

# 注意力机制驱动的多源数据融合配网估计

邱桂华,汤志锐,陈宇婷

(南方电网广东佛山供电局,广东 佛山 528000)

**摘要:**针对配网中数据异构性与多源性挑战,提出一种基于编码-解码注意力机制自监督多源量测数据融合方法。该方法通过自监督学习自动捕捉数据间相关性,并利用编码和解码注意力机制提取加权融合特征,增强数据关联性、完整性与可用性,此方法能够自适应不同类型输入数据,进而确保在多源数据场景下实现高精度配网状态估计。在57节点仿真系统上开展的实验结果表明,所提方法在准确率、AUC和Macro\_F值等核心指标上均优于GraphMDN,RetNode,AdaAtt和DR-GCN等主流算法。其中,准确率达到88%,AUC提升至76.05%,Macro\_F值达到93.02%,整体性能显著提升。相较最优对比算法,平均误差降低47%,最大误差控制在0.017以内。结果验证了所提方法在多源融合、电网数据建模与状态估计中的有效性与泛化能力。

**关键词:**多源数据融合;编码-解码注意力;自监督学习;配网状态估计

中图分类号:TM28 文献标识码:A DOI:10.19457/j.1001-2095.dqed26514

## Distribution Network Estimation Driven By Attention Mechanism for Multi-source Data Fusion

QIU Guihua, TANG Zhirui, CHEN Yuting

(Guangdong Foshan Electric Power Bureau of China Southern Power Grid, Foshan 528000, Guangdong, China)

**Abstract:** Aiming at the challenges of data heterogeneity and multi-source in distribution networks, a self-supervised multi-source measurement data fusion method based on coding-decoding attention mechanism was proposed. This method automatically captured the correlation between data through self-supervised learning, and extracted weighted fusion features by encoding and decoding attention mechanisms to enhance the relevance, integrity and availability of data. This method can adapt to different types of input data, thus ensuring the realization of high-precision distribution network state estimation in multi-source data scenarios. Experimental results on a 57-node simulation system show that the proposed method outperforms mainstream algorithms such as GraphMDN, RetNode, AdaAtt and DR-GCN in terms of accuracy, AUC and Macro\_F value. Among them, the accuracy reached 88%, the AUC increased to 76.05%, the Macro\_F value reached 93.02%, and the overall performance was significantly improved. Compared with the optimal comparison algorithm, the average error is reduced by 47%, and the maximum error is controlled within 0.017. The results verify the effectiveness and generalization ability of the proposed method in multi-source fusion, power grid data modeling and state estimation.

**Key words:** multi-source data fusion; encoding-decoding attention; self-supervised learning; distribution network state estimation

配网状态估计作为能源管理系统的关键之一,广泛应用于监测电网的运行状况,辅助识别设备状态、负荷和发电机的功率信息<sup>[1-2]</sup>。传统状态估计方法主要包括最小二乘法、加权最小二乘法<sup>[3]</sup>、扩展卡尔曼滤波(exended Kalman filter, EKF)<sup>[4]</sup>和无迹卡尔曼滤波(unscented Kalman filter, UKF)<sup>[5]</sup>等。这些方法在一定程度上提升了配网状态估计的稳定性,但其模型结构通常依赖

于线性或准线性假设,难以充分建模配网中非线性、多模态和高维耦合特征,尤其在面对实际工况中的数据异构与数据稀疏问题时表现受限。

目前,配电网状态估计所依赖的数据主要来源于SCADA和WAMS系统,但二者在采样频率、数据精度和信息维度上存在显著差异。SCADA系统虽然覆盖面广,但数据时效性和测量精度较低;WAMS系统具备高频采样能力,但安装成本

基金项目:南方电网公司科技项目(GDKJXM20240450)

作者简介:邱桂华(1987—),男,硕士,正高级工程师,主要研究方向为配电网调度运行,Email:13356959297@163.com

高,布点稀疏,难以全面感知配电网状态<sup>[6]</sup>。因此,如何在保持系统整体可观测性的基础上,实现对 SCADA 与 WAMS 多源异构量测数据的高效融合,成为当前研究关注的重点方向之一。

多源量测数据融合是指将来自不同来源、格式、维度的数据统一整合,以构建更完整和一致的特征表示<sup>[7]</sup>。然而,受限于数据异构性、采样频率不一致、冗余结构和语义偏差等问题,传统融合方法(如拼接、加权、统计规则)在实际应用中面临鲁棒性差、信息损失大、模型过拟合等挑战<sup>[8]</sup>。近年来,随着深度学习的发展,研究者尝试引入图神经网络(GCN)<sup>[9]</sup>、自监督学习<sup>[10]</sup>和注意力机制<sup>[11]</sup>来构建更加灵活、泛化能力更强的融合框架,用以建模多源数据间复杂的时空依赖关系,提升状态估计的准确性与鲁棒性。

其中,编码-解码注意力机制(encoder-decoder attention)作为一种在自然语言处理任务中广泛应用的架构,具备强大的序列建模与特征融合能力。通过编码器提取全局特征,解码器结合注意力权重聚焦输入的关键部分,有效提升了对长距离依赖与异构特征的表示能力<sup>[12-13]</sup>。在配电网多源数据处理场景中,具备天然优势,尤其适合建模 WAMS 与 SCADA 系统间的动态关联特征。

基于此,本文提出一种基于编码-解码注意力机制的自监督多源数据融合配网状态估计方法。通过编码-解码结构构建时空上下文依赖关系,并引入自监督策略以增强在弱标注数据场景下的鲁棒性与泛化能力。实验验证表明,本文方法在多个评估指标下均优于现有主流算法,能够显著提升配电网运行状态估计精度与稳定性。

## 1 加权最小二乘状态估计模型

考虑多源量测数据的加权最小二乘法状态估计目标函数  $J(x)$  可表达为<sup>[14]</sup>

$$\min J(x) = J^P(x) + J^{SA}(x) \quad (1)$$

其中

$$\begin{cases} J^P(x) = [z^P - h^P(x)]^T (R^P)^{-1} [z^P - h^P(x)] \\ J^{SA}(x) = [z^{SA} - h^{SA}(x)]^T (R^{SA})^{-1} [z^{SA} - h^{SA}(x)] \\ \text{s.t. } c(x) = 0 \end{cases}$$

式中: $x$  为状态变量; $J^P(x), J^{SA}(x)$  分别为 D-PMU、传统量测目标函数; $z^P, z^{SA}$  为量测值; $h^P(x), h^{SA}(x)$  为量测函数; $R^P, R^{SA}$  为误差协方差矩阵; $c(x)$  为零注入约束。

构建的增广量测模型可表示为

$$G = \begin{bmatrix} H^T R^{-1} H & C^T \\ C & 0 \end{bmatrix} \quad (2)$$

其中  $H = [(H^P)^T (H^{SA})^T]^T$   
 $R = \text{diag}(R^P, R^{SA})$

式中: $H$  为雅克比矩阵; $R$  为误差协方差矩阵; $C$  为零注入约束雅克比矩阵。

引入拉格朗日函数,式(1)所述的状态估计模型转化为

$$L(x, \lambda) = \frac{1}{2} J(x) - \lambda^T c(x) \quad (3)$$

式中: $\lambda$  为拉格朗日乘子。

采用高斯-牛顿法形成如下迭代方程:

$$\begin{bmatrix} H^T R^{-1} H & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} \Delta x_r \\ -\lambda \end{bmatrix} = \begin{bmatrix} H^T R^{-1} \Delta z_r \\ -c(x_r) \end{bmatrix} \quad (4)$$

其中  $\Delta z_r = z - h(x_r)$

式中: $r$  为迭代次数; $x_r, \Delta x_r$  分别为第  $r$  次迭代状态变量和变化量; $z$  为量测值。

## 2 多源数据融合算法

### 2.1 算法模型结构

所提算法设计了一种基于编码和解码注意力机制的自监督多源量测数据融合方法,用于配网状态估计。该模型包括多个输入层、一个编码层、一个解码注意力层和一个自监督层<sup>[15]</sup>。其中,编码和解码注意力层是模型的核心,能够有效地挖掘多源异构数据中的显著信息,算法模型结构流程如图 1 所示。该模型的输入数据类型包括 WAMS 与 SCADA 系统量测数据。

### 2.2 多源异构数据融合

本文设计的多源异构特征融合网络如图 2 所示。结构上,采用多分支分层融合策略。技术上,采用基于门控机制的特征融合方法。

为实现多源异构信息的联合建模,本文引入门控机制进行特征融合,其计算过程如下:

$$G_j^F, G_j^R, G_j^T = \delta [M_{\theta_j}(F_j^t, R_j^t, T_j^t)] \quad (5)$$

$$F_{j+1}^t = G_j^F \otimes F_j^t + G_j^R \otimes R_j^t + G_j^T \otimes T_j^t \quad (6)$$

式中: $G_j^R, G_j^T$  为门控权重; $G_j^F$  为多源融合特征门控权重; $\delta$  为门控激活函数; $\otimes$  表示乘积; $M_{\theta_j}$  为门控生成网络。

### 2.3 编码与解码注意力

通过多源异构融合网络提取的空间特征,利用 Transformer 编解码器<sup>[16]</sup>进一步学习数据间的时空依赖关系。注意力机制通过计算查询与键的

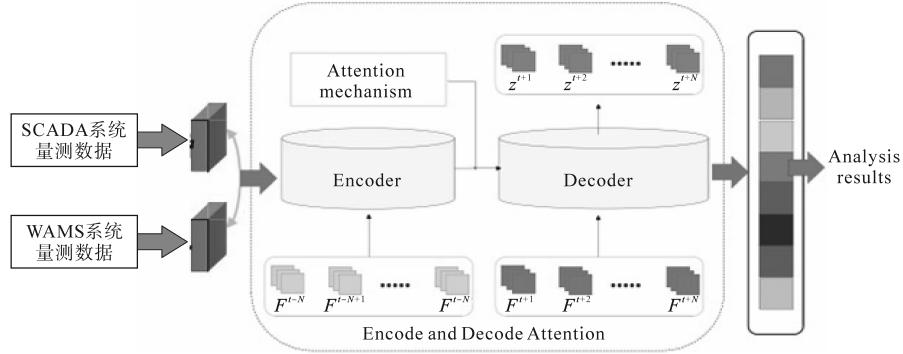
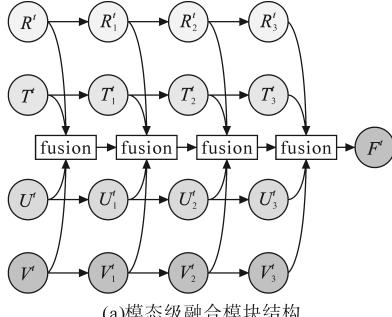
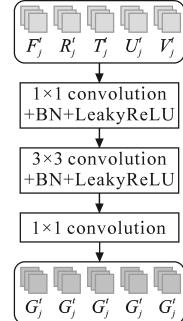


图1 模型结构流程图

Fig.1 Model structure flowchart



(a) 模态级融合模块结构



(b) 融合特征的卷积处理流程

图2 多模态特征融合网络结构示意图

Fig.2 Schematic diagram of the multi-modal feature fusion network structure

相似度以分配注意力权重,并对值向量进行加权求和以生成输出,具体计算如下:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V} \quad (7)$$

当  $\mathbf{Q}, \mathbf{K}$  和  $\mathbf{V}$  相等时,称之为自注意力机制。

为捕获多维特征关联,采用多头注意力机制,其计算如下:

$$\text{MultiHeadAttention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(h_1, h_2, \dots, h_k)\mathbf{W}^o \quad (8)$$

其中

$$h_i = \text{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V) \quad (9)$$

式中:  $\mathbf{W}_i^Q, \mathbf{W}_i^K, \mathbf{W}_i^V$  分别为查询、键和值线性转换的参数矩阵;  $\mathbf{W}^o$  为多头注意力机制最终线性转换的参数矩阵<sup>[17]</sup>。

图3展示了本文设计的轴向注意力机制,通过并行执行时空轴向注意力以提升训练与推理速度。从计算角度看,注意力操作在三个维度上独立执行,显著降低了时空特征图中每个位置的计算复杂度至  $O(T \times H \times W)$ 。同时,经过注意力操作后的输出仍然保留全局感知能力,轴向注意力的输出  $z_{ij,t}$  计算公式如下<sup>[17]</sup>:

$$z_{ij,t}^1 = \text{softmax}\left(\frac{\mathbf{q}_{ij,t}\mathbf{K}_1^T}{\sqrt{d}}\right)\mathbf{V}_1 \quad (10)$$

$$z_{ij,t}^2 = \text{softmax}\left(\frac{\mathbf{q}_{ij,t}\mathbf{K}_2^T}{\sqrt{d}}\right)\mathbf{V}_2 \quad (11)$$

$$z_{ij,t}^3 = \text{softmax}\left(\frac{\mathbf{q}_{ij,t}\mathbf{K}_3^T}{\sqrt{d}}\right)\mathbf{V}_3 \quad (12)$$

$$z_{ij,t} = z_{ij,t}^1 + z_{ij,t}^2 + z_{ij,t}^3 \quad (13)$$

式中:  $\mathbf{q}_{ij,t}$  为查询向量;  $\mathbf{K}_1, \mathbf{V}_1 \in \mathbf{R}_{H \times d}$  为垂直方向键值矩阵;  $\mathbf{K}_2, \mathbf{V}_2 \in \mathbf{R}_{W \times d}$  为水平方向键值矩阵;  $\mathbf{K}_3, \mathbf{V}_3 \in \mathbf{R}_{T \times d}$  为时间方向键值矩阵。

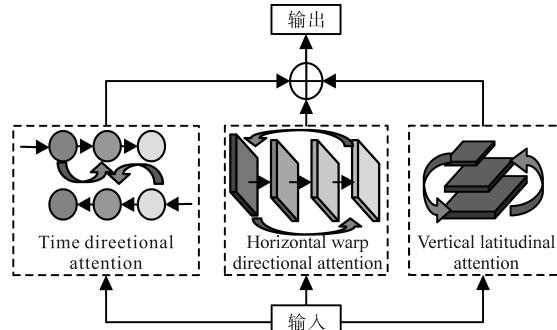


图3 注意机制结构

Fig.3 Attention mechanism structure

编码器由多个相同的层堆叠而成,如图4所示。结构上,卷积前馈网络由两层卷积函数和 Leaky ReLU 激活函数组成。为了高效连接这些残差模块,所有子层的输出维度设置  $d$ 。其中,图4中的空间-时间轴向注意力基于图3中的注意力机制。具体公式如下:

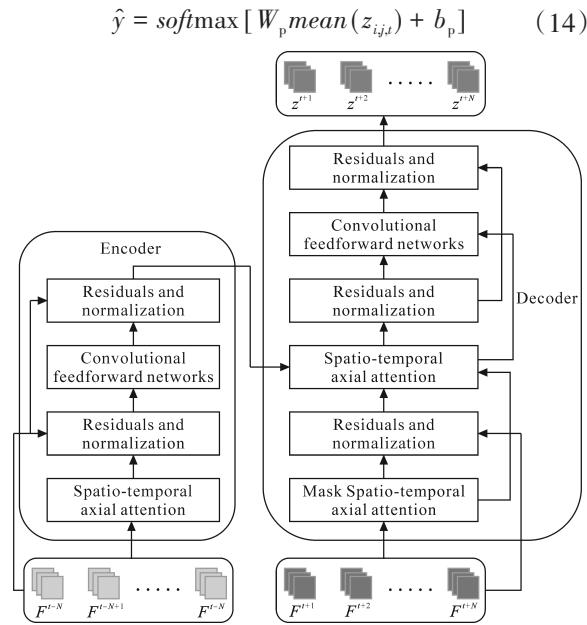


图4 编码器和解码器架构  
Fig.4 Encoder and decoder architecture

## 2.4 自监督优化

损失函数包含三个部分:分类 $L_{CE}$ 、语义约束 $L_{SCL}$ 和自监督对比 $L_{SSL}$ ,如下式所示:

$$L = \lambda_1 L_{CE} + \lambda_2 L_{SCL} + \lambda_3 L_{SSL} \quad (15)$$

式中: $\lambda_1, \lambda_2$ 和 $\lambda_3$ 为平衡参数,用于控制不同损失的影响。

在本文中,使用交叉熵损失函数来计算分类损失,如下式所示:

$$L_{CE} = -\sum_{i=1}^L \sum_{j=1}^m y_j \ln \hat{y}_j \quad (16)$$

式中: $y, \hat{y}$ 分别为真实概率和预测概率。

在进行多源异构数据融合时,需确保语义一致性,即多源异构数据融合不会导致预测标签发生较大变化。为此,设计了语义约束损失来衡量不同源数据与融合数据之间的语义接近程度:

$$L_{SCL} = \frac{1}{V} \sum_{i=1}^L W(\hat{y}, y) \quad (17)$$

多源异构数据经过特征融合模块和编码解码模块后得到多元数据。第 $i$ 个时刻的嵌入表示用 $z$ 表示,即第 $i$ 个时刻嵌入矩阵中对应的行 $Z_u^{(i)}$ 。在本文中,将同一来源在不同时刻的表示视为正样本对,不同来源在不同时刻的表示视为负样本对。通过最大化正样本对的相似度并最小化负样本对的相似度,得到自监督对比损失公式如下:

$$L_{SSL} = \frac{1}{N \cdot V} \sum_{u \in V} \sum_{i,j}^V L(Z_u^{(i)}, Z_u^{(j)}) \quad (18)$$

## 3 示例与仿真分析

### 3.1 实验设置

本文参照文献[6],利用matpower仿真生成SCADA和WAMS系统样本数据库,数据库详情如表1所示,SCADA数据种类包括:线路功率、节点注入功率、节点电压幅值及线路电流幅值。WAMS数据种类包括:节点电压相量、线路电流相量、节点注入电流相量。其中PMU安装节点分布如图5所示<sup>[15]</sup>。算法参数如表2所示<sup>[15]</sup>。

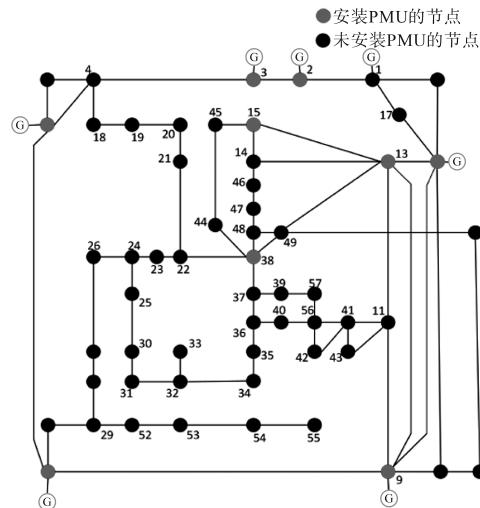


图5 57节点系统  
Fig.5 57-node system

表1 数据库详情  
Tab.1 Database details

参数名称	SCADA	WAMS
样本数据	1 000	1 000
数据种类	4	3
训练集样本	700	700
验证集样本	200	200
测试集样本	100	100

表2 算法参数设置  
Tab.2 Algorithm parameter settings

参数名称	数值	参数名称	数值
嵌入维度	1 024	批量大小	64
隐藏层维度	512	优化功能	Adam
学习速率	0.001	注意网络层数	2
迭代次数	100	—	—

### 3.2 算法性能对比

#### 3.2.1 评价指标构建

为测试本文所提算法性能,与GraphMDN, RetNode, AdaAtt 和 DR-GCN 四种自监督算法进行对比分析。并构建以下三个评估指标:准确率

(Accuracy)、曲线下面积(AUC)和宏平均F值(Macro\_F)。准确率衡量模型正确分类的样本数占总样本数的比例,定义如下:

$$Accuracy = \frac{\sum_{i=1}^N TP_i}{\sum_{i=1}^N (TP_i + TN_i + FP_i + FN_i)} \quad (19)$$

式中: $TP_i$ 、 $TN_i$ 、 $FP_i$ 和 $FN_i$ 分别为类别*i*的真正例数、真反例数、假正例数和假反例数;N为测试样本的总数。

AUC是ROC曲线(受试者工作特征曲线)下的面积,用于评估模型区分正负样本的能力。计算如下:

$$AUC = \int_0^1 TPR(FPR) dFPR \quad (20)$$

$$TPR = \frac{TP}{FN + TP} \quad (21)$$

$$FPR = \frac{FP}{TN + FP} \quad (22)$$

对于多分类问题,AUC需要对每个类别计算AUC值,然后取无权平均(Macro\_AUC):

$$Macro\_AUC = \frac{1}{K} \sum_{i=1}^K AUC_i \quad (23)$$

式中: $K$ 为类别总数; $AUC_i$ 为第*i*类的AUC值。

宏平均F值是精确率(Precision)和召回率(Recall)的调和平均:

$$F = 2 \times \frac{Precision + Recall}{Precision \times Recall} \quad (24)$$

其中  $Precision = \frac{TP}{FP + TP} \quad (25)$

$$Recall = \frac{TP}{FN + TP} \quad (26)$$

对于多分类问题,宏平均F计算方法如下:

$$Macro\_F = \frac{1}{K} \sum_{i=1}^K F1_i \quad (27)$$

式中: $F1_i$ 为针对类别*i*计算的F1值。

### 3.2.2 算法收敛性能分析

为进一步说明本文所提方法的优越性,在本次实验中,使用训练步数(或迭代次数)作为横坐标,使用不同算法的收敛值(准确率)作为纵坐标,进行对比分析。不同模型训练收敛过程对比分析结果如图6所示。

从图6可以看出,所提方法在80轮左右收敛,并达到最高值88,整体最优。GraphMDN收敛最慢,最终效果最差(83),适用性较低。RetNode和AdaAtt收敛速度适中,最终值略低于本文方

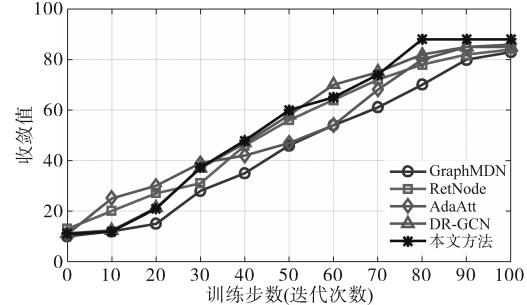


图6 模型训练收敛过程

Fig.6 Model training convergence process

法,仍然是较优方法。DR-GCN表现稳定,收敛值接近85,但仍然不及本文方法。

### 3.2.3 算法参数影响分析

为了探究方程自监督优化参数 $\lambda_1$ 、 $\lambda_2$ 和 $\lambda_3$ 对分类结果的影响,其结果如图7所示。在实验中,固定两个平衡参数为最优设置,而另一个平衡参数在0至1的范围内,以0.2为步长取不同值,并验证其性能。

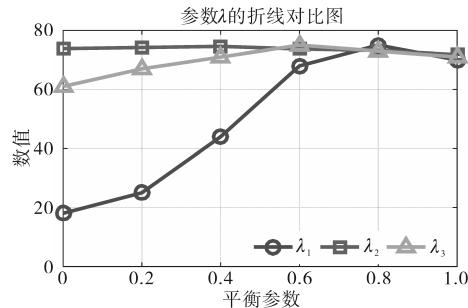


图7 不同参数影响效果

Fig.7 Effect of different parameters

由图7可知, $\lambda_1$ 对性能的影响最为显著,其次是 $\lambda_2$ 。当 $\lambda_1$ 从0开始增大时,性能会有所提升。当 $\lambda_1$ 达到0.8时,性能达到最优,随后则呈下降趋势。 $\lambda_2$ 整体变化不大,在74左右波动,在参数0.4时表现最佳。 $\lambda_3$ 从0到0.6逐渐增加,在0.6处达到峰值(75),随后开始下降,但下降幅度较小,说明其对参数变化的稳定性较好。

### 3.2.4 算法数据处理能力分析

不同算法对SCADA和WAMS系统样本数据分析结果分别如图8所示,其中图8a为SCADA数据处理分析结果,图8b为WAMS数据处理分析结果。

从图8可以看出,本文方法在Accuracy、AUC和Macro\_F三项指标上均优于其他算法,整体表现最优,充分体现了模型的特征提取与泛化能力。

在图8a高密度SCADA数据中,本文方法的

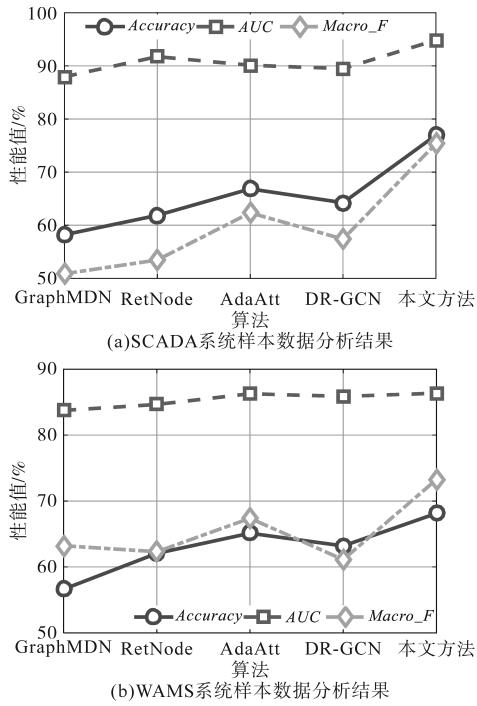


图8 不同算法数据处理效果

Fig.8 Data processing effects of different algorithms

Accuracy 约为 78%, 较 AdaAtt(68%) 提升约 10 个百分点, 较 GraphMDN(59%) 提升约 19 个百分点; AUC 稳定在 96% 左右, Macro\_F 约 78%, 验证了模型在稠密数据场景下的高精度与稳定性。

在图 8b 稀疏异构 WAMS 数据中, 本文方法的 Accuracy 约 70%, 较 AdaAtt(67%) 提升约 3 个百分点; AUC 达 88%, Macro\_F 为 73%, 均为最高值, 表明模型在弱特征与样本不均条件下仍保持稳健表现。

总体来看, 本文方法在不同类型数据中均展现出显著的鲁棒性与适应性。这主要得益于模型中引入的自监督机制与注意力结构, 能够深度挖掘多源数据的内在关联并自动补足监督信息空缺, 尤其在小样本与少数类表示学习场景中表现出明显优势。

### 3.3 不同融合策略的评估

为进一步说明多源数据融合网络的有效性, 对比分析了四种融合策略:

策略 1: 在输入层将异构特征拼接, 并利用相同的卷积网络进行特征提取, 该方法通常应用于深度模型中融合多源数据。

策略 2: 输出拼接, 每个多源数据特征由独立网络分支提取, 最终通过  $1 \times 1$  卷积融合各数据源的特征。

策略 3: 卷积与分层融合——采用分层特征

融合策略, 并使用  $1 \times 1$  卷积替代门控机制进行特征融合。

策略 4: 门控与分层融合(本文方法)——结合门控机制与分层融合策略, 以调控不同数据源对整体特征的贡献。不同特征融合策略结果如表 3 所示。

表3 不同策略比较

Tab.3 Different strategy comparisons

样本数据	策略	Accuracy/%	AUC/%	Macro_F/%
SCADA	策略 1	70.85	68.92	90.12
	策略 2	73.10	71.25	91.75
	策略 3	75.05	73.80	93.02
	策略 4	77.90	76.05	93.02
WAMS	策略 1	60.55	53.24	81.54
	策略 2	62.57	54.11	83.80
	策略 3	66.93	57.92	86.21
	策略 4	68.81	59.85	87.95

实验结果分析如下:

1) SCADA 数据分析。从表 3 中可见, 在 SCADA 数据场景下: 策略 1(输入拼接)在 Accuracy, AUC 和 Macro\_F 上分别为 70.85%, 68.92% 和 90.12%, 性能最低。策略 2(输出拼接)相比策略 1, 在 Accuracy 上提升了 2.25 个百分点, AUC 提升 2.33 个百分点, Macro\_F 提升 1.63 个百分点。策略 3(卷积与分层融合)进一步提高, Accuracy 达到 75.05%, 较策略 1 提升 4.2 个百分点。本文提出的策略 4(门控与分层融合)在 Accuracy 上达到 77.90%, 较策略 3 提升 2.85 个百分点, 在 AUC 上提升 2.25 个百分点, 在 Macro\_F 上与策略 3 持平(均为 93.02%), 为所有策略中的最优结果。

2) WAMS 数据分析。在 WAMS 数据场景下, 策略性能差距更为明显: 策略 1 各项指标均为最低, Accuracy 仅为 60.55%, AUC 为 53.24%, Macro\_F 为 81.54%。策略 2 相比策略 1, 在 Accuracy 上提升 2.02 个百分点, Macro\_F 提升 2.26 个百分点。策略 3 显著优于策略 1 和 2, 其中 Accuracy 提升至 66.93%, 较策略 1 提升 6.38 个百分点, AUC 提升 4.68 个百分点。本文所提策略 Accuracy 达到 68.81%, 较策略 3 提升 1.88 个百分点, AUC 提升 1.93 个百分点, Macro\_F 达到 87.95%, 为全场最佳。

从整体上看, 本文提出的门控与分层融合策略(策略 4)在两种数据场景下均取得最高指标值, 证明其在多源数据的特征提取和融合方面具备更优的效果。相比传统拼接类策略, 所提方法的平均准确率提升约 7%, AUC 平均提升超 6%,

且 *Macro\_F* 始终处于领先位置,体现出其在提高分类判别力、保持语义一致性方面的有效性。

### 3.4 状态评估效果分析

上述算例分析,有效验证了本文所提方法在数据融合方面的优越性。在本节中,为验证本文所提方法在配网运行状态估计中的有效性,通过与最小加权二乘法(WLS)、全连接神经网络(DNN)和图神经(GCN)<sup>[6]</sup>三种算法进行对比,结果如图9所示。

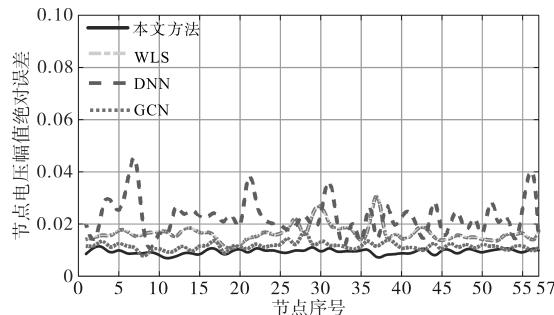


图9 不同算法状态评估效果

Fig.9 State evaluation effects of different algorithms

图9中展示了本文方法与三种主流算法(WLS, DNN, GCN)在57个节点上的节点电压幅值相对误差分布情况。从整体趋势来看,本文方法在各节点上的误差始终维持在0.012以下,曲线最为平稳,体现出优异的估计精度与鲁棒性。相比之下,DNN误差波动最明显,峰值接近0.05,在多源异构特征融合下易出现过拟合;WLS误差整体位于0.015~0.03区间,虽较稳定但受线性假设限制,难以刻画复杂关联;GCN表现相对平稳,但部分节点误差仍超过0.015,总体精度不及本文方法。

从表4中可以看出,本文方法在平均误差和最大误差两项指标上均明显优于其他方法。与DNN相比,平均误差降低了约73.5%,最大误差降低了超过80%;即使与性能相对较好的GCN算法相比,本文方法的平均误差也低了近47%,最大误差降低了47%左右。

表4 不同算法误差数据提取均值与最大值

Tab.4 Extract mean and maximum values from error data of different algorithms

方法	平均误差	最大误差
本文方法	0.009	0.017
GCN	0.017	0.032
WLS	0.021	0.045
DNN	0.034	0.088

本文方法的优越性在于结合了编码-解码注意力机制与自监督学习框架。注意力机制可根据数据相关性动态调整SCADA与WAMS等多源数据的权重,有效过滤冗余信息;编码-解码结构强化了对时空依赖的建模,自监督学习则通过数据自身生成训练信号,无需大量标注,提升了泛化性与稳定性。综合应用上述机制,显著降低了估计误差,增强了配网状态估计的鲁棒性。

## 4 结论

本文提出一种面向配电网状态估计的自监督多源数据融合方法,针对多源异构数据融合难题,融合了编码-解码注意力机制与自监督学习。主要结论如下:

方法设计方面:引入编码-解码注意力机制,有效提升多源数据加权融合能力。结合自监督学习,降低对标注数据的依赖,增强模型的泛化能力。

收敛性与鲁棒性验证:所提方法在80轮内快速收敛,准确率达到88%,优于对比算法。在不同数据扰动下保持稳定性能,体现出了良好的鲁棒性。

数据处理能力分析:在SCADA和WAMS数据处理中,准确率分别提升4%和25%以上。*AUC*和*Macro\_F*值全面领先,适应稀疏与冗余数据场景。

融合策略评估结果:策略4(门控与分层融合)在*Accuracy*, *AUC*, *Macro\_F*上表现最优。相比传统拼接策略,平均提升准确率约7%, *AUC*提升超6%。

状态估计误差分析:在57节点系统中,平均误差降至0.009,最大误差控制在0.017。相比DNN,误差降低超73%;相较GCN,误差降低近47%。

综上,本文方法在配网多源异构数据环境下具备优良的估计精度与稳定性,适合推广应用于实际配网运行状态评估任务中。

## 参考文献

- [1] 蒋文辉,林冬,于力,等.基于状态估计方法的相量测量单元电压稳定性研究[J].电气传动,2023,53(11):53-59.  
JIANG Wenhui, LIN Dong, YU Li, et al. Research on voltage stability of phasor measuring unit based on state estimation method[J]. Electric Drive, 2023, 53 (11): 53-59.

- [2] YUAN Q, ZHOU K, YAO J. A new measure of wind power variability with implications for the optimal sizing of standalone wind power systems[J]. *Renewable Energy*, 2020, 150: 538-549.
- [3] 满延露, 刘敏. 基于WLS-AUKF混合算法的主动配电网联合状态估计[J]. *电子科技*, 2025, 38 (2): 93-102.
- MAN Yanlu, LIU Min. Joint state estimation of active distribution network based on WLS-AUKF hybrid algorithm[J]. *Electronic Science and Technology*, 2025, 38 (2): 93-102.
- [4] 董磊, 赖纪东, 苏建徽, 等. 基于IMAFFRLS-EKF的锂电池在线参数辨识和SOC估计方法[J]. *太阳能学报*, 2024, 45 (6): 66-74.
- DONG Lei, LAI Jidong, SU Jianhui, et al. On-line parameter identification and SOC estimation method of lithium battery based on IMAFFLS-EKF[J]. *Journal of Solar Energy*, 2024, 45 (6): 66-74.
- [5] 李锦满, 李儒欢, 李浩南, 等. 基于无迹卡尔曼滤波的动力电池状态估计 [J]. *电池*, 2024, 54 (3): 340-343.
- LI Jinman, LI Ruhuan, LI Haonan, et al. State estimation of power battery based on unscented Kalman filter[J]. *Chinese Journal of Batteries*, 2024, 54 (3): 340-343.
- [6] 吴兆宇. 考虑SCADA和WAMS数据融合的图神经网络暂态稳定评估方法[D]. 南京: 东南大学, 2023.
- WU Zhaoyu. Transient stability evaluation method of graph neural networks considering SCADA and WAMS data fusion[D]. Nanjing: Southeast University, 2023.
- [7] ZHANG Pengfei, LI Tianrui, WANG Guoqiang, et al. Multi-source information fusion based on rough set theory: a review [J]. *Information Fusion*, 2021, 68: 85-117.
- [8] YAN J, HU Y, GUO C. Rotor unbalance fault diagnosis using DBN based on multi-source heterogeneous information fusion [J]. *Procedia Manufacturing*, 2019, 35: 1184-1189.
- [9] SCARSELLI F, GORI M, TSOI A C, et al. The graph neural network model[J]. *IEEE Transactions on Neural Networks*, 2008, 20(1): 61-80.
- [10] GAO W, ZHAO Z. Self-supervised multi-source heterogeneous data fusion using encode and decode attention for intelligent medical device communication analysis[J]. *IEEE Trans. on Consumer Electronics*, 2024, 70(1): 1318-1325.
- [11] CHEFER H, GUR S, WOLF L. Generic attention-model explainability for interpreting bi-modal and encoder-decoder transformers[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, Canada: IEEE, 2021: 397-406.
- [12] FENG L, ZHAO C, SUN Y. Dual attention-based encoder-decoder: a customized sequence-to-sequence learning for soft sensor development[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(8): 3306-3317.
- [13] ZHU K, LI Y, MAO W, et al. LSTM enhanced by dual attention-based encoder-decoder for daily peak load forecasting[J]. *Electric Power Systems Research*, 2022, 208: 107860.
- [14] 刘喆林. 考虑多源量测数据的数字配电网状态估计方法与应用[D]. 天津: 天津大学, 2022.
- LIU Z L. State estimation method and application of digital distribution network considering multi-source measurement data [D]. Tianjin: Tianjin University, 2022.
- [15] XIONG Z, MA J, CHEN B, et al. Multi-source data recognition and fusion algorithm based on a two-layer genetic algorithm-back propagation model[J]. *Frontiers in Big Data*, 2025, 7: 1520605.
- [16] 向德萍, 张普, 向世明, 等. 基于Transformer的多模态气象预测[J]. *计算机工程与应用*, 2023, 59 (10): 94-103.
- XIANG Deping, ZHANG Pu, XIANG Shiming, et al. Multi-modal weather prediction based on transformer[J]. *Computer Engineering and Applications*, 2023, 59 (10): 94-103.
- [17] ZHU K, LI Y, MAO W, et al. LSTM enhanced by dualattention-based encoder-decoder for daily peak load forecasting[J]. *Electric Power Systems Research*, 2022, 208: 107860.

收稿日期:2025-03-17

修改稿日期:2025-04-25