部分首先对风电机组文本类数据进行数据清洗以及向量转换,随后通过双向长短期记忆网络条件随机场(bi-directional long short-term memory-conditional random field, BiLSTM-CRF)的命名实体识别模型进行实体抽取并构建三元组。再通过 Neo4j 进行可视化得到风电运维知识图谱。在获得风电运维知识图谱后,为了充分利用其中包含的故障信息进行机组状态监测,可以通过查询机组各子系统及传感器之间的故障传递因果关系,构建风电机组的图结构,从而为后续分析奠定结构化、可解释的关系基础。

2) 在风电时序图数据生成部分,为实现文本数据与 SCADA 数据的深度融合,在对 SCADA 时序数据进行清洗等预处理分析后,将对应的传感器数据嵌入至风电

机组图结构中,生成风电时序图数据,有效解决了传统方法未能充分运用文本数据的问题。

3) 提出了一种基于 KG-STGNN 的正常行为建模框架。首先,使用 HGAT 层聚合高阶邻域信息进行节点特征更新,有效捕捉图结构中蕴含的故障传播空间特征。其次,将更新后的时序图数据输入至 Transformer 层,通过 Transformer 的全局自注意力机制捕捉各节点自身在时序图数据中的时间特征。最后,通过线性层输出图中各节点的预测结果,并结合正常行为建模原理,通过分析正常运行状态下的数据,确定风电机组状态监测的健康阈值。

4) 在状态监测分析阶段,提出节点级状态监测策略,通过异常时刻各节点表现与图结构中故障传递链的匹配分析,在实现故障预警的同时,提供了由故障传播路径支撑的可解释结果,从而显著降低误报率。

## 2 知识图谱驱动的风电机组图数据生成

风电时序图数据生成主要分以下两部分内容:风电机组图结构构建与时序 SCADA 数据嵌入,其中图结构基于风电文本类数据的分析构建运维知识图谱后查询得到。

### 2.1 知识图谱驱动的风电机组图结构构建

知识图谱本质上作为一个结构化语义网络<sup>[28-30]</sup>,通过“实体”、“属性”及“联系”等要素构成三元组的形式<sup>[31-32]</sup>,以实现系统化整合大规模网络信息以及可视化和图谱管理任务<sup>[33-35]</sup>。知识图谱的构建分为自顶向下和自底向上两种方法,其中自顶向下的构建方法主要依据已整合完成的高质量数据源,从中提取相关的知识图谱本体及其相关信息;而自底向上的构建方法需要通过一些技术手段,从数据集中提取出所需要的知识,适用于领域知识图谱构建<sup>[36]</sup>。由于风电机组具有丰富的运维文本数据,缺少整合完成的高质量数据源,因此使用自底向上的方法对风电机组运维知识图谱进行构建。

首先,将与风电机组领域相关的设备、流程、故障、维护记录及论文专利等文本数据进行数据预处理。再通过实体抽取技术将文本数据转换为三元组结构。最后,通过 Neo4j 软件可视化三元组,得到风电运维知识图谱。该知识图谱利用图形化的数据结构将风电机组运维领域中的文本数据统一管理,并在此基础上实现知识的更新、分析及应用。

尽管风电文本数据包含复杂且多样的非结构化信息,但其核心要素相对固定,能够提前定义其知识架构中的关系<sup>[37]</sup>。因此,采用实体抽取方法结合已定义的各实体间关系构建风电机组运维知识图谱。首先,对文本数据进行数据预处理:缺失值删除以及采用已有文献<sup>[28]</sup>中的文本词向量转化方法。其次,使用命名实体识别模型对文本数据进行实体抽取。由于风电文本数据中包含较多的故障部件名称及其故障原因,因此需要命名实体识别模型既能充分考虑上下文关系,又能精准捕捉句子中的语义信息。基于此,文中采用 BiLSTM-CRF 命名实体识别模型<sup>[38]</sup>进行实体抽取任务。BiLSTM-CRF 将文本数据打上“B、I、O”标签进行风电运维三元组的构建,其中 B (Beginning) 表示某个实体词的开始, I (Inside) 表示某个实体词的中间, O (Outside) 表示非实体词。

以“发电机超速”为例,通过 BiLSTM-CRF 进行序列标注任务的流程为:1) 将文本转换为词向量后,输入到一个前向和后向的 LSTM 中进行前后语义捕捉;2) 再通过 CRF 模型输出符合条件的最大可能的预测标注序列;3) 最终得到的结果用 B、I、O 表示为“B、I、I、B、I”,如图 2 所示。

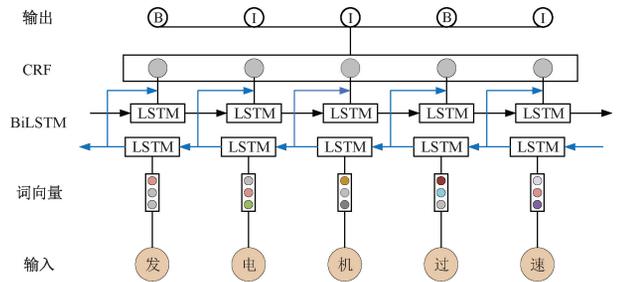


图 2 BiLSTM-CRF 模型示例

Fig. 2 Example of the BiLSTM-CRF model

其中“发电机”作为‘故障设备’被识别为头实体,“超速”作为‘故障情况’被识别为尾实体,将“故障原因”作为关系得到(“发电机”,“故障原因”,“超速”)三元组,将发电机部分故障情况通过可视化平台得到其知识图谱如图 3 所示。

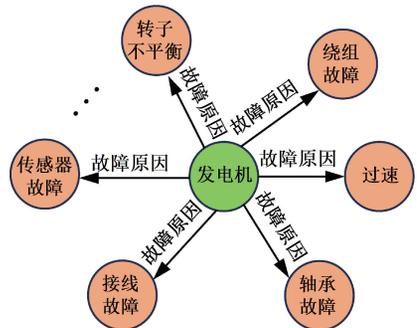


图 3 三元组可视化示例

Fig. 3 Example of the triplet visualization

在文本数据预处理后,利用命名实体识别模型提取实体,进而得到风电运维三元组,并通过 Neo4j 进行可视化,从而构建风电运维知识图谱。与自顶向下的先验知识构图方法相比,风电运维知识图谱融合了风电机组运维手册、故障记录表、学术论文、专利等多种文本数据,并通过在不同类别实体间建立联系,构成了先验知识的集合。同时,知识图谱能够根据实时物料数据进行更新,确保其正确性和准确性。

通过查询风电运维知识图谱中当前机组子系统与传感器之间的故障传递因果关系,并保留故障现象频繁出现的传感器节点,可以将查询得到的节点划分为 3 层,如图 4 所示。

其中,顶层为风电机组整机节点,各部件的故障最终都会反映在风电机组整机中。第 2 层为子系统层节点,由于风电机组各子系统在真实的空间环境中存在结构连接,当某子系统发生故障时,与其有结构连接或者真实物理世界中相距较近的子系统也会受到故障影响。第 3 层为风电机组各子系统内置的多种传感器层,目前的

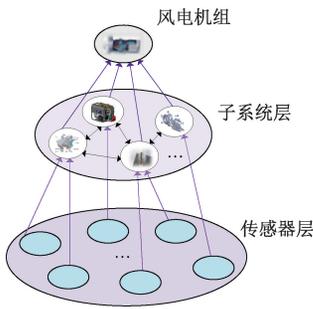


图 4 风电机组图结构

Fig. 4 Wind turbine graph structure

SCADA 系统预警信息均来自于传感器的预警,因此其代表故障信息传递的源头。该图结构可用于描述风电机组在整机、各子系统与传感器之间的关联。将时序 SCADA 数据嵌入到对应节点中,生成风电时序图数据,以描述风电机组在一定时间范围内,传感器监测信息、子系统信息、整机信息之间的相互关系。

## 2.2 风电时序图数据生成

考虑到风电机组中各子系统及传感器之间的故障传播空间相关性,由整机、子系统和传感器组成的知识图谱作为风电机组图结构。将上述图结构表达为  $G=(V,E)$  的图数据作为 GNN 的输入,其中  $V$  代表知识图谱图结构中节点的集合, $E$  是代表节点间信息传递的边的集合。 $v_i \in V$  表示第  $i$  个节点, $e_{ij}=(v_i,v_j) \in E$  表示一条从  $v_i$  指向  $v_j$  的有向边。为了更好地储存知识图谱所内含的故障因果关系,采用邻接矩阵  $A \in R^{N \times N}$  和特征矩阵  $X \in R^{N \times F}$  分别作为知识图谱图结构中节点信息和节点间关系信息,其中  $N$  代表节点个数, $F$  代表节点特征的维度。因为在知识图谱中,各节点之间的边均为有向边,代表在故障信息传递途径中,节点  $v_i$  中的信息可以沿着有向边传递到另一个节点  $v_j$  中,因此在邻接矩阵  $A$  中,如果节点  $i$  的有向边指向节点  $j$  则在矩阵  $A$  中,元素  $A_{ij}$  为 1, $A_{ji}$  为 0。

在数据预处理阶段,对原始 SCADA 数据进行清洗<sup>[39]</sup>和归一化处理。清洗前后 SCADA 数据对比如图 5 所示。选取图  $G$  中各节点对应的时序 SCADA 特征数据作为图数据的特征矩阵,完成时序 SCADA 数据的嵌入,形成风电时序图数据,如图 6 所示。

## 3 基于 KG-STGNN 的风电机组正常行为建模与状态监测分析

由于风电机组运行数据中健康数据远多于故障数据,存在严重的数据不平衡问题。因此,必须重点关注如何有效利用健康数据所代表的正常行为进行建模,以确保监测模型能够准确学习机组在正常运行状态下的行为特征,建

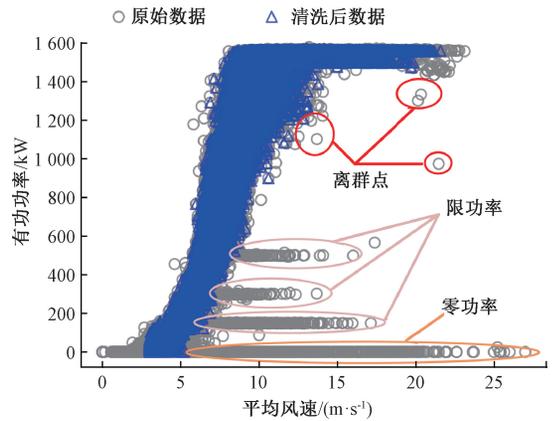


图 5 SCADA 数据清洗前后对比

Fig. 5 Comparison of SCADA data before and after cleaning

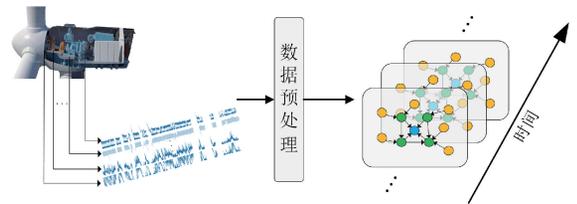


图 6 SCADA 数据嵌入

Fig. 6 SCADA data embedding

立正常行为下的健康阈值。该方法可使监测模型在掌握正常运行模式的同时,增强对异常数据的敏感度<sup>[40]</sup>。

在基于正常行为建模的监测分析过程中,当 SCADA 特征的预测值与实际值间的偏差超过正常行为阈值时,可将风电机组的运行状态判别为异常<sup>[40]</sup>。因此,结合图神经网络的风电机组的状态监测方法可以描述为:将一个时间窗口中的图  $G_{1:T}$  输入至函数  $f$ ,通过一个窗口内的时间序列与图结构提取风电机组的正常行为特征,对窗口后的下一个图进行预测,如式(1)所示。

$$f: [G_1, \dots, G_T] \Rightarrow \hat{G}_{T+1} \quad (1)$$

最后,得到每个窗口的预测图与实际图之间的绝对误差 (absolute error, AE),如式(2)所示。

$$G_e = |\hat{G}_{T+1} - G_{T+1}| \quad (2)$$

式中:  $G_{T+1}$  代表窗口后的下一幅实际图; $\hat{G}_{T+1}$  代表对应的预测图; $G_e$  为反映图异常情况的绝对误差图。在此基础上,提出了 KG-STGNN 的风电机组状态监测模型。该模型结合了知识图谱与时空图网络的优势,可有效捕捉风电机组运行中的时空信息,从而有效提升状态监测的性能。

### 3.1 KG-STGNN 模型

所提出的时空图网络模型主要由 3 个模块组成,分别为 HGAT 层、Transformer 层和线性输出层。首先, HGAT 层用于捕捉时序图数据的空间特征;其次,

Transformer 层用于捕捉滑动窗口中各节点自身的全局时间特征;最后,线性输出层将空间特征与时间特征相结合,预测各节点的下一步状态。

在 HGAT 层模块中,由于 HGAT 方法面对多节点时在避免过度平滑问题和提取图中高阶信息方面具有显著优势,因此将其应用于知识图谱驱动的风电机组状态监测中。通过多头注意力机制计算自适应边权重对不同阶邻居节点的特征进行聚合。需要注意的是所提出的 HGAT 模块不同于 GAT,如图 7 所示。

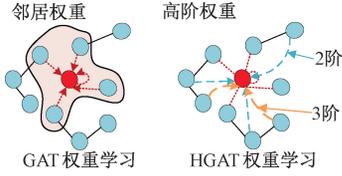


图 7 GAT 与 HGAT 在权重学习上的差异

Fig. 7 The difference between GAT and HGAT in weight learning

GAT 在学习权重的过程中,仅为自身以及邻居节点计算权重,中心节点仅可接收到邻居范围内的信息,而 HGAT 不仅可以计算邻居范围的权重信息,还可以为不同阶邻居节点分配权重。随着阶数增加,可以学习到更多的高阶节点信息来更新节点。

节点更新过程中节点之间的注意力系数计算如式(3)所示。

$$m_{i,l} = \frac{\exp(\mathbf{u}_i \mathbf{x}_{i,l})}{\sum_{k=0}^L \exp(\mathbf{u}_i \mathbf{x}_{i,k})} \quad (3)$$

式中:矩阵  $\mathbf{U} \in \mathbb{R}^{n \times n}$  表示注意力网络的参数矩阵; $\mathbf{u}_i$  为矩阵  $\mathbf{U}$  中第  $i$  行; $m_{i,l} \in \mathbb{R}^{L+1}$  表示包含自环的  $i$  节点的  $l$  阶邻居域权重; $L$  表示阶数; $\mathbf{x}_{i,l}$  表示  $i$  节点  $l$  阶聚合后的节点特征。

将不同阶邻域的特征与获得的节点间权重进行聚合,以获得特征更新后的节点表示。设  $\mathbf{h}_i \in \mathbb{R}^{F'}$  为节点  $i$  的  $l$  阶邻居特征聚合后的节点特征, $F'$  表示图中每个节点的新维度,则节点的输出结果可以表示为:

$$\mathbf{h}_i = \frac{1}{M} \sigma \left( \sum_{m=1}^M \sum_{l=0}^L m_{i,l} \mathbf{x}_{i,l} \right) \quad (4)$$

式中: $M$  表示注意力头数; $\sigma$  表示激活函数。为了对该时刻所有节点的空间特征进行拼接整合,最终 HGAT 层的输出为:

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_N] \quad (5)$$

式中: $N$  为图结构中节点个数; $\mathbf{H} \in \mathbb{R}^{N \times F'}$  表示图结构通过 HGAT 层构建的空间特征。

在捕捉空间特征信息之后,利用 Transformer 层计算每个节点特征对整个窗口的贡献,捕捉滑动窗口中各

节点的全局时间特征信息。由于 HGAT 的输出是具有一定窗口长度  $S$  的图序列。则  $\mathbf{H}_i \in \mathbf{H}_{1:S}$  ( $i=1,2,\dots,S$ ) 表示 HGAT 层某时刻  $t$  的图特征矩阵。

Transformer 层进行节点全局时间信息更新,如图 8 所示。图数据通过 3 个独立的全连接层进行转换,生成相应的查询矩阵  $\mathbf{Q} \in \mathbb{R}^{N \times d_k}$ 、键矩阵  $\mathbf{K} \in \mathbb{R}^{N \times d_k}$ 、值矩阵  $\mathbf{V} \in \mathbb{R}^{N \times d_v}$ , $d_k$  与  $d_v$  为单注意力头数下矩阵  $\mathbf{K}$  和  $\mathbf{V}$  维度。使用单注意力头数在  $t$  时刻对节点  $i$  进行全局时间信息更新时,通过全连接层进行转换的公式为:

$$\mathbf{Q}_i^k = \mathbf{W}_1 \mathbf{h}_i^k + \mathbf{b}_1 \quad (6)$$

$$\mathbf{V}_i^t = \mathbf{W}_2 \mathbf{h}_i^t + \mathbf{b}_2 \quad (7)$$

$$\mathbf{K}_i^t = \mathbf{W}_3 \mathbf{h}_i^t + \mathbf{b}_3 \quad (8)$$

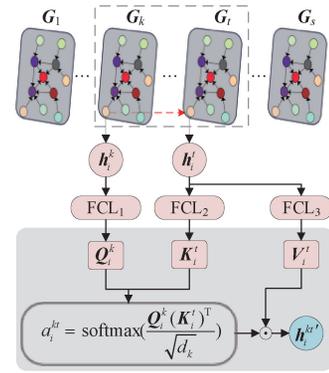


图 8 Transformer 层节点更新示意图

Fig. 8 Illustration of node update in the Transformer layer

式中: $\mathbf{W}_1 \in \mathbb{R}^{1 \times d_k}$ ,  $\mathbf{W}_2 \in \mathbb{R}^{1 \times d_v}$ ,  $\mathbf{W}_3 \in \mathbb{R}^{1 \times d_k}$  为可学习矩阵; $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$  为对应的偏置; $\mathbf{Q}_i^k \in \mathbb{R}^{1 \times d_k}$  为  $k$  时刻节点  $i$  的查询矩阵; $\mathbf{V}_i^t \in \mathbb{R}^{1 \times d_v}$  与  $\mathbf{K}_i^t \in \mathbb{R}^{1 \times d_k}$  为  $t$  时刻节点  $i$  的值矩阵与键矩阵; $\mathbf{h}_i^k$  与  $\mathbf{h}_i^t$  分别为  $k$  时刻节点  $i$  特征与  $t$  时刻节点  $i$  特征。 $k$  时刻对  $t$  时刻的节点  $i$  注意力权重由  $\mathbf{Q}_i^k$  和  $\mathbf{K}_i^t$  计算得到,即:

$$a_i^{kt} = \text{softmax} \left( \frac{\mathbf{Q}_i^k (\mathbf{K}_i^t)^T}{\sqrt{d_k}} \right) \quad (9)$$

通过上述计算,可以得到节点更新后特征为:

$$\mathbf{h}^{kt} = a_i^{kt} \cdot \mathbf{V}_i^t \quad (10)$$

式中: $\mathbf{h}^{kt} \in \mathbb{R}^{1 \times d_v}$  为  $k$  时刻对  $t$  时刻的节点  $i$  信息更新特征。则  $t$  时刻通过 Transformer 进行节点全局时间信息更新后节点  $i$  特征为:

$$\mathbf{h}_i^t = \sum_{j=1}^S a_i^{jt} \mathbf{V}_i^t \quad (11)$$

式中: $\mathbf{h}_i^t \in \mathbb{R}^{1 \times d_v}$  表示  $t$  时刻节点  $i$  的更新特征,整合了其他时间步中的信息。通过单注意力头数的 Transformer 层后  $t$  时刻图数据特征表示为:

$$\mathbf{H}^t = [\mathbf{h}_1^t, \mathbf{h}_2^t, \dots, \mathbf{h}_N^t] \quad (12)$$

式中:  $\mathbf{H}' \in \mathbb{R}^{N \times d_v}$ 。

为了捕获多样化的特征模式,Transformer 层采用了多头注意力机制,将各注意力头数输出的图数据拼接并进行线性变换后得到最终输出:

$$MHead(\mathbf{H}') = \text{Concat}(\mathbf{H}'_1, \mathbf{H}'_2, \dots, \mathbf{H}'_Z) \mathbf{W}^0 \quad (13)$$

式中: $Z$  为注意力头数;  $\mathbf{W}^0 \in \mathbb{R}^{(Z \times d_v) \times d_{\text{model}}}$  为全局的线性投影矩阵;  $d_{\text{model}}$  表示隐藏层维度。

最终,通过线性输出层计算 Transformer 层输出的各节点特征向量,预测图结构中各节点的未来状态。

### 3.2 风电机组正常行为建模

使用健康数据训练 KG-STGNN 模型进行正常行为模型构建,从而得到健康阈值。因为状态监测图结构中存在多个节点,因此使用每个节点自身的预测残差表示各节点自身的异常情况。通过采用正常行为模型导出的参数(AE)作为各节点异常情况指标,如式(14)所示。

$$e_N = |N_{\text{predict}} - N_{\text{real}}| \quad (14)$$

式中: $N_{\text{predict}}$  为节点的模型预测值;  $N_{\text{real}}$  为节点真实值;  $e_N$  为节点绝对误差。将各节点在正常状态下的最大绝对误差作为状态监测阈值。当进行状态监测时若该节点存在超过阈值现象,则判定该节点出现异常行为,并触发预警。

### 3.3 基于 KG-STGNN 的节点级状态监测策略

考虑到风电机组在实际运行中存在变工况、环境变化快速及传感器种类繁多等情况,单一节点的监测结果往往因信息量不足,无法准确反映风电机组的整体健康状况,从而出现错误预警。为克服这一局限,提出一种基于节点级故障传递链的监测策略,如图 9 所示。该策略利用状态监测图结构中的故障传递因果关系,通过对多个节点进行联动监测,全面追踪故障信息在不同节点间的传递路径,从而提供更加准确的故障诊断和预测。

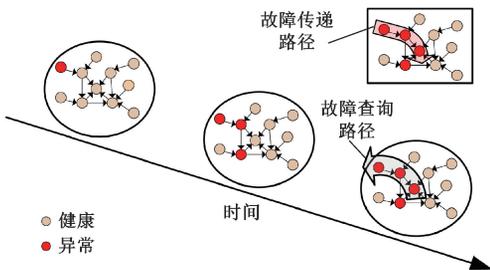


图 9 节点级故障传递链示意图

Fig. 9 Illustration of node-level fault propagation chain

在对风电机组进行节点级状态监测时,考虑到图神经网络的信息聚合特性,若某一节点的绝对误差超出预定阈值成为异常节点,接下来一段时间内该节点会将异常信息传递至图结构中相邻的节点,进而导致相邻节点

出现异常情况。因此,基于状态监测图结构的风电机组监测方法首先对顶端风电机组节点进行状态检测,若发现异常预警,再向下一层级节点发起异常查询,并沿着树状结构逐层向下查询,直至到达最底层的传感器节点,形成自顶向下的故障查询链路。若逆向的故障查询链路与有向的状态监测图结构相匹配,则可视为故障传递链建立。故障传递链为故障查询链路的逆向过程,如图 9 所示。在建立故障传递链后,将链路中的起始传感器节点视为本次异常的起点,即本次风电机组异常情况的故障源或靠近故障源的节点。该方法能节约计算成本,聚焦于有异常的部分,有效地提高监测的精准度和效率。

## 4 风电机组案例分析与讨论

通过对两台双馈式风电机组的数据分析来验证所提状态监测方法的有效性,相关机组信息如表 1 所示。机组的额定功率均为 1.5 MW,其中来自东北平原的案例 1 机组为正常运行的风电机组,来自西南高原的案例 2 机组 SCADA 系统在 2017 年 2 月 5 日监测到发电机轴承温度过高而进行停机保护,检修时发现该轴承发生了故障。鉴于二者源自同一厂家且额定功率相同,可将其整合运维知识,构建统一风电运维知识图谱以实现高效的状态监测。

表 1 案例分析用的机组信息

Table 1 Information of the wind turbines used in the case study

机组信息	案例 1	案例 2
额定功率	1.5 MW	1.5 MW
训练集时间	2016-01 至 2016-02	2016-01 至 2016-03
验证集时间	2016-02 至 2016-03	2016-09 至 2016-11
测试集时间	2016-04 至 2016-05	2017-01 至 2017-02
采样频率	每 5 min 采样一次	每 5 min 采样一次
健康情况	健康	故障(2017-02-05 发电机故障)

### 4.1 案例 1-健康机组状态监测

根据 2.1 节中所述的风电机组图结构构建方法,对文本数据进行缺失值删除后,通过 Word2vec 将文本中词表示为向量形式。最终通过命名实体识别方法进行实体抽取,得到风电运维知识图谱,如图 10 所示。

对风电运维知识图谱进行关键子系统与传感器的故障因果信息查询,得到当前机组以风电机组为中心节点的风电机组图结构,如图 11 所示。图中各节点代表的连续变量传感器特征或子系统名称参见表 2 中的信息。

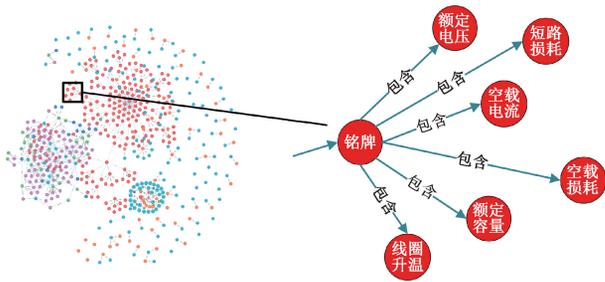


图 10 风电运维知识图谱

Fig. 10 Wind turbine operation and maintenance knowledge graph

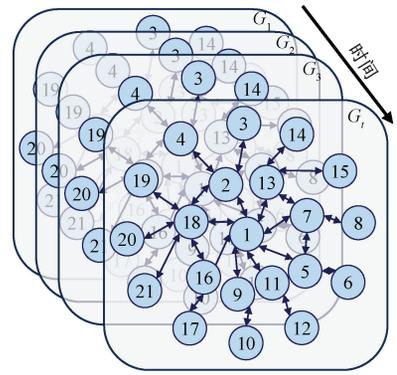


图 12 风电机组时序图数据

Fig. 12 Time-series data of wind turbines

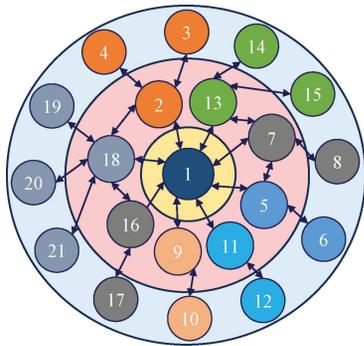


图 11 风电机组图结构

Fig. 11 Wind turbines graph structure

表 2 风电图结构各节点信息

Table 2 Information about each node in the wind turbine graph structure

节点	节点信息	节点	节点信息
1	风电机组	12	塔底控制柜温度
2	发电机	13	外部环境传感器
3	发电机转速	14	平均风速 30 s
4	发电机非驱动端轴承温度	15	环境温度
5	液压系统	16	主轴
6	液压系统压力	17	主轴转速
7	偏航系统	18	齿轮箱
8	偏航夹角	19	齿轮箱泵出口压力
9	机舱	20	齿轮箱泵进口压力
10	机舱温度	21	齿轮箱油温
11	塔架	—	—

表 3 时序图数据中三元组语义关系

Table 3 Semantic relationships of triples in temporal graph data

头实体	关系	尾实体	语义关系
发电机转速	故障传递	发电机	发电机转速传感器发生故障会传递至发电机
齿轮箱油温	故障传递	齿轮箱	齿轮箱油温传感器发生故障会传递至齿轮箱
发电机	部件组成	风电机组	发电机属于风电机组
偏航系统	故障传递	液压系统	偏航系统发生故障会传递至液压系统
液压系统	能量传递	偏航系统	液压系统为偏航系统提供能量
主轴	故障传递	齿轮箱	主轴发生故障会传递至齿轮箱
外部环境传感器	部件组成	风电机组	外部环境传感器属于风电机组

中每个子系统及中心风电机组节点都有相应的状态。如  $G_t$  时刻节点 4 与节点 2 及二者的连边表示发电机状态与发电机转速监测值相关,因此根据故障传递关系,得到该三元组语义关系,即发电机转速节点发生故障会影响发电机节点的状态,图结构中部分节点语义关系如表 3 所示。通过上述三元组语义关系,结合 HGAT 层,即可完成传感器节点特征信息沿着图结构传递至子系统节点与风电机组节点,使得各子系统节点包含对应的传感器信息,风电机组节点包含所有传感器节点的特征信息,以此进行节点级风电机组状态监测。

参照构建完成的风电机组图结构,建立图模型邻接矩阵,将对应预处理后的 SCADA 数据嵌入到风电机组图结构(节点及节点间的连边,如图 11 所示)中得到风电机组时序图数据,如图 12 所示。

在后续的状态监测过程中,图 12 给出的时序图数据

### 1) 健康机组的状态监测分析

在数据预处理以及时序图数据生成后,使用健康时段的时序图数据在 Pytorch 深度学习框架上进行正常行为模型的训练。所构建的风电机组图结构包含 21 个节点,模型需对 21 个节点进行多输入多输出的回归任务。

为了实现对各时刻节点数据的预测,采用滑动窗口法划分时间序列数据:具体来说,滑动窗口长度设为 96,滑动步长设为 1,以满足测试阶段对每个时间步均需进行预测的需求,如图 13 所示为滑动窗口大小选择依据,其中纵坐标为均方根误差(root mean squared error, RMSE)。

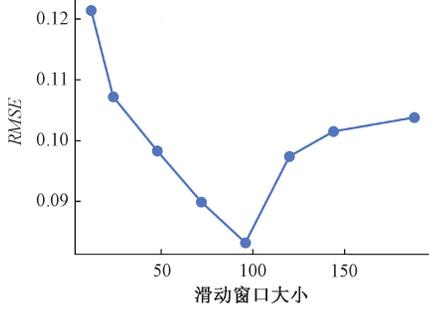


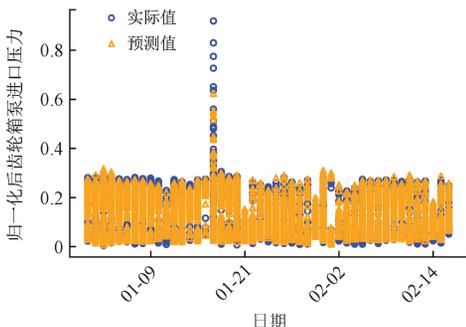
图 13 滑动窗口大小选取

Fig. 13 Selection of sliding window size

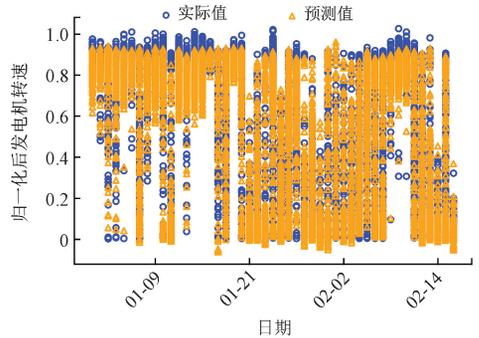
由于计算资源受限,该模型未采用自动化的超参数优化方法,而是基于经验法设置了学习率、训练轮数和提前终止等训练参数。学习率初值设为  $1 \times 10^{-3}$ ,并引入学习率衰减机制:考虑到多输入多输出任务的训练较为复杂,将学习率衰减周期设为 20 轮(约为总训练轮数的  $1/10 \sim 1/2$ )。最大训练轮数通过模型初步运行后观察损失变化趋势确定为 200 轮,同时将提前停止轮数设置为 20 轮以防止过拟合。此外,模型结构中包含多头注意力机制,注意力头数设为 4,各头的维度取总隐藏层维度除以头数。

如图 14 所示,部分节点的预测值与真实值对比结果表明,模型能够有效捕捉 SCADA 数据的动态变化特征,可以看出所提状态监测方法在预测时序数据方面有一定的优势。

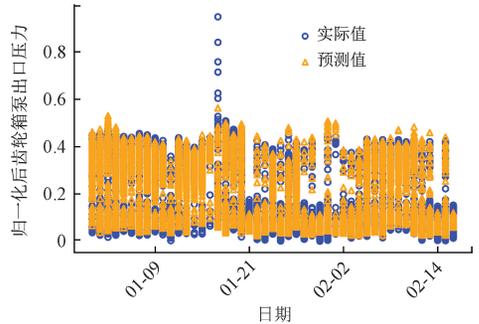
根据 3.3 节描述的节点级状态监测方法,结合模型正常行为学习后各节点的阈值,对测试集数据进行预测和绝对误差的计算,节点级状态监测存在异常的节点结果如图 15 所示。



(a) 齿轮箱泵进口压力拟合效果  
(a) The fitting result of the temperature of the



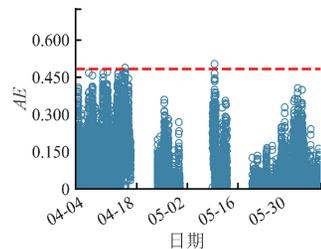
(b) 发电机转速拟合效果  
(b) The fitting result of the generator speed



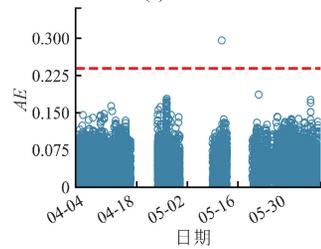
(c) 齿轮箱泵出口压力拟合效果  
(c) The fitting result of the gearbox pump outlet pressure

图 14 案例 1 机组正常行为建模分析结果

Fig. 14 Normal behavior modeling analysis results of the wind turbine for case 1



(a) 节点 1  
(a) Node 1



(b) 节点 3  
(b) Node 3

图 15 案例 1 机组节点级监测结果

Fig. 15 Node-level condition monitoring results of the wind turbine for case 1

通过监测风电机组节点可以看出,仅在 2017 年 5 月 9 日 12:45 发出预警,同一时刻风电机组状态监测图结

构中并不存在其余异常节点,无法构成故障传递链预警。当进行所有节点查询时,发现发电机转速节点在2017年5月11日13:20发出预警,同一时刻图结构中也不构成故障传递链,风电机组处于正常运行状态,这与风电机组SCADA系统的检测效果一致,验证了所提状态监测方法的有效性。

## 2) 健康机组误报案例分析

为了进一步验证节点级状态监测与故障传递链路在提高监测结果的准确性方面的作用,对健康机组的风电机组节点与发电机转速节点产生的误报情况进行分析。

由于节点1、3发生故障的时刻其余节点均无异常,因此选择这些时间点的发电机转速数据进行分析,结果如图16所示。在13:20左右风速突然产生较大波动导致节点3在该时刻超出阈值。虽然可以对发电机转速产生的误报进行解释,但对于节点1在5月9日12:45触发异常预警时并无其余节点产生异常情况,并且SCADA系统在此时也没有触发异常报警,因此很难对其异常行为进行解释。

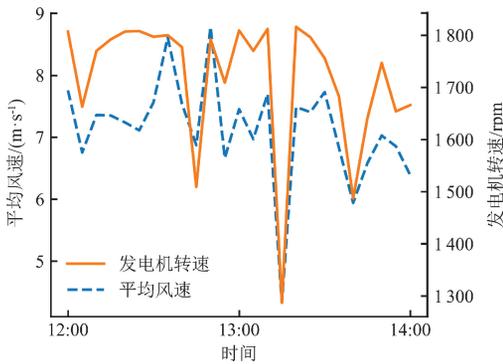


图16 案例1发电机转速原始数据故障时刻前后变化趋势

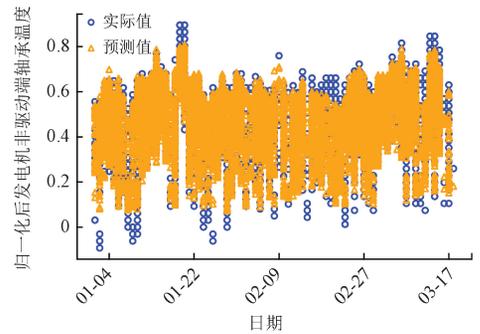
Fig. 16 The trend of the original data before and after the failure time of the generator speed for case 1

上述分析表明,单独的对某个节点或特征进行监测易受到环境以及动态工况的工作条件影响,产生难以解释的误报现象。该健康机组的监测结果也证明所提出的节点级状态监测策略能有效降低误报情况,增加监测的准确性和鲁棒性。

## 4.2 案例2-故障机组

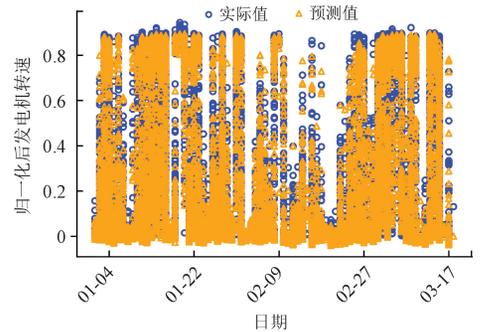
由于机组额定功率一致,且均为双馈式风电机组,机组物理结构和运行特性近似,因此风电机组图结构保持一致。根据风电机组图结构,结合数据预处理后的SCADA数据,生成案例2机组的风电时序图数据。在时序图数据生成后,基于PyTorch深度学习框架利用训练集时序图数据完成正常行为模型的训练,模型的学习率初始值为0.0001,其余超参数设置与案例1保持一致。如图17所示,部分节点的预测值与真实值对比结果表

明,模型能够有效捕捉SCADA数据的动态变化特征,验证了时序建模的有效性。



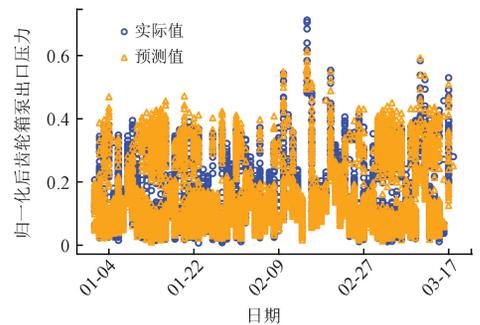
(a) 发电机非驱动端轴承温度拟合效果

(a) The fitting result of the temperature of the non-driven end bearing of the generator



(b) 发电机转速拟合效果

(b) The fitting result of the generator speed



(c) 齿轮箱泵出口压力拟合效果

(c) The fitting result of the gearbox pump outlet pressure

图17 案例2机组正常行为建模分析结果

Fig. 17 Node-level normal behavior modeling analysis results of the wind turbine for case 2

故障机组中存在异常结果节点如图18、19所示。通过图18、19的结果以及表4的各异常节点异常预警时间可以得出,节点1分别在在 $t_1$ 、 $t_4$ 、 $t_5$ 、 $t_6$ 时刻开始触发异常预警,其余时刻为图结构中其余节点开始出现异常预警时间。

在 $t_1$ 时刻,节点1、2、18均出现短暂的异常预警现象,而SCADA系统此时刻并未出现异常情况。针对这一现象,对此时刻的风电机组图结构进行分析。在产生异常的节点中,虽然没有传感器层的故障源头节点,但节

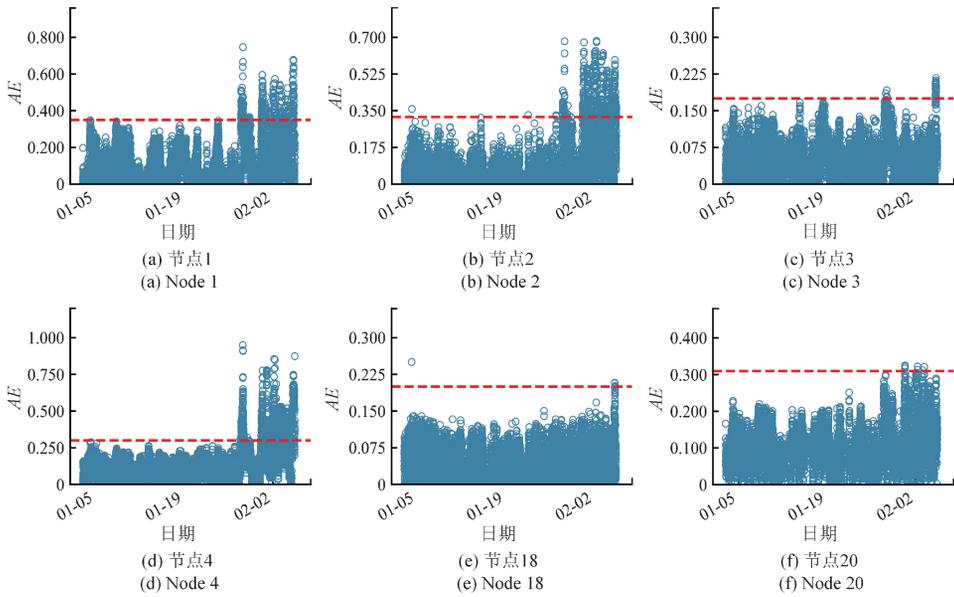


图 18 案例 2 机组节点级状态监测结果

Fig. 18 Node-level condition monitoring results of the wind turbine for case 2

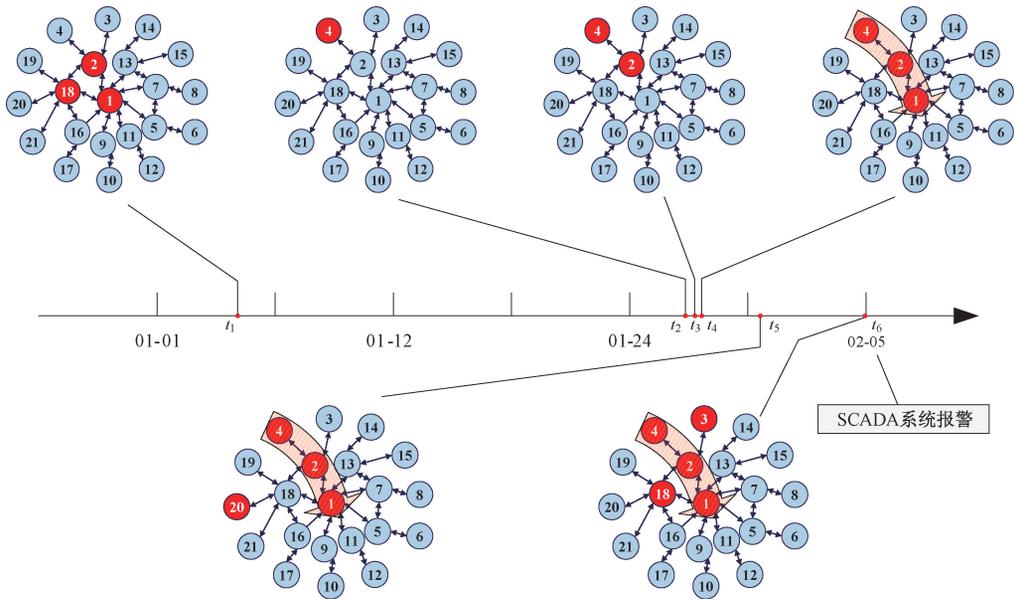


图 19 案例 2 机组节点级预警时间线及节点故障传递路径

Fig. 19 Node-level early warning timeline and node fault propagation path of the wind turbine for case 2

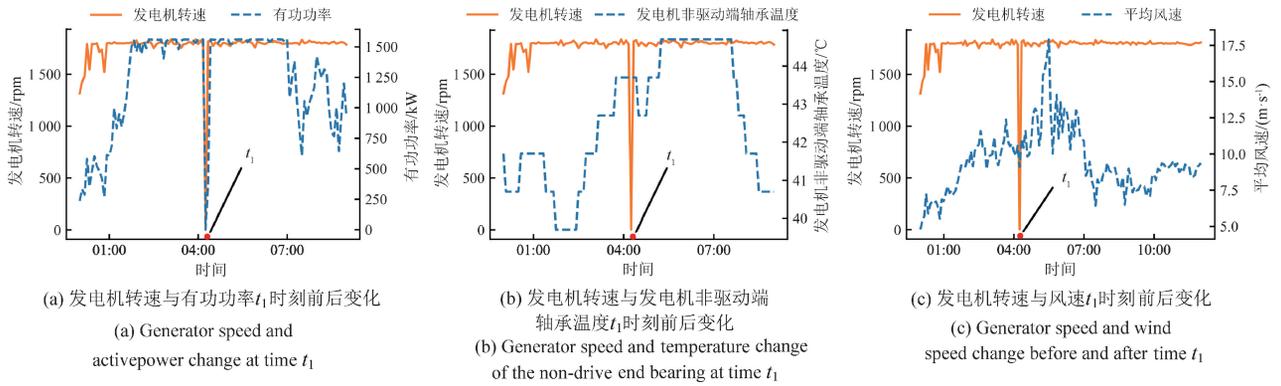
表 4 案例 2 机组节点故障预警时间

Table 4 Node failure warning time of the wind turbine for case 2

时刻符号	预警时间	预警节点
$t_1$	01-04T04:25	1, 2, 18
$t_2$	01-27T20:15	4
$t_3$	01-27T22:40	2, 4
$t_4$	01-27T23:00	1, 2, 4
$t_5$	01-31T07:40	1, 2, 4, 20
$t_6$	02-05T05:35	1, 2, 4, 18, 3

点 1、2、18 却在没有任何传感器产生异常的情况下发出了异常预警,不符合故障发生的规律。因此对与节点 2、18 相关的传感器原始数据进行可视化。发电机转速于  $t_1$  时刻出现骤降至 0 的情况,而发电机轴承温度在该时刻以及往后的一段时间内均为正常状态,如图 20 所示。

之后发电机有功功率与发电机转速同步骤降至 0,随着发电机转速恢复,有功功率也回归正常状态。通过对原始数据的分析发现,该机组在 1 月 4 日后随着风速的逐渐增加,有功功率也随之攀升至额定功率输出。

图20 案例2机组  $t_1$  时刻误报情况原始数据变化图Fig. 20 False alarm data change chart at time  $t_1$  of the wind turbine for case 2

因此,  $t_1$  时刻节点 1、2、18 产生异常报警的原因为: 由于风速在短时间内迅速上升, 发电机转速以及发电机有功功率均处于较高水平, 发电机非驱动端轴承温度逐渐上升, 导致风电机组运行负载较高, 触发机组对其输出功率进行调控, 减少运行负载进而保护机组, 因此后续时间段内非驱动端轴承温度逐渐下降。由于风电机组工作环境的复杂性, 节点 3 在训练集 (如图 17 (b) 所示) 中针对性地学习到了此类风电机组正常状态下机组策略引起的数据变化情况, 因此, 节点 3 在  $t_1$  时刻的绝对误差虽然超过其绝对误差均值, 但未超出阈值线引起异常预警。而节点 1、2、18 因为 KG-STGNN 的高阶信息聚合能力, 能接收到多个传感器数据信息, 在风电机组运行负载较高的状态下, 相关传感器节点在  $t_1$  时刻的绝对误差将处于超出绝对误差均值的水平, 因此节点 1、2、18 在该时刻产生短暂的异常预警, 出现误报。

继续对节点 1 进行监测, 发现节点 1 在  $t_4$  时刻之前均低于阈值, 之后迅速上升。因此对  $t_4$  时刻的图数据按照自顶向下的故障查询方法进行节点异常查询, 发现该时刻子系统层级的节点 2 与传感器层级的节点 4 也出现了异常预警现象, 通过后续时刻的节点级状态监测结果显示, 节点 1、2、4 在该时刻均产生异常预警, 形成故障查询链路。通过观测故障查询链路的逆向, 在  $t_4$  时刻开始与风电机组图结构相匹配, 形成故障传递链, 视为风电机组发出异常预警, 且故障源头或靠近故障源传感器为发电机非驱动端轴承温度。后续节点 20、18、3 发生异常预警的时刻, 节点 1、2、4 的故障传递链依旧存在。通过查询机组 SCADA 系统发现, 系统于 2017 年 2 月 5 日 11:20 发出了发电机非驱动端轴承温度过高的异常预警, 经运维工作人员检修后发现, 此次异常预警是由于发电机轴承发生跑圈故障所致。

上述结果的分析表明所提出的节点级状态监测策略

可以有效捕捉到风电机组的异常情况, 并提前 8 天发出预警。

### 4.3 模型性能分析

为了验证所提出方法的有效性, 采用相同的数据集, 阈值选取方法和深度学习参数, 将所提方法与非知识图谱构图方法及现有的模型方法如 CNN-LSTM<sup>[14]</sup>、GRU<sup>[41]</sup>、HGAT、GSTN 等在风电机组状态监测中的模型预测性能、误报率 (健康月份预警样本量与总健康数据样本量比值) 和异常预警能力 (最早正式异常预警时间) 方面的性能进行对比实验, 结果如表 5 所示。实验结果表明, 将风电文本类数据中故障传递信息与 SCADA 时序数据相结合的时空融合模型在 RMSE 与预警时间的关键指标上性能优于其他模型; 此外, 利用蕴含故障传递信息的图结构, 所建立的故障传递链误报验证机制在误报率指标上也显著优于对比模型。

表5 不同模型性能比较

Table 5 Comparison of performance across different models

模型	RMSE	误报率/%	预警时间
GSTN	0.071 3	10.30	1月29日
HGAT	0.072 4	13.60	1月27日
GRU	0.085 9	43.70	1月31日
CNN	0.090 1	46.90	2月2日
CNN-LSTM	0.081 7	21.50	1月31日
KG-STGNN	<b>0.067 7</b>	<b>5.63</b>	<b>1月27日</b>

为进一步验证所提出各模块在模型整体性能中的贡献, 进行了一系列消融实验。通过逐步移除模型中的关键组件, 构造不同的模型变体, 以量化各子模块对模型预测性能的具体影响, 实验结果如表 6 所示。

表 6 KG-STGNN 各模块消融实验

Table 6 Module-level ablation experiments for KG-STGNN

模型变体	RMSE	下降率/%
KG-STGNN	<b>0.067 7</b>	—
-KG	0.071 4	<b>5.47</b>
-Transformer	0.070 3	3.84
-HGAT	0.070 1	3.55
-多头注意力机制	0.069 1	2.07

从表中可以看出,完整模型 KG-STGNN 在预测任务中取得了最优的 RMSE。当移除知识图谱(-KG)辅助构图机制,改为采用传统全连接方式构图时,模型性能明显下降, RMSE 升至 0.071 4,性能下降了 5.47%,说明引入知识图谱有效增强了图结构的表达能力,对模型性能提升具有关键作用。仅采用 HGAT 进行空间特征提取、去除 Transformer 模块的变体(-Transformer) RMSE 为 0.070 3,性能下降 3.84%,表明 Transformer 模块能有效捕捉时序数据的时间特性,验证了时间特性建模模块的重要性。将高阶图注意力机制替换为传统 GAT(-HGAT)后,模型 RMSE 上升至 0.070 1(性能下降 3.55%),说明 HGAT 通过建模节点间多阶邻接关系,能够有效捕捉复杂的图拓扑信息,而传统 GAT 的单阶注意力机制难以表征此类高阶空间特征。此外,当模型中仅采用单头注意力机制时(-多头注意力机制),RMSE 升高至 0.069 1,性能下降 2.07%,表明多头注意力机制在提升模型表达能力方面具有一定正面作用。

综上所述,消融实验充分表明:知识图谱构图策略、Transformer 时间建模模块、高阶图注意力机制及多头注意力机制均对模型最终性能具有正向促进作用,验证了各模块设计的合理性与有效性。

#### 4.4 节点级监测策略误报抑制分析

在机组运行过程中,每次误报现象都会给运营单位带来经济损失。因此,减少误报情况可以减少线下人工维护次数,有效提升风电机组运维的可靠性。为量化所提节点级监测策略对风电机组误报的抑制效果,提出通过计算误报抑制率(即监测策略实施前后误报次数的减少量与总误报量的比值)进行评估,其中总误报量为未经监测策略时各监测节点超阈值预警的累计频次。由于节点级监测策略需要通过图结构进行分析,因此在对故障传递链的误报抑制效果进行分析时,均使用通过知识图谱构图的 GNN 模型进行对比。对同一风电机组某一健康运行月份进行状态监测,分析结果如表 7 所示。可以看出节点级监测策略在各 GNN 方法中误报抑制率均在 85% 以上,验证了该策略的有效性。

表 7 监测策略误报抑制分析

Table 7 Analysis of false alarm suppression in monitoring strategies

模型	总误报量	误报减少量	误报抑制率/%
KG-GSTN	78	69	88.5
KG-HGAT	116	101	87.1
KG-GAT	139	124	89.2
KG-STGNN	45	39	86.7

## 5 结 论

针对现有风电机组状态监测方法存在的输入特征相互独立、忽略输入特征之间潜在的关联,以及未充分运用风电机组文本类数据的问题,本研究提出了一种知识图谱与时空图网络融合驱动的风电机组状态监测方法。该方法首先通过文本类数据构建风电运维知识图谱,并通过查询得到风电机组图结构;继而结合 SCADA 时序数据生成风电时序图数据,实现文本数据在风电机组状态监测中的应用。在此基础上,利用 HGAT 聚合图结构中高阶邻域节点信息,结合 Transformer 注意力层捕捉各节点在时序图数据中的全局依赖关系,实现时间序列预测。同时通过节点级状态监测略分析各节点在异常时刻的图结构与故障传递链的匹配情况,从而确定预警时间、解释监测结果并降低误报率。实际应用结果表明,该方法可有效实现机组状态监测,相较传统监测策略能更早发现故障异常,并能通过其可解释性有效抑制误报警。

当前构建的知识图谱主要聚焦于风电机组运维领域,并且在构建的过程中部分方法依赖先验知识进行调控。未来可进一步拓展该知识图谱的适用范围,通过构建一个覆盖风电机组全领域的数据集并充分利用深度学习技术构建风电机组领域知识图谱库,以有效降低人为干预可能带来的图谱构建错误情况概率。这样不仅能针对不同型号的风电机组生成专属的监测图结构,还能高质量地反映各子系统及传感器之间的关联性,从而显著提升状态监测模型的监测能力和泛化能力。

## 参考文献

- [1] 金晓航,孙毅,单继宏,等. 风力发电机组故障诊断与预测技术研究综述[J]. 仪器仪表学报, 2017, 38(5): 1041-1053.  
JIN X H, SUN Y, SHAN J H, et al. Fault diagnosis and prognosis for wind turbines: An overview[J]. Chinese Journal of Scientific Instrument, 2017, 38(5): 1041-1053.
- [2] 刘军,安柏任,张维博,等. 大型风力发电机组健康

- 状态评价综述[J]. 电力系统保护与控制, 2023, 51(1): 176-187.
- LIU J, AN B R, ZHANG W B, et al. Review of health status evaluation of large wind turbines [J]. Power System Protection and Control, 2023, 51(1): 176-187.
- [ 3 ] KUSIAK A, VERMA A. A data-mining approach to monitoring wind turbines [J]. IEEE Transactions on Sustainable Energy, 2011, 3(1): 150-157.
- [ 4 ] 金晓航, 许壮伟, 孙毅, 等. 基于生成对抗网络的风电机组在线状态监测[J]. 仪器仪表学报, 2020, 41(4): 68-76.
- JIN X H, XU ZH W, SUN Y, et al. Online condition monitoring of wind turbine based on generative adversarial network[J]. Chinese Journal of Scientific Instrument, 2020, 41(4): 68-76.
- [ 5 ] HAN H Y, YANG D SH. Correlation analysis based relevant variable selection for wind turbine condition monitoring and fault diagnosis [J]. Sustainable Energy Technologies and Assessments, 2023, 60: 103439.
- [ 6 ] QU F M, LIU J H, ZHU H F, et al. Wind turbine fault detection based on expanded linguistic terms and rules using non-singleton fuzzy logic [J]. Applied Energy, 2020, 262: 114469.
- [ 7 ] 赵光权, 刘小勇, 姜泽东, 等. 基于深度学习的轴承健康因子无监督构建方法[J]. 仪器仪表学报, 2018, 39(6): 82-88.
- ZHAO G Q, LIU X Y, JIANG Z D, et al. Unsupervised health indicator of bearing based on deep learning[J]. Chinese Journal of Scientific Instrument, 2018, 39(6): 82-88.
- [ 8 ] 柏林, 晏锐, 刘小峰, 等. 基于多层深度互信息变分网络的风电机组轴承故障预警方法[J]. 太阳能学报, 2022, 43(11): 194-202.
- BO L, YAN R, LIU X F, et al. Wind turbine bearing fault early warning based on multi-layer depth mutual information variational network[J]. Acta Energetica Sinica, 2022, 43(11): 194-202.
- [ 9 ] CHEN H S, LIU H, CHU X N, et al. Anomaly detection and critical SCADA parameters identification for wind turbines based on LSTM-AE neural network[J]. Renewable Energy, 2021, 172: 829-840.
- [ 10 ] CHEN J SH, LI J, CHEN W G, et al. Anomaly detection for wind turbines based on the reconstruction of condition parameters using stacked denoising autoencoders[J]. Renewable Energy, 2020, 147: 1469-1480.
- [ 11 ] 吕游, 封烁, 郑茜, 等. 基于 LeNet5like 的迁移学习风电机组叶片覆冰故障诊断研究[J]. 仪器仪表学报, 2024, 45(3): 128-143.
- LYU Y, FENG SH, ZHENG X, et al. Research on fault diagnosis of wind turbine icing characteristics based on LeNet5like transfer learning [J]. Chinese Journal of Scientific Instrument, 2024, 45(3): 128-143.
- [ 12 ] 王梓齐, 刘长良, 刘帅. 基于集成 NSET 和模糊软聚类的风电机组齿轮箱状态监测[J]. 仪器仪表学报, 2019, 40(7): 138-146.
- WANG Z Q, LIU CH L, LIU SH, et al. Condition monitoring of wind turbine gearbox based on ensemble nonlinear state estimation technique and soft fuzzy clustering[J]. Chinese Journal of Scientific Instrument, 2019, 40(7): 138-146.
- [ 13 ] 金晓航, 许壮伟, 孙毅, 等. 基于 SCADA 数据分析和稀疏自编码神经网络的风电机组在线运行状态监测[J]. 太阳能学报, 2021, 42(6): 321-328.
- JIN X H, XU ZH W, SUN Y, et al. Online condition monitoring for wind turbine based on SCADA data analysis and sparse autoencoder neural network[J]. Acta Energetica Sinica, 2021, 42(6): 321-328.
- [ 14 ] XIANG L, YANG X, HU AI J, et al. Condition monitoring and anomaly detection of wind turbine based on cascaded and bidirectional deep learning networks [J]. Applied Energy, 2022, 305: 117925.
- [ 15 ] WANG Z K, JIANG X, XU Z F, et al. Early anomaly detection of wind turbine gearbox based on SLFormer neural network [J]. Ocean Engineering, 2024, 311: 118925.
- [ 16 ] TONG H J, QIU R C, ZHANG D X, et al. Detection and classification of transmission line transient faults based on graph convolutional neural network [J]. CSEE Journal of Power and Energy Systems, 2021, 7(3): 456-471.
- [ 17 ] LI CH Y, MO L F, YAN R Q. Fault diagnosis of rolling bearing based on WHVG and GCN [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 3519811.
- [ 18 ] SCARSELLI F, GORI M, TSOI A CH, et al. The graph neural network model [J]. IEEE Transactions on Neural Networks, 2009, 20(1): 61-80.
- [ 19 ] ZHAN P M, QIN X R, ZHANG Q, et al. A novel structural damage detection method via multisensor spatial-temporal graph-based features and deep graph convolutional network [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 2504814.
- [ 20 ] 郑毅, 王承民, 刘保良, 等. 基于多层级时空图神经网络的风电机组在线异常检测[J]. 电力系统自动化, 2024, 48(5): 107-119.
- ZHENG Y, WANG CH M, LIU B L, et al. Online

- anomaly detection of wind turbines based on hierarchical spatio-temporal graph neural network[J]. *Automation of Electric Power Systems*, 2024, 48(5): 107-119.
- [21] HE Z L, CHEN P F, LI X Y, et al. A spatiotemporal deep learning approach for unsupervised anomaly detection in cloud systems[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 34(4): 1705-1719.
- [22] HE L CH, BAI L, YANG X, et al. High-order graph attention network[J]. *Information Sciences*, 2023, 630: 222-234.
- [23] FENG CH L, LIU CH, JIANG D X. Unsupervised anomaly detection using graph neural networks integrated with physical-statistical feature fusion and local-global learning[J]. *Renewable Energy*, 2023, 206: 309-323.
- [24] JIN X H, LYU SH Y, KONG Z Q, et al. Graph spatio-temporal networks for condition monitoring of wind turbine[J]. *IEEE Transactions on Sustainable Energy*, 2024, 15(4): 2276-2286.
- [25] WANG J ZH, KOU M G, LI R Z, et al. Ultra-short-term wind power forecasting jointly driven by anomaly detection, clustering and graph convolutional recurrent neural networks[J]. *Advanced Engineering Informatics*, 2025, 65: 103137.
- [26] ZHANG Y M, ZHOU W Y, HUANG J CH, et al. Temporal knowledge graph informer network for remaining useful life prediction[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 3528610.
- [27] WU P CH, TU H B, MOU X, et al. An intelligent energy management method for the manufacturing systems using the knowledge graph and large language model[J/OL]. *Journal of Intelligent Manufacturing*, 1-12[2025-06-01].
- [28] 李晓璐, 王崑, 赵冰, 等. 面向风电机组运维数据的知识图谱构建方法[J]. *动力工程学报*, 2024, 44(6): 886-894.
- LI X L, WANG K, ZHAO B, et al. Knowledge graph construction method for wind turbine operation and maintenance data[J]. *Journal of Chinese Society of Power Engineering*, 2024, 44(6): 886-894.
- [29] 王萌, 王昊奋, 李博涵, 等. 新一代知识图谱关键技术综述[J]. *计算机研究与发展*, 2022, 59(9): 1947-1965.
- WANG M, WANG H F, LI B H, et al. Survey on key technologies of new generation knowledge graph[J]. *Journal of Computer Research and Development*, 2022, 59(9): 1947-1965.
- [30] TRISEHYA B D, QI J ZH, WANG W, et al. GCP: Graph encoder with content-planning for sentence generation from knowledge bases[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(11): 7521-7533.
- [31] 叶欣智, 尚磊, 董旭柱, 等. 面向配电网故障处置的知识图谱研究与应用[J]. *电网技术*, 2022, 46(10): 3739-3749.
- YE X ZH, SHANG L, DONG X ZH, et al. Knowledge graph for distribution network fault handling[J]. *Power System Technology*, 2022, 46(10): 3739-3749.
- [32] VERMA S, BHATIA R, HARIT S, et al. Scholarly knowledge graphs through structuring scholarly communication: A review[J]. *Complex & Intelligent Systems*, 2023, 9: 1059-1095.
- [33] 叶远波, 李端超, 谢民, 等. 基于知识图谱的二次设备测试自动配置方法[J]. *电力系统保护与控制*, 2022, 50(12): 162-171.
- YE Y B, LI D CH, XIE M, et al. Automatic configuration method of secondary equipment test based on a knowledge graph[J]. *Power System Protection and Control*, 2022, 50(12): 162-171.
- [34] SOUSA D, COUTO F M. Biomedical relation extraction with knowledge graph-based recommendations[J]. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26(8): 4207-4217.
- [35] 王军平, 张文生, 王勇飞, 等. 面向大数据领域的事理认知图谱构建与推断分析[J]. *中国科学: 信息科学*, 2020, 50(7): 988-1002.
- WANG J P, ZHANG W SH, WANG Y F, et al. Constructing and inferring event logic cognitive graph in the field of big data[J]. *Scientia Sinica Informationis*, 2020, 50(7): 988-1002.
- [36] 李刚, 李银强, 王洪涛, 等. 电力设备健康管理知识图谱: 基本概念、关键技术及研究进展[J]. *电力系统自动化*, 2022, 46(3): 1-13.
- LI G, LI Y Q, WANG H T, et al. Knowledge graph of power equipment health management: Basic concepts, key technologies and research progress[J]. *Automation of Electric Power Systems*, 2022, 46(3): 1-13.
- [37] 黄简, 杨程, 冯天波, 等. 面向风电机组运维的知识图谱构建研究与应用[J]. *电力系统保护与控制*, 2024, 52(8): 167-177.
- HUANG J, YANG CH, FENG T B, et al. Research and application of knowledge graph construction for wind turbine operation and maintenance[J]. *Power System Protection and Control*, 2024, 52(8): 167-177.
- [38] 冯艳红, 于红, 孙庚, 等. 基于BiLSTM的命名实体识别方法[J]. *计算机科学*, 2018, 45(2): 261-268.

FENG Y H, YU H, SUN G, et al. Named entity recognition method based on BiLSTM [J]. Computer Science, 2018, 45(2): 261-268.

- [39] 金晓航, 喻轩昂, 关汉林. 基于回声状态网络的风电机组运行状态监测[J]. 仪器仪表学报, 2025, 46(1): 258-269.

JIN X H, YU X ANG, GUAN H L. Condition monitoring of wind turbine based on echo state networks [J]. Chinese Journal of Scientific Instrument, 2025, 46(1): 258-269.

- [40] 王梓齐. 基于正常行为建模的风电机组状态监测方法研究[D]. 北京: 华北电力大学, 2022.

WANG Z Q. Research on condition monitoring method of wind turbine based on normal behavior modeling [D]. Beijing: North China Electric Power University, 2022.

- [41] 马永光, 冯勇升. 基于 ICEEMDAN-PCA-GRU 的风电机组齿轮箱故障预警方法研究[J]. 太阳能学报, 2023, 44(4): 67-73.

MA Y G, FENG Y SH. Research on fault early warning method of wind turbine gearbox based on ICEEMDAN-PCA-GRU [J]. Acta Energiæ Solaris Sinica, 2023, 44(4): 67-73.

## 作者简介



金晓航(通信作者), 2014年于香港城市大学获得博士学位, 现为浙江工业大学教授, 主要研究方向为工业大数据分析, 复杂装备的状态监测、故障诊断、寿命预测、智能运维。

E-mail: xhjin@zjut.edu.cn

**Jin Xiaohang** (Corresponding author) received his Ph. D. degree from City University of Hong Kong in 2014. He is currently a professor at Zhejiang University of Technology. His main research interests include industrial big data analysis, condition monitoring, fault diagnosis and prognosis, and intelligent maintenance for complex electro-mechanical equipment.



张元鸣, 2010年于日本宇都宫大学获得博士学位, 现为浙江工业大学计算机学院副教授, 主要研究方向为工业知识图谱、工业大模型和故障预测与健康管理等。

E-mail: zym@zjut.edu.cn

**Zhang Yuanming** received his Ph. D. degree from Utsunomiya University of Japan in 2010. He is currently an associate professor in the College of Computer Science and Technology at Zhejiang University of Technology. His main research interests include industrial knowledge graphs, industrial large models, prognosis and health management for equipment.



孔子迁, 2024年于浙江大学获得博士学位, 现为杭州电子科技大学特聘副研究员, 主要研究方向为工业时序大数据分析与复杂装备智能运维。

E-mail: kzq@hdu.edu.cn

**Kong Ziqian** received his Ph. D. degree from Zhejiang University in 2024. He is currently a specially appointed associate researcher at Hangzhou Dianzi University. His main research interests include industrial time-series big data analysis and intelligent operation and maintenance of complex equipment.



徐正国, 2009年于清华大学自动化系获得工学博士学位, 现为浙江大学控制学院教授, 主要研究方向为复杂工程系统的安全性运行理论与技术、工业系统大数据分析、智能感知元件与系统。

E-mail: xzg@zju.edu.cn

**Xu Zhengguo** received his Ph. D. degree from Tsinghua University in 2009. He is currently a professor in the College of Control Science and Engineering at Zhejiang University. His main research interests include safe theory and technology of complex engineering systems, industrial big data analysis, and intelligent sensing systems.