

语义图引导的非侵入式负荷监测

田 旭*, 陈砚圃

西京学院计算机学院, 陕西 西安

收稿日期: 2026年9月6日; 录用日期: 2026年10月7日; 发布日期: 2026年10月15日

摘 要

目的: 针对分钟级总表数据中的设备状态混叠与累计能耗偏差, 提出语义图引导的非侵入式负荷监测方法。方法: 融合低频谱时特征与设备关系先验, 构建图谱偏置注意力, 联合预测设备状态和分项功率, 以总表上界约束抑制超额分解, 并通过原型校准复核异常事件。结果: 在REFIT、UK-DALE和BLOND上, 状态宏平均 $F1$ 分别达到91.7%、93.1%和87.4%, 功率平均绝对误差分别为20.4、15.6和7.8 W; 相较于潜在抽象桥Transformer, $F1$ 提高1.2、0.5和1.7个百分点。消融实验表明, 图谱偏置改善状态区分与功率估计, 上界约束降低累计能耗偏差。结论: 设备关系与电气观测具有互补性, 联合计量约束能够提高低频负荷监测精度, 为分项用能分析提供支持。

关键词

非侵入式负荷监测, 负荷语义图, 多任务学习, 原型拒识, 能量约束

Semantic Graph Guided Nonintrusive Load Monitoring

Xu Tian*, Yanpu Chen

School of Computer Science, Xijing University, Xi'an Shaanxi

Received: September 6, 2026; accepted: October 7, 2026; published: October 15, 2026

Abstract

Objective: To address ambiguous appliance states and cumulative energy-consumption deviations in minute-level aggregate power data, this study proposes a semantic graph guided method for nonintrusive load monitoring. **Methods:** Low frequency spectral-temporal features are integrated with prior knowledge of appliance relationships to construct graph-biased attention for joint prediction of appliance states and disaggregated power, with an aggregate-power upper-bound constraint suppressing

*通讯作者。

文章引用: 田旭, 陈砚圃. 语义图引导的非侵入式负荷监测[J]. 计算机科学与应用, 2026, 16(10): 46-56.
DOI: 10.12677/csa.2026.1610309

over-disaggregation and prototype calibration supporting the review of anomalous events. Results: On REFIT, UK-DALE, and BLOND, the state macro-averaged F1 scores reach 91.7%, 93.1%, and 87.4%, respectively, and the corresponding mean absolute power errors are 20.4, 15.6, and 7.8 W; compared with the latent abstraction bridge Transformer, F1 increases by 1.2, 0.5, and 1.7 percentage points, respectively. Ablation experiments show that graph biases improve state discrimination and power estimation, while the upper-bound constraint reduces cumulative energy-consumption deviations. Conclusion: Appliance relationships and electrical observations are complementary, and incorporating the aggregate-power upper-bound constraint can improve the accuracy of low frequency load monitoring and support appliance-level energy-use analysis.

Keywords

Nonintrusive Load Monitoring, Load Semantic Graph, Multitask Learning, Prototype Rejection, Energy Constraint

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

非侵入式负荷监测(Nonintrusive Load Monitoring, NILM)利用建筑总表推断设备状态与分项功率。Hart [1]建立了聚合信号与设备特征的关联框架。邓晓平等[2]总结了数据处理、负荷分解与评价方法。分钟级记录适用于长期用能分析,但时间平均削弱启动瞬态,相似功率平台与并发运行增加归属歧义,因此需要挖掘有限观测中的区分性信息。

序列到点学习(Sequence-to-Point, S2P) [3]根据总表窗口预测设备中心点功率;多状态双卷积神经网络(Multi-State Dual CNN, MSDC) [4]利用状态转移信息组织预测。NILMFormer [5]以子序列平稳化与时间位置编码处理非平稳性,潜在抽象桥(Latent Abstraction Bridge, LAB)Transformer [6]通过潜在表示约束改善泛化。上述方法侧重序列表征,设备关系与电气观测的协同仍有提升空间。

国内研究从结构信息与特征加权等方面改进负荷分解。庄卫金等[7]通过图信号处理挖掘设备运行状态;李利娟等[8]将外部注意力融入序列到点模型,以强化关键特征;白星振等[9]面向低频数据构建级联宽度学习模型,兼顾分解精度与训练效率。

分钟级监测面临两个问题:相似波形可能对应不同设备组合,需要关系先验补充归属证据;持续功率偏差会累积为能耗误差,需要协调逐点预测与累计计量。设备关系随建筑和作息变化,先验强度也应可调节。

为此,本文提出LoadGraph-Former,结合低频谱时特征与图谱偏置注意力,联合预测状态和功率,以总表上界损失抑制超额分解。三套数据的同域比较、模块消融及双向迁移用于评价监测精度、模块贡献与场景适应性,并分析计算效率。

2. 方法

2.1. 观测输入与任务定义

给定总表有功功率 $P(t)$ 、一阶差分 $\Delta P(t)$ 、有效观测掩码 $m(t)$ 及上下文 $c(t)$,模型读取以时刻 t 为中心的窗口,输出 K 类已知设备的状态概率、非负功率和事件嵌入:

$$X_t = \{P(\tau), \Delta P(\tau), m(\tau), c(\tau)\}_{\tau=t-L}^{t+L}, f_\theta(X_t, G) = \{\hat{s}_t, \hat{p}_t, z_t\}. \quad (1)$$

其中, G 为训练期负荷语义图, $c(t)$ 包含时间戳及可用静态信息。采样间隔为 1 min, $L = 32$, 对应 65 点居中窗口和 32 min 未来观测等待, 适用于离线分析与延迟回放。分表功率用于监督, 推理使用总表、固定设备字典和图谱。

状态采用多标签表示, 标签由训练期确定的分表阈值和最短持续时间生成。功率分支保留待机分量, 多设备切换重叠处保留分项监督, 清晰单事件用于原型学习。整体流程见图 1。

2.2. 低频谱时特征与负荷语义图

谱时编码包括均值、标准差、分位数、前后均值差及边沿斜率。标准化参数仅由训练集估计, 窗口均值与尺度通过旁路保留幅值信息, 各字段均附有效掩码。参数设置见第 3.2 节。

频域分支通过快速傅里叶变换(Fast Fourier Transform, FFT)提取去趋势窗口的能量比例、谱质心和平坦度, 表征 1/120 Hz 以内的波动。事件阈值取绝对功率与局部中位数绝对偏差(MAD)门槛的较大值; 固定步长滑窗估计功率, 事件分支复核边界。

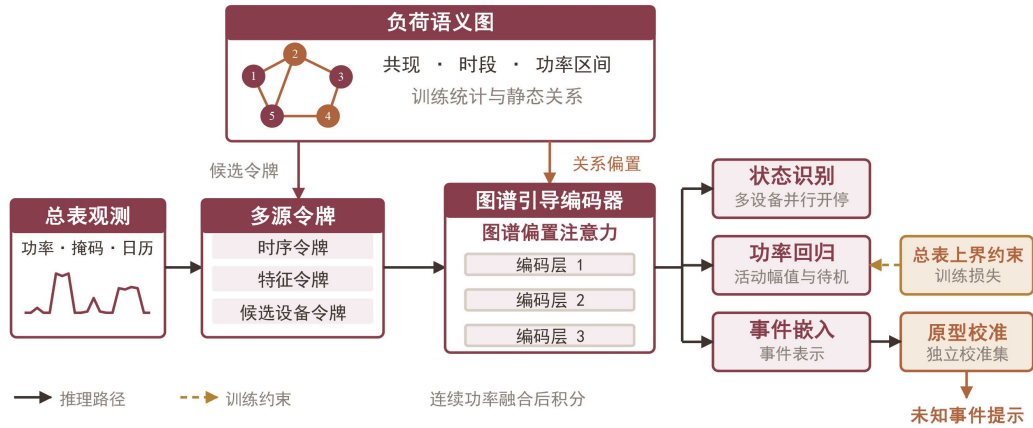


Figure 1. Semantic graph guided multitask load monitoring workflow
图1. 语义图引导的多任务负荷监测流程

稳态令牌为[均值, 标准差, Q_{10} , Q_{50} , Q_{90} , 前后均值差, 覆盖率], 边沿令牌为[最大绝对差分, 前段斜率, 后段斜率, 候选数]; 斜率由功率对秒数最小二乘拟合。频谱令牌为3个频带能量占比、谱质心和平坦度。统计使用有效点, 差分要求相邻点均有效; 短缺口仅为频谱插值, 其余缺口屏蔽频谱令牌, 分表真值与积分不插值。

去趋势序列乘以Hann窗 $w_n = 0.5[1 - \cos(2\pi n/64)]$, $n = 0, \dots, 64$ 。令 X_k 为实数FFT (rFFT) 系数, $S_k = |X_k|^2$; 剔除直流后取 $k = 1, \dots, 32$, $f_k = k/(65 \times 60)$ Hz。频带占比为带内功率和/总功率和, 谱质心为 $\sum f_k S_k / (\sum S_k + \varepsilon)$, 平坦度为 $\exp(\text{mean}(\log(S_k + \varepsilon))) / (\text{mean}(S_k) + \varepsilon)$ 。各字段按训练均值与标准差标准化, 缺测归零并附掩码。

负荷语义图以已知设备为候选节点, 编码训练共现、运行时段、功率区间及有依据的测量归属。每类关系保留来源与可用性掩码, 未提供的拓扑或相别关系不参与计算。第 r 类关系表示为:

$$A_{ij}^{(r)} = M_{ij}^{(r)} \psi_r(v_i, v_j; D_{\text{train}}), \quad (2)$$

其中, M 为可用性掩码, ψ 为静态映射或训练统计, D_{train} 为训练数据。共现关系按边际频率归一化, 并经

平滑、支持度筛选和截断处理。测试时固定设备字典、索引顺序及源域图谱, 偏置映射见第2.3节。

2.3. 图谱偏置注意力

Transformer [10]通过注意力机制建模序列交互。本文依次拼接时间、特征和候选设备令牌(token), 数量分别为 T 、 F 、 K , 总长 $N = T + F + K$ 。设备关系按统一索引映射为分块矩阵:

$$\mathbf{B}_{rel} = \begin{bmatrix} \mathbf{0}_{T \times T} & \mathbf{0}_{T \times F} & \mathbf{B}_{tc} \\ \mathbf{0}_{F \times T} & \mathbf{0}_{F \times F} & \mathbf{B}_{fc} \\ \mathbf{B}_{ct} & \mathbf{B}_{cf} & \mathbf{B}_{cc} \end{bmatrix}. \quad (3)$$

式(3)定义关系偏置的统一索引。共现和静态关系填入候选-候选块, 时段先验填入时间-候选块, 幅值匹配填入特征-候选块; 反向块取转置, 其余位置置零。三类偏置分别构建后相加, 每类关系仅计入一次。掩码控制观测与关系的可用性, 索引对齐方式见图2。

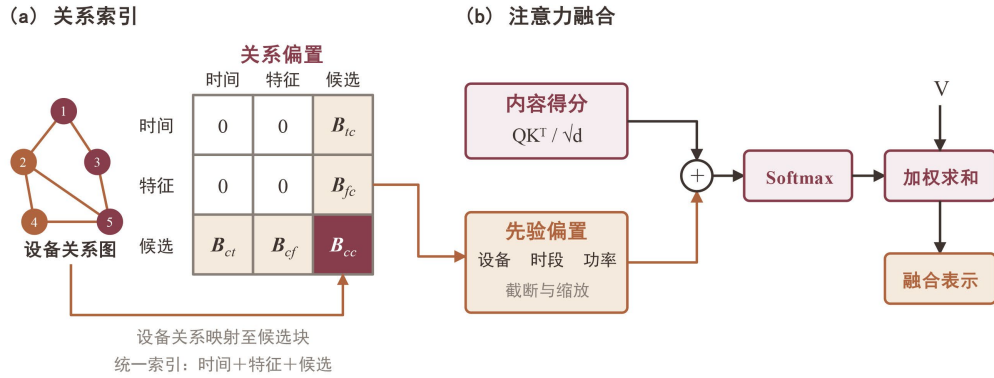


Figure 2. Index alignment between relation biases and attention scores

图2. 关系偏置与注意力分数的索引对齐

根据式(3)的索引, 共现、功率区间重叠和静态关系构成 $\mathbf{B}_G$, 时段与幅值匹配分别构成 $\mathbf{B}_T$ 和 $\mathbf{B}_E$, 注意力计算为:

$$\text{Attn}_G = \text{softmax} \left(\frac{\mathbf{QK}^T}{\sqrt{d}} + \mathbf{B}_G + \mathbf{B}_T + \mathbf{B}_E \right) \mathbf{V}, \quad (4)$$

其中, $\mathbf{Q}$ 、 $\mathbf{K}$ 、 $\mathbf{V}$ 分别为查询、键和值, d 为每个注意力头的维度; 键矩阵 $\mathbf{K}$ 与设备数量 K 由上下文区分。关系强度经温度缩放和截断后加入内容得分, 测试时固定图谱及学习得到的系数。

训练时按关系类型随机失活, 以减轻模型对先验的过度依赖; 推理时保留全部有效关系。

设备对在双方均有效的训练分钟内计数: n_i 、 n_j 为开启数, n_{ij} 为共同开启数。支持度不足或关系缺失时 $M_{ij} = 0$, 对角置零。共现分数及第 h 个头的候选块为:

$$c_{ij} = M_{ij} \frac{n_{ij} + \alpha}{\sqrt{(n_i + \alpha)(n_j + \alpha)} + \varepsilon} \quad (5)$$

$$\mathbf{B}_{cc,ij}^{(h)} = \text{clip} \left(\frac{g_h c_{ij} + g_P o_{ij} + g_S a_{ij}}{\tau_G}, -2, 2 \right) \quad (6)$$

o_{ij} 为活动功率 $[Q_5, Q_{95}]$ 区间交并比, a_{ij} 为同支路指示, 无资料取零; g_h 、 g_P 和 g_S 可学习, 按第3.2节配置初始化。时段分布 $q_{i,u} = (n_{i,u} + 1) / (\sum_u n_{i,u} + 24)$, u 为小时; 时间-候选块取 $\text{clip}(\gamma_T \log(24q_{i,u}), -2, 2)$ 。

边沿幅值 a 取前后均值差绝对值, d_i 为其到活动功率区间的距离, $s_i = \max(Q_{05} - Q_5, 10 \text{ W})$; 特征 - 候选块取 $\text{clip}(-\gamma_E (d_i/s_i)^2, -2, 0)$ 。反向块取转置, 其余块置零, 缺失字段乘掩码。训练后固定统计量、系数与设备顺序。

2.4. 多任务输出与总表上界约束

共享主干之后设置任务特定投影: 状态分支读取语义表示, 功率分支保留幅值旁路, 原型分支生成事件嵌入。设备 i 的功率由状态概率、非负活动幅值和待机分量共同表示:

$$\hat{p}_i(t) = \hat{s}_i(t) a_i(t) + b_i(t), \quad a_i(t) = \text{softplus}(u_i(t)), \quad b_i(t) \geq 0. \quad (7)$$

其中, a_i 为非负活动幅值, b_i 为待机分量。状态分支采用二元交叉熵, 回归损失在共同有效分表掩码内计算, 多设备切换重叠处不计算单类别原型损失。联合目标为:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{cls}} + \lambda_2 \mathcal{L}_{\text{reg}} + \lambda_3 \mathcal{L}_{\text{open}} + \lambda_4 \mathcal{L}_{\text{energy}}, \quad (8)$$

$\mathcal{L}_{\text{cls}}$ 、 $\mathcal{L}_{\text{reg}}$ 和 $\mathcal{L}_{\text{open}}$ 分别为状态分类、功率回归与已知原型紧致损失, $\mathcal{L}_{\text{energy}}$ 用于惩罚超额分解。各项损失按训练尺度归一化, 权重与训练配置见第3.2节。上界损失定义为:

$$\mathcal{L}_{\text{energy}} = \frac{1}{|\mathcal{V}|} \sum_{t \in \mathcal{V}} \max\left(0, \sum_{i \in \mathcal{K}} \hat{p}_i^{\text{aw}}(t) - P(t) - \epsilon_m\right). \quad (9)$$

其中, $\mathcal{V}$ 为有效时刻集合, ϵ_m 为固定计量容差。该损失仅惩罚已知分项之和超出总表容差的部分, 为未计量背景保留空间; 总表与分项之差作为异常核查残差。

实现时, 分类按有效设备 - 时刻平均; 回归误差按训练功率尺度归一化后计算Huber损失; 原型项计算清晰单事件的归一化平方距离; 式(9)按有效时刻平均。各项再除以初始100批次均值加 ϵ 。联合阶段取 $\lambda_4(e) = 0.20 \min(1, e/10)$, 其余权重按第3.2节固定。模型包保存归一化统计、阈值、关系系数、设备映射与随机种子。

2.5. 原型校准与连续推理

原型网络[11]依据嵌入与类别中心之间的距离进行分类。本文利用清晰、可靠的已知训练事件估计原型中心 μ_c 和类内尺度 σ_c , 嵌入 z 的归一化距离定义为:

$$D_c(z) = \frac{\|z - \mu_c\|_2}{\sigma_c + \epsilon}. \quad (10)$$

固定主干与原型中心后, 在独立已知类校准集上计算最小距离分数。设有效校准分数数量为 N_{cal} , 名义误拒目标 $\delta = 0.05$, 将分数按升序排列并确定阈值:

$$\tau = s_{(k)}, \quad k = \min\left\{N_{\text{cal}}, \lceil (N_{\text{cal}} + 1)(1 - \delta) \rceil\right\}. \quad (11)$$

最小归一化距离大于 τ 时, 将事件标记为未知候选, 否则归入最近的已知类别。 δ 确定已知校准分布上的名义工作点。事件复核与功率估计并行, 为异常模式识别提供辅助依据。

训练依次包括表示预训练、已知类监督和联合微调, 随后冻结主干并独立校准。重叠窗口的功率输出按位置和有效性融合, 在统一时间轴上积分; 事件分支同步复核状态切换边界。

3. 结果

3.1. 数据与评价设置

实验采用REFIT住宅数据[12]、英国住宅设备级用电数据集(UK Domestic Appliance-Level Electricity,

UK-DALE) [13]和建筑级办公设备数据集(Building-Level Office Environment Dataset, BLOND) [14]进行评价。表1列出采集规模与低频观测条件, 评价记录按通道、功率类型和覆盖率筛选。

Table 1. Data sources and low frequency observation conditions

表1. 数据来源与低频观测条件

数据集	公开采集规模	原始低频记录	本文评价用途
REFIT	20个家庭; 每户9路设备测量	8 s功率	住宅同域与共同设备迁移
UK-DALE	5个家庭	分表6 s; 部分总表有功1 s	独立住宅同域与双向迁移
BLOND	16类设备	1 s汇总功率	办公设备同域监测

记录重采样为1 min平均有功功率, 覆盖率不低于0.90。UK-DALE采用家庭1、2、5的总表有功通道。预处理保留总表背景; 设备更换时分段确定有效期, 同类实例按固定映射聚合, 缺失分表记录由掩码标识。

同域记录按时间顺序划分为训练、验证、校准和测试集, 比例为6:1:1:2, 并剔除跨分区窗口。标准化与图谱仅由训练集估计, 验证集用于参数选择, 校准集用于阈值确定; 结果汇总5个配对随机种子。迁移实验选取名称、测量和运行定义一致的设备, 在源域训练与校准, 在目标域评价冻结模型。

状态宏平均 $F1$ ($Macro-F1$)为各设备开停二分类 $F1$ 的平均值, 功率平均绝对误差(Mean Absolute Error, MAE)在共同有效时刻计算。设备能量由同一有效区间内的功率积分得到:

$$MAE = \frac{1}{K} \sum_i \frac{\sum_{t \in \mathcal{V}_i} |\hat{p}_i(t) - p_i(t)|}{|\mathcal{V}_i|}, \quad E_i = \frac{\sum_{t \in \mathcal{V}_i} p_i(t) \Delta t}{3.6 \times 10^6}. \quad (12)$$

其中, $\mathcal{V}_i$ 为设备 i 的有效时刻集合, 功率、 Δt 和 E_i 分别以W、s和kWh计。累计误差采用设备能量加权的聚合误差(Signal Aggregate Error, SAE):

$$SAE_w = \frac{\sum_i |\hat{E}_i - E_i|}{\sum_i E_i + \epsilon} \times 100\%. \quad (13)$$

参照NILMTK [15]的数据与指标一致性原则, 各项指标使用相同设备和有效时段。 SAE 先计算各设备能量误差的绝对值, 再汇总归一化, 避免正负误差抵消; 该指标与状态 $F1$ 、功率 MAE 共同评价监测性能。

3.2. 比较方法与训练配置

基线包括S2P、MSDC、NILMFormer和LAB Transformer, 均采用共同有效时段与信息条件。无状态输出的模型按统一阈值判定开停, 序列与中心点输出均在有效中心区间评价。主要参数见表2。

Table 2. Main training and calibration settings of LoadGraph-Former

表2. LoadGraph-Former主要训练与校准参数

参数	取值	参数	取值
采样间隔	1 min	窗口点数	65
嵌入维度	128	编码器层数	3
注意力头数	4	前馈层维度	256
随机失活率	0.10	批大小	64
初始学习率	1×10^{-3}	优化器	AdamW
训练轮数	80	名义校准误拒目标 δ	0.05

各模型采用相同的数据划分、采样率和评价掩码，并统一训练预算与早停规则。特征工程、图谱和损失参数见表3。

Table 3. Feature engineering and model training parameters
表3. 特征工程与模型训练参数

项目	参数设置	项目	参数设置
令牌/步长	$T = 65, F = 3$; 步长 1 点	日历输入	小时/星期各用正余弦编码
前后统计区间	中心前、后各 8 点，不含中心	分位数/标准差	0.10/0.50/0.90; 线性插值; $\text{ddof} = 0$
频谱窗口	线性去趋势; 对称 Hann 窗 65 点	FFT/频带	65 点 rFFT; $k = 1 \sim 4/5 \sim 12/13 \sim 32$
缺失处理	覆盖 ≥ 0.90 ; 每段缺口 ≤ 2 点	事件检测	MAD 窗 21 点; 门槛 $\max(30 \text{ W}, 3 \text{ MAD})$
候选事件合并	间隔 $\leq 2 \text{ min}$ 时保留最大边沿	数值稳定项	归一化域 $\varepsilon = 10^{-8}$; 功率尺度 $\geq 1 \text{ W}$
关系统计	平滑 $\alpha = 1$; 边际支持数 ≥ 20	图谱偏置	温度 1; 截断 ± 2 ; 关系失活 0.10
关系系数初值	共现/区间/静态: 0.50/0.25/0.25	时段/幅值系数	均为 0.25, 可学习, 测试冻结
损失权重	$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (1, 1, 0.10, 0.20)$	损失尺度/预热	初始 100 批次均值; 联合阶段 10 轮
回归/上界容差	归一化 Huber, $\delta = 1$; $\varepsilon_m = 10 \text{ W}$	状态标签	阈值 $\max(10 \text{ W}, 0.1 \text{ } Q_{95})$; 持续 $\geq 2 \text{ min}$
训练阶段	预训练 10/监督 50/联合 20 轮	AdamW 补充项	$\beta = (0.9, 0.999)$, $\varepsilon = 10^{-8}$; 衰减 10^{-4}
学习率/梯度	余弦退火至 10^{-5} ; 梯度范数 ≤ 1	配对种子	2021, 2022, 2023, 2024, 2025

3.3. 同域监测性能

同域结果见表4。LoadGraph-Former在三套数据的状态*Macro-F1*、功率*MAE*和设备加权*SAE*上均优于所比较的基线。

Table 4. Within-domain monitoring performance on three datasets
表4. 三套数据的同域监测性能比较

数据集	模型	状态 <i>Macro-F1</i> /%	<i>MAE</i> /W	设备加权 <i>SAE</i> /%
REFIT	S2P	84.6	31.8	8.9
	MSDC	87.1	27.4	7.4
	NILMFormer	89.8	22.9	6.2
	LAB Transformer	90.5	21.8	5.9
	LoadGraph-Former	91.7	20.4	5.4
UK-DALE	S2P	86.9	25.7	8.1
	MSDC	90.2	20.1	6.0
	NILMFormer	92.0	17.2	5.2
	LAB Transformer	92.6	16.3	4.9
	LoadGraph-Former	93.1	15.6	4.7
BLOND	S2P	78.2	12.4	10.6
	MSDC	81.0	10.6	9.2

续表

NILMFormer	85.0	8.8	7.5
LAB Transformer	85.7	8.5	7.2
LoadGraph-Former	87.4	7.8	6.8

注：加粗表示同一数据集上的最优结果。

相较于LAB Transformer (图3), 本文方法在三套数据上的 $F1$ 提高1.2、0.5和1.7个百分点, MAE 降低1.4、0.7和0.7 W, SAE 降低0.5、0.2和0.4个百分点, 表明状态与分项计量精度同步改善。

3.4. 关键模块消融

在相同REFIT划分、窗口、预算与配对种子下, 分别移除图谱偏置、低频特征、上界损失与表示预训练, 结果见表5。

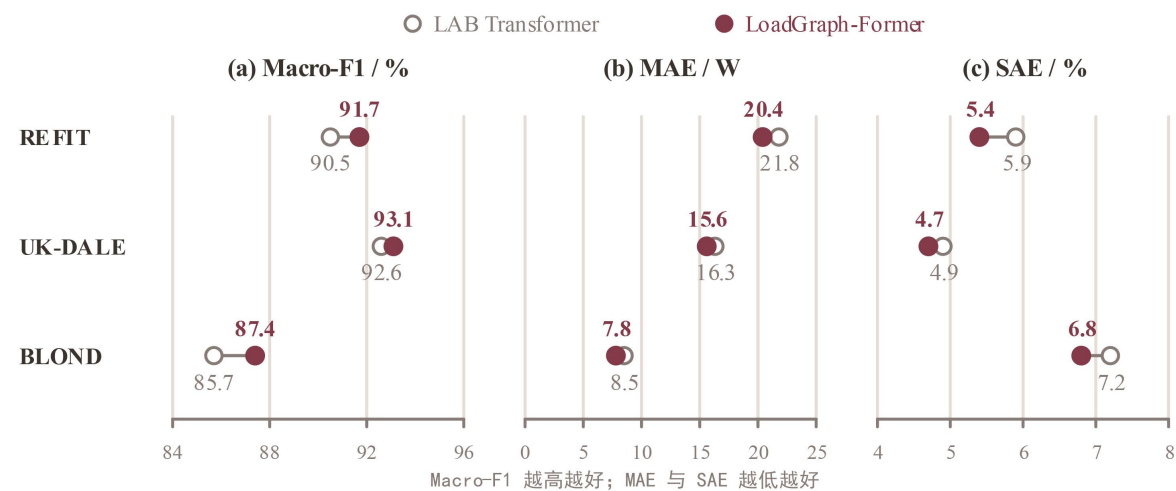


Figure 3. Within-domain comparison with the LAB Transformer
图3. 与LAB Transformer的同域性能对比

Table 5. Ablation results for key components on REFIT
表5. REFIT上的关键模块消融结果

设置	状态Macro-F1/%	MAE/W	设备加权SAE/%
完整方法	91.7	20.4	5.4
移除图谱偏置	90.3	22.5	6.1
移除低频特征分支	90.8	21.6	5.8
移除训练上界损失	91.6	21.0	6.0
移除表示预训练	90.7	22.1	6.2

注：消融实验采用相同REFIT划分、窗口与训练预算。

如图4所示, 移除图谱偏置使 $F1$ 下降1.4个百分点、 MAE 增加2.1 W, 其状态和功率性能退化最明显。移除低频特征使 MAE 增加1.2 W, 说明关系先验需要电气证据支持。移除上界损失对 $F1$ 影响较小, 却使 SAE 由5.4%升至6.0%, 体现了该约束对累计偏差的控制作用。

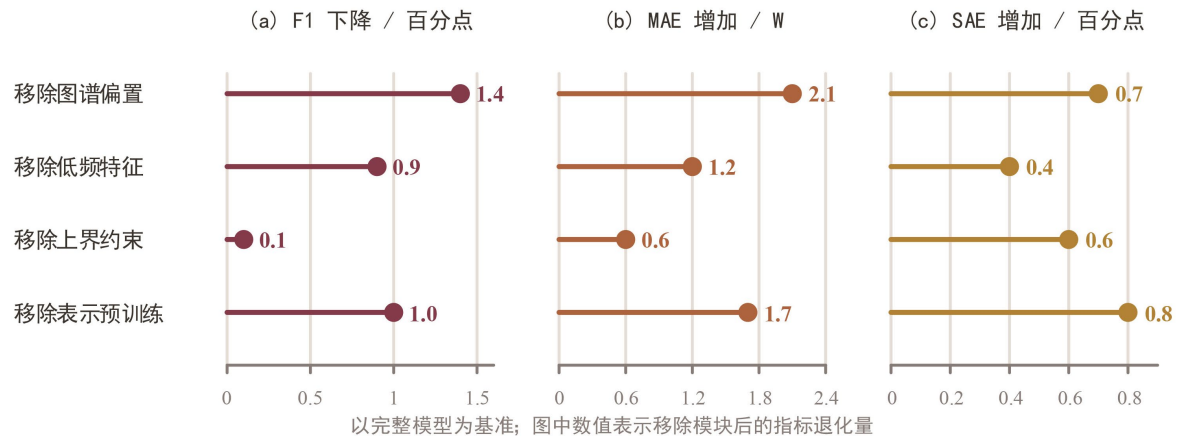


Figure 4. Performance changes after removing individual components
图4. 移除关键模块后的性能变化

3.5. 共同设备双向迁移

共同设备双向迁移结果见表6。训练、验证与校准在源域完成，目标域评价冻结模型，以检验跨住宅泛化能力。

Table 6. Bidirectional transfer performance for common appliances
表6. 共同设备的双向迁移性能

迁移方向	方法	状态Macro-F1/%	MAE/W
REFIT→UK-DALE	NILMFormer	81.7	27.2
	LAB Transformer	83.2	25.6
	LoadGraph-Former	83.0	25.9
UK-DALE→REFIT	NILMFormer	79.4	36.8
	LAB Transformer	81.1	34.5
	LoadGraph-Former	82.0	33.4

注：加粗表示同一迁移方向上的最优结果。

REFIT至UK-DALE迁移时，本文F1和MAE为83.0%与25.9 W，略逊于LAB Transformer的83.2%与25.6 W；反向达到82.0%与33.4 W，分别改善0.9个百分点和1.1 W。两个方向均优于NILMFormer，表明关系收益取决于源、目标域运行模式的匹配程度。

3.6. 模型复杂度与推理效率

设编码层数为 M 、嵌入维度为 $D = hd$ ，主干复杂度为 $O(MND^2 + MN^2D)$ ，其中 $N = T + F + K$ 。密集偏置存储为 $O(N^2)$ ，原型距离计算为 $O(KD)$ 。图谱离线构建，推理按固定索引读取；额外偏置加法为 $O(MhN^2)$ ， h 为头数。

效率评估采用65点窗口、9个设备、3个特征令牌、批量1及32位浮点(FP32)。单设备基线汇总9台设备的参数、运算量与输出耗时。参数量为可训练标量数；浮点运算量(Floating-Point Operations, FLOPs)覆盖编码、注意力和输出，乘加计2次运算。

中央处理器(CPU)为Intel Core i7-12700H(单线程、32 GB内存), 图形处理器(GPU)为NVIDIA RTX 3060 Laptop GPU (6 GB显存)。软件为Python 3.11、PyTorch 2.5.1和CUDA 12.1。关闭梯度, 预热100次、计时1000次, 重复5轮取均值; GPU计时前后同步, 数据预置于内存或显存。结果见表7。

Table 7. Model complexity and inference latency
表7. 模型复杂度与推理耗时对比

模型	参数量/M	FLOPs/G	CPU 均值/ms	GPU 均值/ms
S2P	12.60	1.280	26.4	4.8
MSDC	9.80	0.960	22.1	4.1
NILMFormer	0.48	0.082	6.0	1.6
LAB Transformer	0.55	0.090	6.4	1.7
LoadGraph-Former	0.60	0.100	6.9	1.9

注: 耗时为批量1的模型前向均值, 不含文件读取、主机传输、特征提取、偏置组装及输出融合。

与LAB Transformer相比, 本文模型增加0.05 M参数和0.010 GFLOPs, CPU与GPU平均耗时分别增加0.5和0.2 ms。结合表4的精度增益, 关系建模以有限计算开销改善了监测性能。

4. 讨论

4.1. 关系先验与电气特征的互补性

图谱偏置为相似电气观测补充设备关系证据。移除图谱或低频特征均降低状态与功率估计性能, 说明两类信息具有互补性。BLOND上的 FI 增益为1.7个百分点, 体现了关系信息对办公设备辨别的作用; 能耗误差还受设备功率尺度影响。

4.2. 多任务目标与能量偏差控制

移除上界损失对 FI 影响较小, 却使 SAE 增加0.6个百分点, 说明持续功率高估需要计量约束。分表监督确定设备归属, 上界约束保留背景空间, 原型距离与功率残差辅助异常核查。

4.3. 迁移适应性与应用拓展

迁移排序变化表明关系先验具有场景依赖性, 支持度筛选与偏置调节有助于协调源域先验和目标观测。后续可通过留出类别评价拒识, 并按建筑或日期分块估计性能不确定性。

4.4. 非因果限制与因果化路径

65点居中窗口需要32 min未来观测, 表7的毫秒级耗时反映模型计算效率。响应时间还包括观测等待、采样聚合、排队、预处理与融合; 当前模型适用于离线监测和延迟分析。

因果化可采用 $[t-64, t]$ 尾随窗口预测末端, 仅用已到达数据计算统计特征、FFT和事件阈值, 并施加因果掩码。固定训练图谱后重训主干并校准原型。未来形态信息减少可能影响启停判别, 动态令牌更新也增加键值缓存维护难度; 应同时比较 FI 、 MAE 、 SAE 及端到端平均延迟和第95百分位延迟。

5. 结论

本文提出LoadGraph-Former, 以语义图偏置、多任务输出与总表上界约束联合建模, 在三套数据上取得87.4%~93.1%的状态 FI , 并降低功率与能耗误差。消融表明关系信息、电气特征和计量约束具有互补

性; 效率与迁移结果说明精度收益应结合计算开销和场景匹配程度评价。后续将研究因果建模与未知事件识别。

参考文献

- [1] Hart, G.W. (1992) Nonintrusive Appliance Load Monitoring. *Proceedings of the IEEE*, **80**, 1870-1891. <https://doi.org/10.1109/5.192069>
- [2] 邓晓平, 张桂青, 魏庆来, 等. 非侵入式负荷监测综述[J]. 自动化学报, 2022, 48(3): 644-663.
- [3] Zhang, C., Zhong, M., Wang, Z., Goddard, N. and Sutton, C. (2018) Sequence-to-Point Learning with Neural Networks for Non-Intrusive Load Monitoring. *Proceedings of the AAAI Conference on Artificial Intelligence*, **32**, 2604-2611. <https://doi.org/10.1609/aaai.v32i1.11873>
- [4] He, J., Liu, J., Zhang, Z., Chen, Y., Liu, Y., Khossainov, B., et al. (2023) MSDC: Exploiting Multi-State Power Consumption in Non-Intrusive Load Monitoring Based on a Dual-CNN Model. *Proceedings of the AAAI Conference on Artificial Intelligence*, **37**, 5078-5086. <https://doi.org/10.1609/aaai.v37i4.25636>
- [5] Petralia, A., Charpentier, P., Kadhi, Y. and Palpanas, T. (2025) NILMFormer: Non-Intrusive Load Monitoring That Accounts for Non-Stationarity. *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*, Toronto, 3-7 August 2025, 4761-4772. <https://doi.org/10.1145/3711896.3737251>
- [6] Liu, Y., Kong, Z., Huang, T., Yang, Y., Song, C., Han, Q., et al. (2026) Latent Abstraction Bridge Transformer for Generalizable Nonintrusive Load Monitoring. *Scientific Reports*, **16**, Article No. 20046. <https://doi.org/10.1038/s41598-026-48516-0>
- [7] 庄卫金, 张鸿, 方国权, 等. 基于设备运行状态挖掘的非侵入式负荷分解方法[J]. 电力建设, 2020, 41(8): 9-16.
- [8] 李利娟, 刘海, 刘红良, 等. 融合外部注意力机制的序列到点非侵入式负荷分解[J]. 上海交通大学学报, 2024, 58(6): 846-854.
- [9] 白星振, 康家豪, 尚继伟, 等. 基于级联宽度学习与麻雀算法的非侵入式负荷分解方法[J]. 山东科技大学学报(自然科学版), 2024, 43(2): 102-111.
- [10] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017) Attention Is All You Need. *Advances in Neural Information Processing Systems*, **30**, 5998-6008. https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
- [11] Snell, J., Swersky, K. and Zemel, R. (2017) Prototypical Networks for Few-Shot Learning. *Advances in Neural Information Processing Systems*, **30**, 4077-4087. https://proceedings.neurips.cc/paper_files/paper/2017/hash/cb8da6767461f2812ae4290eac7cbc42-Abstract.html
- [12] Murray, D., Stankovic, L. and Stankovic, V. (2017) An Electrical Load Measurements Dataset of United Kingdom Households from a Two-Year Longitudinal Study. *Scientific Data*, **4**, Article No. 160122. <https://doi.org/10.1038/sdata.2016.122>
- [13] Kelly, J. and Knottenbelt, W. (2015) The UK-DALE Dataset, Domestic Appliance-Level Electricity Demand and Whole-House Demand from Five UK Homes. *Scientific Data*, **2**, Article No. 150007. <https://doi.org/10.1038/sdata.2015.7>
- [14] Kriechbaumer, T. and Jacobsen, H. (2018) BLOND, a Building-Level Office Environment Dataset of Typical Electrical Appliances. *Scientific Data*, **5**, Article No. 180048. <https://doi.org/10.1038/sdata.2018.48>
- [15] Batra, N., Kelly, J., Parson, O., Dutta, H., Knottenbelt, W., Rogers, A., et al. (2014) NILMTK: An Open Source Toolkit for Non-Intrusive Load Monitoring. *Proceedings of the 5th International Conference on Future Energy Systems*, Cambridge, 11-13 June 2014, 265-276. <https://doi.org/10.1145/2602044.2602051>