

基于多算法融合的非侵入式负荷监测模型

史茗元¹, 钱本华¹, 宋自强¹, 王睿¹, 刘尧^{2,3}

(1. 东北大学信息科学与工程学院, 沈阳 110819; 2. 东北大学计算机科学与工程学院, 沈阳 110169; 3. 麦考瑞大学计算机学院, 新南威尔士州 2109)

摘要: 尽管非侵入式负荷监测(non-intrusive load monitoring, NILM)已经得到了广泛的研究, 然而现有的非侵入式负荷监测模型存在多工作状态电器预测困难的问题, 导致预测精度显著降低。为此设计了一种基于多算法融合的非侵入式负荷监测模型。首先, 对REDD低频数据集进行基于时间的插值、过采样等方式预处理数据。其次, 模型采用图卷积神经网络(graph convolutional networks, GCN)和卷积神经网络(convolutional neural network, CNN)提取功率特征, 将功率特征输入到自注意力机制和长短期记忆网络(long short-term memory, LSTM)中, 有效提取了输入信号中的关键特征, 提高多工作状态电器的预测精度。最后, 利用预处理后REDD低频数据集进行仿真验证, 实验结果表明所提出的模型在MAE、SAE和 R^2 指标上均优于对比模型, 能够有效实现负荷分解。

关键词: 非侵入式负荷监测; 自注意力机制; 图卷积神经网络

Non-Intrusive Load Monitoring Model Based on Multi-Algorithm Fusion

SHI Mingyuan¹, QIAN Benhua¹, SONG Ziqiang¹, WANG Rui¹, LIU Yao^{2,3}

(1. College Of Information Science and Engineering, Northeastern University, Shenyang 110819, China; 2. School of Computer Science and Engineering, Northeastern University, Shenyang 110169, China; 3. School of Computing, Macquarie University, NSW 2109, Australia)

Abstract: Despite extensive research on non-intrusive load monitoring(NILM), existing models face challenges in accurately predicting multiple operational states of appliances, leading to significant decreases in prediction accuracy. To address this issue, this paper proposes non-intrusive load monitoring model based on multi-algorithm fusion. Firstly, the REDD low-frequency dataset is preprocessed using time-based interpolation and oversampling. Secondly, the model employs graph convolutional networks(GCN) and convolutional neural networks(CNN) to extract power features, which are then fed into a self-attention mechanism and long short-term memory(LSTM) networks. This effectively captures the key features of the input signals, thereby improving the prediction accuracy for appliances with multiple operational states. Finally, simulation verification is conducted using the preprocessed REDD low-frequency dataset. The experimental results indicate that the proposed model outperforms comparative models in terms of MAE, SAE, and R^2 , demonstrating its effectiveness in load disaggregation.

Key words: non-intrusive load monitoring; self-attention mechanism; graph convolutional neural network

0 引言

能源体系的创新和需求侧管理的日益重要性已经成为推动能源转型和优化的关键因素^[1-2]。用户用电数据作为需求侧管理的基石, 其准确性和完整性对管理策略的精细度和有效性有着直接影响^[3]。

通过实时监测用户的电力负荷和用电行为, 可以得到宝贵的实时用电数据^[4]。

传统的侵入式监测方法需要在每个用电设备内部安装传感器, 通过收集这些设备的实时用电信息来监测整个电力负荷。这种方法尽管可以得到真实可靠的数据, 但高昂的成本和不便于推广的缺点使得非侵入式负荷监测成为了研究的热点。非侵入式监测方法通过分析整个家庭的电力使用情况, 无需在单个设备上安装传感器, 就能实现对各个电器能

基金项目: 国家自然科学基金资助项目(U23B20118)。

Foundation item: Supported by the National Natural Science Foundation of China (U23B20118).

耗的监测和识别,大大降低了成本,提高了可访问性,为家庭用户提供了更加经济、高效的能源管理解决方案^[5-6]。

非侵入式负荷监测首先由 Hart^[7]在 1992 年提出。这种方法通过分析特定区域在特定时间内的总功耗,来识别该区域内各个独立负荷的状态和能耗情况。早期研究主要采用组合优化和机器学习方法。组合优化通过将电器状态组合成一个状态集,将问题转化为整数规划问题。例如,文献[8]采用粒子群优化算法将总电力数据分解到各电器,解决了少数电器稳态电流的分解问题,但误差较大。机器学习中的聚类算法和隐马尔可夫模型也被广泛应用。文献[9]利用动态聚类算法实现了对大部分电器功率的分解,但不适用于功率波动较小且运行时间短的电器。

近年来,随着智能电表和物联网的快速发展,非侵入式负荷监测的研究取得了显著进展^[10]。深度学习技术在各种模式识别问题中表现优异,这吸引了大量研究人员将这些人工智能技术应用于非侵入式负荷监测领域^[11-13]。Kelly^[14]等人首次将深度学习技术用到非侵入式负荷监测中,模型分解精度较以往方法大大提高。文献[15]提出的序列到序列和序列到点模型具有更好的分解性能,但存在网络层数少、无法提取更高级别负荷特征的问题。文献[16]开发了一种新型的串联卷积神经网络,可以显著减少由于长期的、低效的背景负载而导致的分析误差,从而大幅度改善了负荷分解的精度。文献[17]引入了一种基于时序卷积网络和长短时记忆网络的新型半监督多标签深度学习技术,用于从有标签和无标签数据中对家电运行状态进行分类。文献[18]则首次使用图卷积神经网络进行非侵入式负荷监测,根据反映了相应设备信号的不同操作状态的熵度量将聚合的信号编码为不同的图节点,创建了包含时不变依赖关系的编码,在聚合信号中实现了高效的信息表达,但是所提出的模型只能在多状态设备上效果良好。文献[19]提出了一种基于贝叶斯优化和双向长短期记忆网络的回归模型,利用双向长短期记忆网络来捕捉时间序列数据中的长期依赖关系,并通过贝叶斯优化算法来优化模型参数,然而模型仅使用双向长短期记忆网络,无法充分捕捉到电器之间的复杂关联性,导致模型在多电器同时工作时识别准确性下降。

针对以上问题,本文提出一种基于多算法融合的非侵入式负荷监测方法。包括数据预处理和 NILM 负荷监测模型。本文贡献主要有以下 3 个方面。

1)提出了一种基于多算法融合的非侵入式负荷监测模型结构,旨在解决现有方法无法充分捕捉电器间复杂关联性的问题。该模型通过结合图卷积神经网络与长短期记忆网络从电器间的关联性与时间序列特征两个角度进行建模,显著提升了多状态电器的预测精度。同时模型引入卷积神经网络来提取时间序列中的局部特征,并通过自注意力机制解决了长时间序列中关键特征信息提取不足的问题,进一步增强了预测精度。

2)将 NILM 任务等价为状态识别和负荷分解两个子任务,通过设置阈值判断电器状态并使用一种分类网络和回归网络并行的负荷分解模型结构。这种结构能够同时处理状态识别和负荷分解任务,提高了模型的综合性能和泛化能力。

3)针对 REDD 低频数据集的数据缺失与类别不平衡问题,提出了一种结合时间插值、过采样与动态滑动窗口阈值的混合预处理方法,有效提升了数据质量。通过上述方法的综合应用模型在 REDD 数据集上的性能表现优于对比模型,验证了方法的有效性和优越性。

1 理论分析

1.1 负荷分解问题模型

非侵入式负荷监测的任务是从电表总线处测量的家庭总用电数据中,推断出特定电器的单独使用情况。在给定的时间段 $[1, \dots, T]$ 内,总用电量 $(x_1, \dots, x_T)$ 是由所有活动电器的功率消耗加上一个噪声项组成的。非侵入式负荷监测旨在将总用电量分解为各单独电器的使用情况。其核心目标是通过监测负荷的运行状态,如设备的启停与运行模式,以及电气参数,包括电压、电流和功率等,来实现对各电器的精确检测。本文主要研究电器功率的监测。对于大多数情况,只关心所有电器中一个电器的功率使用情况,因此其他电器可以被建模为一个未知因素 $u = (u_1, u_2, \dots, u_M)$,把这些加在一起就得到了如方程 1 所示的系统模型。

$$x_t = \sum_{i=1}^T y_{it} + \varepsilon_t + u_t \quad (1)$$

式中: x_t 为在 t 时间测量的总的有功功率; y_{it} 为第 i 个电器的单独功率; ε_t 为噪声项; u_t 为目标电器以外的电器。在去噪场景中, 噪声项不存在, 总用电量是已知电器功率消耗的总和。然而, 在现实世界中, 由于测量误差和未监测电器的存在, 噪声项是不可避免的。

1.2 电器功率分析

居民家庭中的电器可以分为单工作模式电器和多工作模式电器两种。单工作模式电器例如电磁炉、电热水壶、冰箱等, 它们在工作时只有一种功率峰值, 可以简单地认为有开启和关闭两种状态; 而多工作模式电器例如洗衣机、洗碗机、微波炉等, 在工作时有多种工作模式。

4种常见电器功率曲线如图1所示。

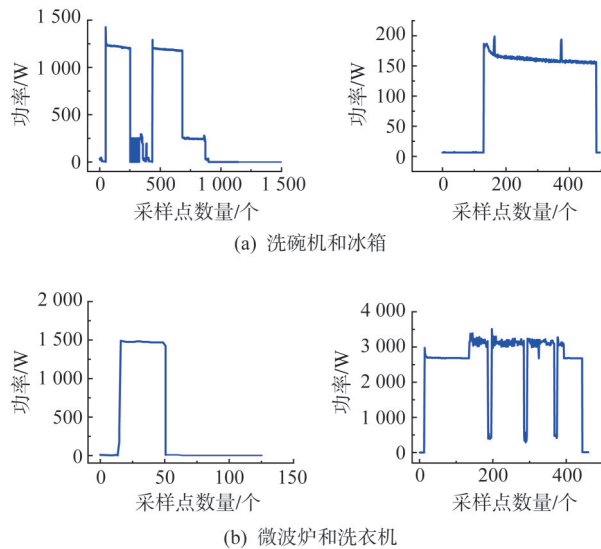


图1 4种电器运行功率曲线

Fig. 1 Operating power curves of four electrical appliances

本文重点关注电器功率消耗中存在的复杂关联性, 特别是在多个电器同时运行时的情况。在家庭环境中, 多个电器往往同时工作, 导致各电器功率信号发生叠加, 从而使得总功率变化复杂且难以区分。例如, 单个电器如冰箱可能呈现周期性的启动和停止, 而多状态电器如洗衣机和微波炉在不同运行模式下则会显示出多峰值和剧烈波动的特征; 同时, 一些电器(如空调)的启动还可能间接影响其他设备的功率消耗, 形成明显的协同效应。这种多电器间的相互干扰与协同作用使得负荷状态识别变得尤为复杂, 需要考虑电器之间的相互影响和环境因素对功率波动的综合作用。

由上述可知, 在处理具有多工作模式的家用电器时, 由于电器在不同工作状态下的功率表现差异, 准确预测其功率消耗变得尤为复杂。为了解决这一挑战, 本文采用了阈值法对电器的开关状态进行分类。阈值法的实现较为简单, 具有一定的通用性, 电器阈值设置为最低工作功率的值。然而, 这种方法也存在一定的局限性。在实际应用中, 电器运行时可能会产生噪声, 这些噪声有时会超过预设的最低功率阈值, 从而影响预测的准确性。为克服这一问题, 本文将状态识别和负荷分解结合起来, 利用分类网络对电器的工作状态进行识别, 以判断电器的开关状态; 通过回归网络对功率进行预测。通过这种方式两类网络协同工作, 能够更有效地预测具有多工作状态电器的功率消耗。

为实现精确的负荷分解, 需要充分分析时间序列数据中的多种特征。首先, 局部特征如电器开启或关闭时出现的短时功率峰值和局部波动, 是捕捉瞬时变化的重要依据; 其次, 长期依赖性特征反映了电器运行过程中周期性和趋势性的变化, 如冰箱压缩机的周期性工作和洗衣机工作时各阶段的动态转换。这些特征有助于模型学习持续性和渐变信息。

2 本文NILM模型

本节基于第1节提出了一种基于多算法融合的非侵入式负荷监测模型, 介绍如下。

2.1 卷积神经网络

CNN能捕捉时间序列中具有平移不变性的局部特征, 在非侵入式负荷监测中, 有助于识别不同电器的功率消耗模式, 提升负荷分解准确性。同时, CNN的权重共享机制可减少参数量, 提高计算效率和泛化能力。

在本文的非侵入式负荷监测任务中, CNN被用于处理家庭用电的功率时间序列数据。输入层接收家庭用电的功率时间序列数据作为模型输入。卷积层用来捕捉功率时间序列中的局部特征, 如电器启动时的功率峰值、稳态时的功率变化等。池化层提高计算效率和模型的泛化能力, 并保留重要特征。全连接层将提取的特征向量转换为预测的电器功率, 实现负荷分解。softmax层公式如式(2)所示。

$$\text{softmax}(z_i) = \frac{\exp(z_i)}{\sum_{j=1}^n \exp(z_j)} \quad (2)$$

式中: $\exp()$ 为指数函数; z_i 和 z_j 为输入向量 $\mathbf{z}$ 中的第 i 个和第 j 个元素; n 为输入向量 $\mathbf{z}$ 的元素数量。

2.2 长短期记忆网络

长短期记忆网络 (long short-term memory, LSTM)^[20] 是一种特殊的循环神经网络 (recurrent neural network, RNN)^[21], 用于来处理和预测时间序列数据。长短期记忆神经网络结构如图 2 所示。

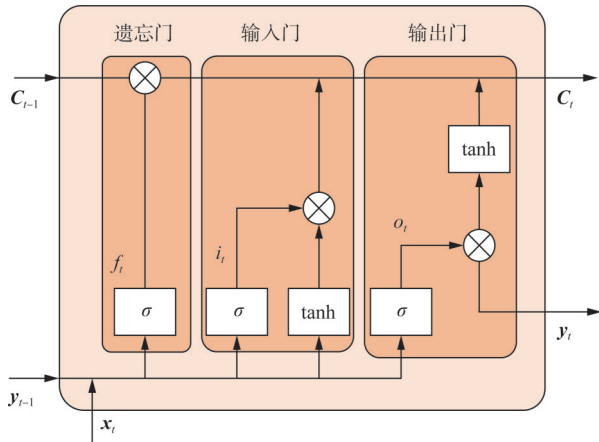


图 2 LSTM 结构图

Fig. 2 LSTM structure diagram

LSTM 通过引入包含 3 种门控机制的记忆单元来克服传统 RNN 在处理长序列时遇到的梯度消失问题。这些记忆单元具有 3 种门: 遗忘门、输入门和输出门。

遗忘门控制着前一个记忆状态的保留程度, 决定了在当前时间步长是否要忘记之前的记忆状态。其计算公式如式 (3) 所示。

$$f_t = \sigma(W_f[y_{t-1}, x_t]) + b_f \quad (3)$$

输入门控制着新信息的输入程度, 决定了在当前时间步更新记忆状态时新信息的输入量。其计算公式如式 (4)~(6) 所示。

$$i_t = \sigma(W_i[y_{t-1}, x_t]) + b_i \quad (4)$$

$$\tilde{C}_t = \tanh(W_c[y_{t-1}, x_t]) + b_c \quad (5)$$

$$C_t = i_t \times \tilde{C}_t + f_t \times C_{t-1} \quad (6)$$

输出门控制着记忆状态到隐藏状态的输出程度, 决定了当前时间步的隐藏状态如何基于更新后的记忆状态进行输出。其计算公式如式 (7)~(8) 所示。

$$O_t = \sigma(W_o[y_{t-1}, x_t]) + b_o \quad (7)$$

$$y_t = O_t \times \tanh(C_t) \quad (8)$$

式中: 下标 f 、 i 、 o 、 c 分别为遗忘门、输入门、输出门和细胞状态; 下标 t 和 $t-1$ 为当前时刻和前一时刻; W 、 b 分别为激活函数的权重和对应的偏差; f_t 、 i_t 、 C_t 和 y_t 分别为遗忘门、输入门、当下细胞层状态和当下细胞整体的输出; x_t 为当前时刻的输入向量; $\tilde{C}_t$ 为候选细胞状态向量; σ 和 $\tanh$ 分别为 Sigmoid 激活函数和双曲正切激活函数。

在本文的非侵入式负荷监测任务中, 双向 LSTM 被选用以增强模型的上下文感知能力。正向 LSTM 和反向 LSTM 分别从前向和反向读取序列, 最终将两者的输出结合, 从而为后续的解码器提供全面的上下文信息。双向 LSTM 在处理具有复杂时间依赖关系的电器功率数据时, 能够显著提高预测精度。

2.3 自注意力机制

自注意力机制是近年来在深度学习中取得广泛应用的一种机制^[22], 尤其在处理长时间序列数据时表现出色。不同于传统的卷积或递归网络, 自注意力机制能够对输入序列中的每个元素分配不同的权重, 从而使模型能够关注到序列中最重要的部分。

自注意力机制的计算方式如下。

$$\text{Attention} = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (9)$$

式中: Q 、 K 、 V ($Q, K, V \in \mathbb{R}^{m \times d}$) 分别为查询矩阵、键矩阵和值矩阵, 由输入序列所有点的 $query$ 、 key 和 $value$ 组成; m 和 d 分别为输入序列的长度和输入序列各点的维度。

为了放大设备开启状态在总信号中的贡献, 本文使用了一个的简化注意力模型, 该模型使用它的相对重要性组合了编码器的所有隐藏状态。注意力机制可以表述为式 (10)~(12)。

$$\alpha_t = \text{softmax}(e_t) \quad (10)$$

$$e_t = Q_a^T \tanh(V_a h_t + T_a) \quad (11)$$

$$c = \sum_{t=1}^T \alpha_t h_t \quad (12)$$

式中: Q_a 、 V_a 和 T_a 为与架构的其他组件共同学习的注意力参数; α_t 为注意力权重; e_t 为注意力分数; h_t 为输入序列中的第 t 个向量; c 为注意力单元的输

出向量,并用作解码器的输入向量。

对于本文非侵入式负荷监测问题,自注意力机制能够很好的弥补CNN和双向LSTM对于长时间序列特征提取的不足,自注意力机制特别适用于捕捉长时间序列中的关键特征,如电器的启动关闭或工作状态的变化(如洗衣机的洗涤和甩干等)。

2.4 图卷积神经网络

在图卷积神经网络中,节点的特征表示是通过聚合其邻居节点的特征来更新的^[23]。这种聚合操作考虑了节点之间的连接关系,使得每个节点的特征表示都包含了其邻居节点的信息。

对于给定的无向图结构用 $G=(V, E)$ 来表示, V 为网络中所有的节点, E 表示节点之间的边,假设其中有 N 个节点,每个节点都有自己的 D 个特征,故节点的特征组成一个 $N \times D$ 维的矩阵 X ,各个节点之间的关系也会形成一个 $N \times N$ 维的矩阵 A ,这个 A 就被称为邻接矩阵, X 和 A 就是模型的输入。

图卷积神经网络是一个神经网络层,它层与层之间的传播关系公式如下。

$$H^{(l+1)} = \sigma(\tilde{D} - \frac{1}{2}\tilde{A}\tilde{D} - \frac{1}{2}H^{(l)}W^{(l)}) \quad (13)$$

式中: $\tilde{A} = A + I$, I 为单位矩阵; $\tilde{D}$ 为 $\tilde{A}$ 的度矩阵;公式为 $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$; W 为每层的权重矩阵; H 为每一层的特征,对于输入层, H 就是 X ; l 为第 l 层; σ 为非线性激活函数。使用上述过程,在构建的图

上创建了聚合信号的丰富嵌入,其中包含来自其相邻节点的时不变信息。

本文模型利用图卷积神经网络强大的特征提取能力来捕捉电器间的相互关系,从而增强模型的性能。其中,节点为电器的功率,边为电器间的相关性。图卷积神经网络的图结构能够捕捉到多电器同时工作的情况,从而增强模型对复杂场景的适应能力。

2.5 基于多算法融合的非侵入式负荷监测模型

模型采用回归网络和分类网络相结合的架构^[24],用于解决非侵入式负荷监测问题。模型结构如图3所示。

模型的输入来自家庭电力系统的总功率时间序列数据。数据首先经过预处理,以保证输入数据的质量和稳定性。为了处理时间序列数据并捕捉局部特征,模型采用滑动窗口技术,通过滑动窗口方法,长时间序列被切分为多个子序列,为后续模型的训练提供了局部特征数据。接下来,数据同时输入给回归网络和分类网络。

回归网络用于负荷分解,通过预测每个滑动窗口内电器的具体功率值来实现负荷分解。回归网络由编码器、自注意力机制和解码器组成。

编码器部分首先通过4个一维卷积层(每个包含 F 个过滤器,步长为1)从原始时间序列中提取局部特征(如电器启动时的短时功率峰值),然后采用包含 H 个隐藏单元的双向长短期记忆网络捕捉时间

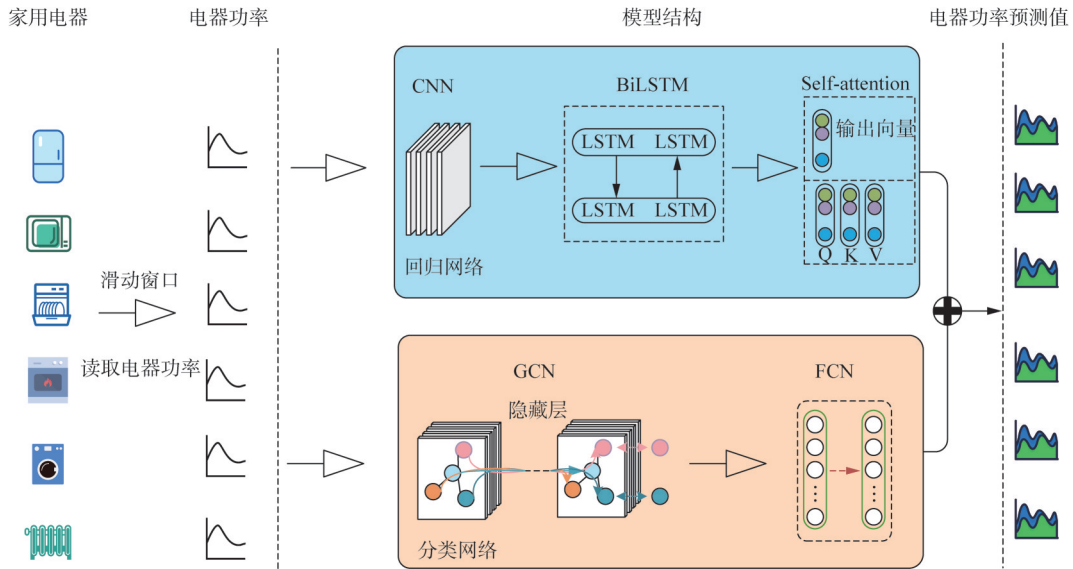


图3 本文模型结构图

Fig. 3 Model structure diagram of this article

序列中的长期依赖关系,通过双向时序处理学习电器功率变化的动态模式(如从关闭到开启),双向长短期记忆网络通过整合时序信息为功率预测提供丰富上下文。自注意力机制采用单层前馈神经网络计算注意力权重,对时间序列中特定关键时刻的信息赋予更高的权重,通过加强关键信息的表示,帮助预测出更准确的功率值。最终,自注意力机制输出的加权平均结果作为上下文向量传递给解码器。解码器部分通过两个全连接层将提取的时序特征映射为各电器的功率预测值。

分类网络用于负荷状态识别,判定各电器的运行模式(如“开启”、“关闭”或其他多种工作状态),以便为功率预测提供状态校正,其核心由图卷积神经网络和全连接网络构成。图卷积神经网络将滑动窗口内的电器视为图节点,节点特征由功率数据特征构成,通过计算每个滑动窗口内的功率曲线的相关性作为相关性阈值建立邻接矩阵。图卷积神经网络通过三阶段处理实现多负荷关联性建模:首先在节点表示阶段嵌入电器特征,接着通过图卷积操作聚合邻居节点信息(例如同时工作的微波炉与电炉灶会相互增强状态特征),最终生成包含全局上下文信息的节点表示。全连接网络将 GCN 输出的高维特征展平后,通过 softmax 函数输出开启或关闭的状态分类概率。这种基于图结构的处理能捕捉到当微波炉工作时,电炉灶大概率同时启动等复杂关联模式。

最终的模型是由回归网络和分类网络的输出相乘得出的。这种结合回归和分类的双网络架构使得模型能够同时考虑电器的功率预测与工作状态的识别,从而得到更加准确的非侵入式负荷监测结果。

3 算例分析

为了验证模型的有效性,本文选用的 REDD 数据集中的低频部分作为实验数据,本文模型使用 Python 语言和 torch 1.7.1 深度学习框架实现,模型的训练与测试在 NVIDIA GTX 4060Ti GPU 上进行。

3.1 数据集

REDD 数据集^[25]包含美国 6 所房屋的几个月的用电数据,在 REDD 低频数据集中,每 1 s 记录一次整个房屋的总功率,在不同的时间段内每 3 s 记录一次各种设备的功率。数据集包含房屋总功率和

10~25 种电器的测量真实值。

实验选取了 REDD 数据集中的冰箱、洗碗机、微波炉、电炉灶、洗衣机、电暖器这 6 种常见家用电器作为实验对象,涵盖了家庭用电的各个方面,包括保鲜、烹饪、清洁、取暖等。

实验选用的 6 种电器使用频率和功率特性各异。冰箱等设备的使用频率高,其功率曲线相对稳定,这有助于模型学习电器运行的基础特征;洗碗机和微波炉等设备的功率波动较大,并且存在多种工作模式,这能够有效训练模型识别复杂功率曲线的能力。此外,电炉灶等设备的功率曲线呈现出明显的脉冲式特征,这对于模型学习电器启动和停止的瞬间特征具有重要价值。通过分析这些电器的功率曲线,模型能够有效提取输入信号中的关键特征。这不仅有助于实现对不同类型电器的准确识别和功率预测,还能显著提升模型的泛化能力,使其在面对不同家庭用电模式时表现出更好的适应性和鲁棒性。数据集中 6 种电器的使用功率占比如图 4 所示。

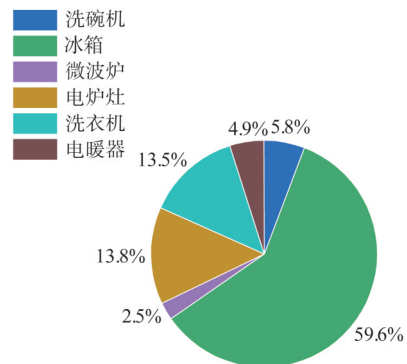


图 4 数据集电器使用功率占比

Fig. 4 Proportions of power used by electrical appliances in the dataset

尽管原始的 REDD 数据集提供了丰富的家庭用电数据,但其中仍存在问题,例如部分数据缺失、数据不平衡以及采样频率不一致等,这些问题都可能对模型的训练效果产生不利影响。因此,本文参考了文献[26]中的预处理方法并进一步采用了插值、过采样等技术对数据进行清洗和增强。预处理后电器功率曲线如图 5 所示。

针对数据中出现的缺失值问题,本文采用基于时间序列的插值方法对缺失数据进行填补,确保每个电器的功率数据保持连续性。对于数据中的异常

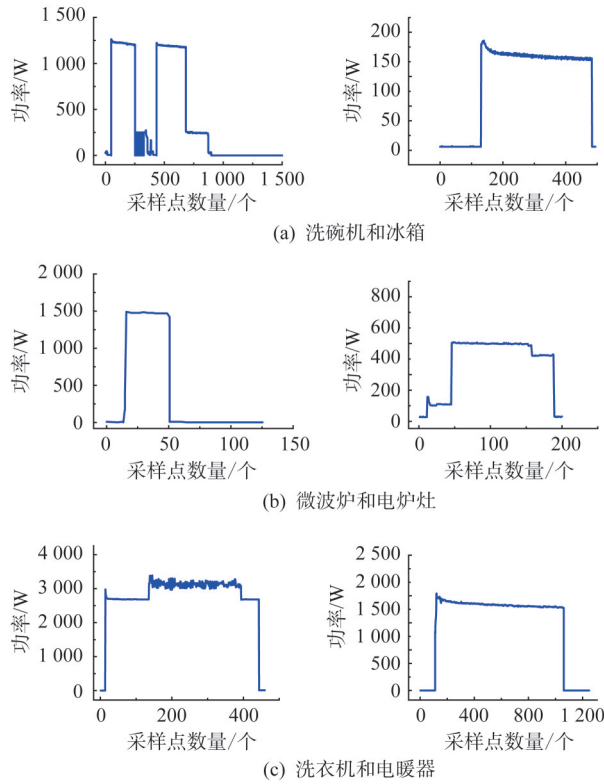


图5 电器预处理后效果图

Fig. 5 Effect diagram of electrical appliances after preprocessing

值,本文采用滑动窗口阈值法进行处理。当滑动窗口内某段电器功率值的变化超出设定范围时,将当前功率值替换为前一个正常值,从而保证数据的完整性和连续性。此外,针对数据不平衡的问题,本文采用过采样技术,通过复制活跃窗口的方式增加数据集中活动状态电器的数量。

为了验证模型的泛化能力,实验选取了不同房屋的电器数据进行训练和测试。具体划分见表1。

表1 训练集和测试集划分

Tab. 1 Split of training and test sets

电器名称	训练集	测试集
微波炉	房屋1、房屋5	房屋2
冰箱	房屋3、房屋5	房屋2
洗碗机	房屋2、房屋5	房屋4
电炉灶	房屋1、房屋2	房屋6
洗衣机	房屋3、房屋5	房屋2
电暖器	房屋1、房屋6	房屋5

3.2 评价指标和损失函数

为了验证本文模型的性能,本文使用决定系数(R^2)、信号聚合误差(S_{AE})、绝对平均误差(M_{AE})来

评估模型预测的效果。单个电器*i*对应的 R^2 、 S_{AE} 、 M_{AE} 定义如下所示。

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (14)$$

$$S_{AE} = \frac{|\sum_{i=1}^N \hat{y}_i - \sum_{i=1}^N y_i|}{\sum_{i=1}^N y_i} \quad (15)$$

$$M_{AE} = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (16)$$

式中: $\hat{y}_i$ 为预测功率值; y_i 为真实功率值; N 为总的功率点的数量; M_{AE} 、 S_{AE} 和 R^2 为评估模型性能的重要指标,它们从不同角度反映了模型的预测效果。 M_{AE} 直接衡量了预测值与真实值之间的平均绝对差异,反映了模型的预测精度,值越小,表示模型对信号的预测越准确; S_{AE} 为一种特定于信号聚合误差的度量,它考虑了信号在时间序列上的聚合效应; R^2 为衡量模型拟合优度的指标, R^2 越接近1,表明模型的拟合度越好。

模型训练时的损失函数 $\mathcal{L}$ 为回归网络的损失函数 $\mathcal{L}_{out}$ 与分类网络损失函数 $\mathcal{L}_{clas}$ 的和。其中 $\mathcal{L}_{out}$ 为网络总输出与单个设备的真实值之间的均方误差(M_{SE}), $\mathcal{L}_{clas}$ 为二进制交叉熵(B_{CE}),用于测量分类网络的开/关状态的分类误差。

$$M_{SE} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (17)$$

$$B_{CE} = -\frac{1}{N} \sum_{i=1}^N (y_i \times \log(p(y_i)) + (1 - y_i) \times \log(1 - p(y_i))) \quad (18)$$

式中: y_i 为二元标签,取值0或1; $p(y_i)$ 为模型预测第*i*个样本标签值为1的概率。

3.3 训练设置

在训练过程中,网络采用了Adam优化算法,训练的周期为5个epoch,初始学习率设置为0.001。模型网络在训练过程中接收包含64个样本的批次数据。

为了从功率时间序列数据中提取有效特征,本文采用滑动窗口方法将数据切分成多个固定大小的片段。每个片段包含一定时间范围内电器的功率数据,窗口大小根据电器的具体类型和功率波动特性进行设置。对于功率波动频繁的电器(如微波炉、

洗衣机等),采用较小的窗口,能够有效捕捉电器状态切换时的快速功率变化;对于功率波动较小且长期稳定运行的电器(如冰箱、电暖器等),则采用较大的窗口,以便更好地捕捉长期的功率变化模式。步长设置为1个数据点,确保窗口之间的连续性,从而完整地读取整个功率时间序列数据。滑动窗口长度设置见表2。

表2 不同电器滑动窗口的长度

Tab. 2 Length of sliding window for different appliances

电器名称	洗碗机	冰箱	微波炉	电炉灶	洗衣机	电暖器
长度 L	1 500	496	128	200	256	1 250

滑动窗口通过将长时间序列数据分割成多个子序列,帮助模型提取电器的局部特征,尤其是在电器启动、停止和状态切换等关键时刻。每个子序列作为模型的输入,捕捉电器功率的局部波动,为负荷分解和电器状态识别提供有力支持。通过调整窗口大小,模型能够适应不同电器的功率曲线特征,提高整体性能和泛化能力。在多个电器同时工作时,滑动窗口还能够帮助模型识别功率叠加的情况,捕捉电器间的相互作用和共存状态,进而提高负荷分解的精度。

为了进一步优化模型性能,本文对网络中的超参数进行了细致的调整,包括滤波器数量(F)、GCN隐藏层维度(G)和LSTM隐藏层的神经元数量(H)。通过网格搜索方法对这些超参数进行了全面优化,并为每类电器单独建模。最终选定了在验证集上表现最佳的参数组合。参数组合见表3。

表3 模型参数设置

Tab. 3 Model parameter settings

电器名称	滤波器数量 F	GCN隐藏层维度 G	LSTM隐藏层神经元数量 H
微波炉	8	8	1 024
冰箱	32	16	1 024
洗碗机	16	16	512
电炉灶	16	16	512
洗衣机	16	32	1 024
电暖器	8	8	512

3.4 实验结果分析

为了验证本文提出模型的优越性,模型与BiLSTM、SGN、GRU、Seq2Seq、VAE、CNN_Attention模型进行了对比分析。本文模型使用预处理

后的数据进行训练。表4展示了本文模型和对比模型在REDD数据集上的实验结果。其中性能最优的模型用粗体表示。

从实验结果看,本文提出的模型在REDD数据集中的6种电器的 M_{AE} 值、 S_{AE} 值和 R^2 指标上均优于所有对比模型。本文模型在 M_{AE} 值评价指标上优于最优对比模型SGN模型9.56%;在 S_{AE} 值上优于最优对比模型VAE模型11.63%;在 R^2 上优于最优对比模型SGN模型26%。本文模型在微波炉和洗碗机(多状态电器)上的 M_{AE} 值分别为12.38和10.24,显著低于SGN模型的16.91和11.20。这些改进主要得益于图卷积神经网络模块对电器间关联性的有效建模,尤其是在多个电器同时工作时,图卷积神经网络显著提高了负荷分解的准确性。尽管长短期记忆网络在建模长时间序列依赖关系方面表现良好,但在预测单状态电器的简单波动时效果有限。因此,本文模型将图卷积神经网络和长短期记忆网络结合起来,并引入自注意力机制关注重要特征,从而在处理不同类型电器负荷监测问题上表现出优越的泛化能力和有效性。

在对比实验中,BiLSTM模型结构相对简单,导致其特征提取能力有限,故在所有电器上的 M_{AE} 值、 S_{AE} 值和 R^2 指标均较差,尤其是在处理具有复杂功率曲线的电器时(如洗碗机和洗衣机),其预测精度更低。SGN模型通过多任务学习在处理多状态电器上表现良好,但在处理功率波动小的电器(如冰箱)时,因信息相似度高而过度依赖局部信息,导致性能受限。GRU模型通过门控机制缓解了RNN模型的梯度消失问题,但在面对复杂功率曲线时效果仍然较差。Seq2Seq模型在处理长序列数据方面表现较好。但其在处理具有多工作状态和复杂功率曲线的电器时(如冰箱和电暖器),其预测精度较低。VAE模型是一种基于深度学习的生成模型,能够学习数据的分布并生成新的数据。但其在处理非侵入式负荷监测任务时,其训练过程相对复杂,需要大量数据进行训练,且模型的泛化能力也相对较弱,这导致VAE模型在冰箱上的 M_{AE} 值、 S_{AE} 值指标较差。CNN_Attention模型融合了卷积神经网络和自注意力机制,能够有效提取时间序列数据中的局部特征和全局特征。但是,CNN_Attention模型使用的CNN无法很好地处理具有不同时间尺度的电器事件(如电炉灶的脉

冲式功率),导致其在处理具有多工作状态和复杂功率曲线的电器时特征提取能力不足,最终导致预测精度下降。

图6展示了本文模型预测的负荷分解曲线。可以看出模型对这6种电器的预测值和真实值相近,表现出良好的适应性,能够根据不同电器的功率特征进行准确的预测。本文模型对多状态电器(如洗碗机和微波炉)的性能提升最为显著,能够准确捕捉功率曲线中的关键状态变化点(如启动、模式切换)。这表明图卷积神经网络通过建模邻接关系有效捕捉了电器之间的复杂关联性,自注意力机制弥补了卷积神经网络在长时间序列特征提取中的不足,从而提升了对多状态电器功率变化的预测能力。相比之下,对于冰箱和电暖器等单状态电器,模型性能提升有限,这是由于这类电器功率曲线波动较小,图卷积神经网络的关联性建模作用不明显。

因此,本文提出的模型融合了图卷积神经网络^[30]、卷积神经网络^[31]、自注意力机制以及双向长短期记忆网络^[32-33]等多种方法。这些技术的结合使得模型能够更全面地提取电器负荷特征,充分捕捉电器间的复杂关联性,从而提高负荷分解的精度。

然而,在训练集和验证集上的模型表现普遍优于测试集,这主要是由于训练数据与测试数据之间存在分布差异。本文使用REDD数据集进行模型训练与测试,该数据集涵盖了多个不同家庭的电力使用数据。不同家庭的电器类型、使用频率、运行时长等因素可能存在显著差异,这些因素会影响电器的功率消耗模式,从而导致电器功率特征的多样性。由于训练集中包含的家庭数量有限,模型可能未能学习到足够多样化的负载特征模式,因此在面对与训练数据差异较大的测试数据时,模型的泛化能力受到限制,进而导致预测误差的增大。

3.5 消融实验

为了进一步分析模型中不同模块对预测结果的影响,本文进行了消融实验。结果如表5所示。

由表5可知,完整模型在3个评价指标上均表现最佳。移除BiLSTM后,模型性能显著下降,这表明双向LSTM在捕捉时间序列的长期依赖关系方面具有重要作用,能够有效降低预测误差并提升模型的拟合能力。移除self-attention机制后,模型误差增大,这说明self-attention机制有助于模型关注关键信息,提高对复杂功率曲线的识别能力。移除GCN后,模型性能也有所下降,这表明GCN在建

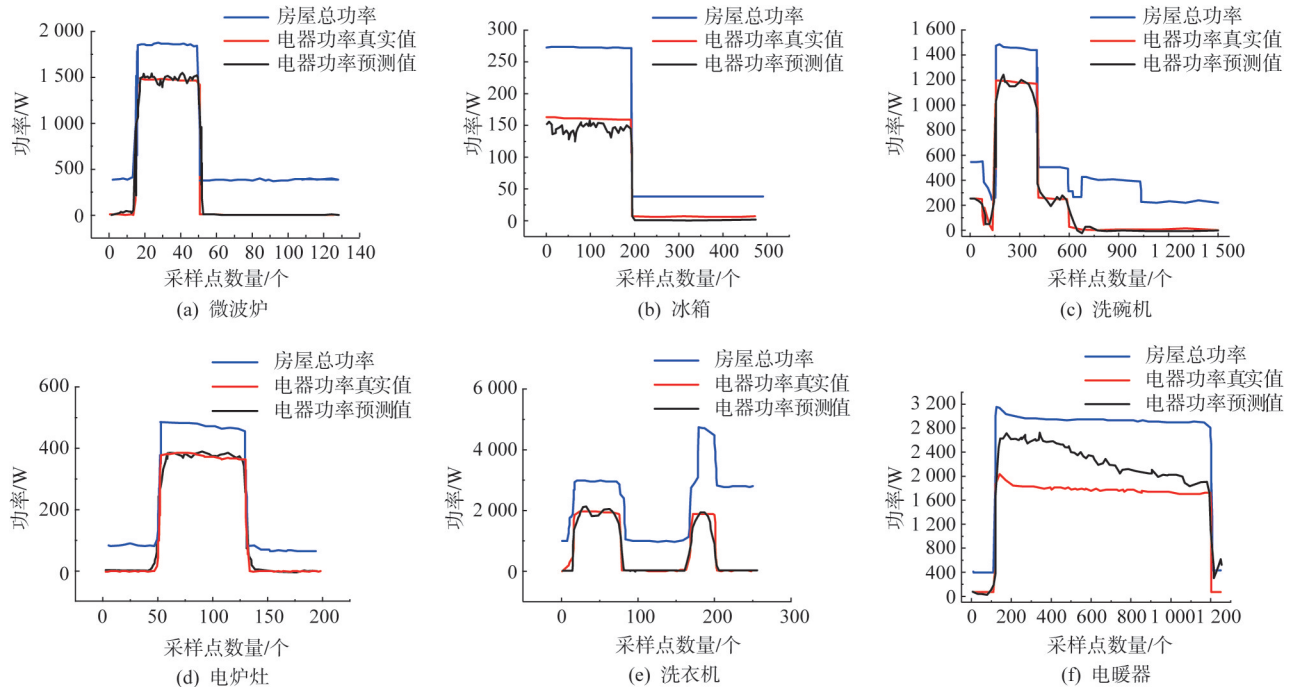


图6 本文提出模型预测曲线图

Fig. 6 Model prediction curves diagram in this article

表 4 模型对比结果

Tab. 4 Model comparison results

模型	评价指标	洗碗机	冰箱	微波炉	电炉灶	洗衣机	电暖器	平均
BiLSTM ^[14]	M_{AE}	39.90	79.43	49.82	37.97	35.82	42.65	47.60
	S_{AE}	0.65	0.71	0.55	0.62	0.58	0.68	0.63
	R^2	-0.90	-0.14	-0.90	-0.87	-0.11	0.03	-0.48
SGN ^[26]	M_{AE}	11.20	38.16	16.91	13.64	19.83	18.79	19.76
	S_{AE}	0.31	0.45	0.52	0.38	0.50	0.47	0.44
	R^2	0.67	0.46	0.43	0.62	0.40	0.44	0.50
GRU ^[27]	M_{AE}	18.41	78.43	33.48	24.85	25.64	23.42	34.04
	S_{AE}	0.40	0.68	0.60	0.49	0.51	0.48	0.53
	R^2	0.29	0.06	-0.11	0.32	0.25	0.28	0.18
Seq2Seq ^[15]	M_{AE}	13.24	40.85	22.74	18.64	21.61	27.39	24.08
	S_{AE}	0.35	0.50	0.58	0.47	0.53	0.60	0.51
	R^2	0.43	0.13	0.24	0.38	0.30	0.20	0.28
VAE ^[28]	M_{AE}	12.73	42.35	16.68	14.52	20.63	21.57	21.41
	S_{AE}	0.32	0.47	0.47	0.38	0.45	0.48	0.43
	R^2	0.42	0.22	0.53	0.47	0.35	0.32	0.39
CNN_Attention ^[29]	M_{AE}	26.86	40.09	17.73	19.21	19.25	20.11	23.88
	S_{AE}	0.45	0.55	0.50	0.52	0.51	0.55	0.51
	R^2	-0.22	0.33	0.39	0.36	0.38	0.34	0.26
本文模型	M_{AE}	10.24	37.06	12.38	11.76	17.23	18.52	17.87
	S_{AE}	0.23	0.37	0.46	0.35	0.42	0.44	0.38
	R^2	0.72	0.61	0.65	0.69	0.58	0.55	0.63

表 5 消融实验结果

Tab. 5 Results of ablation

模块	M_{AE}	S_{AE}	R^2
完整模型	17.87	0.38	0.63
移除 BiLSTM	23.95	0.49	0.29
移除 self-attention	19.98	0.45	0.51
移除 GCN	19.76	0.44	0.49

模电器间关联性方面具有积极作用，能够增强模型对多电器同时工作场景的适应性。

综上所述，模型中的各个模块均对任务具有重要优势，这些模块的综合作用使得模型在本文非侵入式负荷监测任务中表现出更好的性能。

4 结语

本文针对非侵入式负荷监测任务提出了一种基于多算法融合的非侵入式负荷监测模型，有效

解决了现有模型在处理多状态电器预测困难的问题，展现出更高的准确性和泛化能力，模型融合了多种深度学习网络：图卷积神经网络用于建模电器之间的复杂关联性，捕捉电器间的能耗关系；卷积神经网络负责提取时间序列数据中的局部特征，例如电器启动时的功率变化；自注意力机制进一步强化了模型对长时间序列关键特征的提取能力，确保模型能准确关注到重要信息；双向长短期记忆网络通过同时处理正向和反向的时间依赖关系，增强了对时间序列全局特性的建模能力。通过数据预处理去除异常值、噪声等干扰，减少数据中的噪声信息，从而提高模型训练的效率和收敛速度。通过 REDD 数据集进行实验对比。实验结果表明，本文模型在 REDD 数据集 6 种电器的功率预测方面优于所有对比模型。这证明了本文模型在处理不同类型电器负荷监测问题上的通用性和有效性。

通过非侵入式负荷监测模型获取电力信息后，如何将数据应用到实际是一个问题。后续研究将致力于将模型应用到智慧家居等实际情景中，以实现节能降耗、提高能源利用效率等目标。

参考文献

- [1] 丁一, 惠红勋, 林振智, 等. 面向电力需求侧主动响应的商业模式及市场框架设计[J]. 电力系统自动化, 2017, 41(14): 2-9.
DING Yi, HUI Hongxun, LIN Zhenzhi, et al. Business model and market framework design for active response to electricity demand side [J]. Automation of Electric Power Systems, 2017, 41(14): 2-9.
- [2] 何恒靖, 王灏, 肖勇, 等. 基于 Bi-LSTM 算法的非侵入式负荷监测模型[J]. 南方电网技术, 2019, 13(2): 20-26.
HE Hengjing, WANG Hao, XIAO Yong, et al. Non-intrusive load monitoring model based on Bi-LSTM algorithm [J]. Southern Power System Technology, 2019, 13(2): 20-26.
- [3] 周润, 向月, 王杨, 等. 基于智能电表集总数据的家庭电动汽车充电行为非侵入式辨识与负荷预测[J]. 电网技术, 2022, 46(5): 1897-1906.
ZHOU Run, XIANG Yue, WANG Yang, et al. Non-intrusive identification and load forecasting of household electric vehicle charging behavior based on smart meter data [J]. Power System Technology, 2022, 46(5): 1897-1906.
- [4] 余昊杨, 武昕, 郭一凡, 等. 基于事件感知的钢厂工业用户非侵入式负荷检测分解算法[J]. 南方电网技术, 2022, 16(11): 29-36.
YU Haoyang, WU Xin, GUO Yifan, et al. Non-intrusive load detection decomposition algorithm for industrial users of steel mill based on event perception [J]. Southern Power System Technology, 2022, 16(11): 29-36.

- [5] 张石清, 王伟, 钱亚冠, 等. 面向深度学习的非侵入式负荷监测研究进展[J]. 计算机系统应用, 2023, 32(3): 25 – 47.
ZHANG Shiqing, WANG Wei, QIAN Yaguan, et al. Advances in non-intrusive load monitoring research for deep learning[J]. Journal of Computer Systems Application, 2023, 32(3): 25 – 47.
- [6] 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.
DENG Xiaoping, ZHANG Guiqing, WEI Qinglai, et al. Review of non-intrusive load monitoring[J]. IEEE/CAA Journal of Automatica Sinica, 2022, 48(3): 644 – 663.
- [7] HART G W. Non intrusive appliance load monitoring[J]. Proceedings of the IEEE, 1992, 80(12): 1870 – 1891.
- [8] LIN Y H, TSAI M S. Development of an improved timefrequency analysis-based nonintrusive load monitor for load demand identification[J]. IEEE Transactions on Instrumentation and Measurement, 2014, 63(11): 1470 – 1483.
- [9] 赵文清, 张诗满, 李刚. 基于聚类和相关分析的居民用户非侵入式负荷分解[J]. 电力自动化设备, 2020, 40(6): 8 – 19.
ZHAO Wenqing, ZHANG Shiman, LI Gang. Non-intrusive load decomposition of residential users based on cluster and association analysis[J]. Electric Power Automation Equipment, 2020, 40(6): 8 – 19.
- [10] 于超, 覃智君, 阳育德. 基于启停状态识别改进因子隐马尔可夫模型的非侵入式负荷分解[J]. 电网技术, 2021, 45(11): 4540 – 4550.
YU Chao, QIN Zhijun, YANG Yude. Non-intrusive load disaggregation with improved factorial hidden Markov model considering ON-OFF status recognition[J]. Power System Technology, 2021, 45(11): 4540 – 4550.
- [11] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527 – 1554.
- [12] 张淑清, 李君, 姜安琦, 等. 基于FPA-VMD和BiLSTM神经网络的新型两阶段短期电力负荷预测[J]. 电网技术, 2022, 46(8): 3269 – 3279.
ZHANG Shuqing, LI Jun, JIANG Anqi, et al. A novel two-stage model based on FPA-VMD and BiLSTM neural network for short-term power load forecasting[J]. Power System Technology, 2022, 46(8): 3269 – 3279.
- [13] 黄济宇, 管霖, 郭梦轩, 等. 图深度学习技术在智能暂态稳定评估中的应用及展望[J]. 电网技术, 2023, 47(4): 1500 – 1511.
HUANG Jiyu, GUAN Lin, GUO Mengxuan, et al. Application and prospect of graph deep learning technique in intelligent transient stability assessment[J]. Power System Technology, 2023, 47(4): 1500 – 1511.
- [14] KELLY J, KNOTTENBELT W. Neural NILM: deep neural networks applied to energy disaggregation[C]//Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, July 11 – 13, 2015, Seoul, South Korea. New York: ACM, 2015: 55 – 64.
- [15] ZHANG C, ZHONG M, WANG Z, et al. Sequence-to-point learning with neural networks for non-intrusive load monitoring[C]//The 32nd AAAI Conference on Artificial Intelligence, February 2 – 3, 2018, Louisiana. Louisiana: AAAI Press, 2018: 2604 – 2611.
- [16] GOMES E, PEREIRA L. PB-NILM: Pinball guided deep nonintrusive load monitoring[J]. IEEE Access, 2020(8): 48386 – 48398.
- [17] AKBAR M K, AMAYRI M, BOUGUILA N. A novel non-intrusive load monitoring technique using semi-supervised deep learning framework for smart grid[J]. Building Simulation, 2024, 17(3): 441 – 457.
- [18] ATHANASOULIAS S, SYKIOTIS S, KASELIMI M, et al. A first approach using graph neural networks on non-intrusive-load-monitoring[C]//Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments, June 13 – 15, 2022, Korea. Corfu Greece: AICAS, 2022: 601 – 607.
- [19] KASELIMI M, DOULAMIS N, DOULAMIS A, et al. Bayesian-optimized bidirectional LSTM regression model for non-intrusive load monitoring[C]//ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing, May 12 – 17, Brighton, United Kingdom. Brighton: ICASSP, 2019: 2747 – 2751.
- [20] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735 – 1780.
- [21] CONNOR J T, MARTIN R D, ATLAS L E. Recurrent neural networks and robust time series prediction[J]. IEEE Transactions on Neural Networks, 1994, 5(2): 240 – 254.
- [22] STAVROS S, MARIA K, ANASTASIOS D, et al. ELECTRICity: An Efficient Transformer for Non-Intrusive Load Monitoring[J]. Sensors, 2022, 22(8): 2926 – 2926.
- [23] PICCIALLI V, SUDOSO A M. Improving non-intrusive load disaggregation through an attention-based deep neural network[J]. Energies, 2021, 14(4): 847.
- [24] YANG Y, GUAN Z, LI J, et al. Interpretable and efficient heterogeneous graph convolutional network[J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 35(2): 1637 – 1650.
- [25] KOLTER J Z, JOHNSON M J. REDD: a public data set for energy disaggregation research[C]//Workshop on Data Mining Applications in Sustainability. San Diego: DMESS, 2011: 59 – 62.
- [26] SHIN C, JOO S, YIM J, et al. Subtask gated networks for non-intrusive load monitoring[C]//Proceedings of the AAAI Conference on Artificial Intelligence, January 27-February 1, 2019, Honolulu, USA. Honolulu: AAAI Press, 2019: 1150 – 1157.
- [27] CHUNG J, GÜLÇEHRE Ç, CHO K, et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling[J]. CoRR, 2014, abs/1412.3555.
- [28] LANGEVIN A, CHERIET M, et al. Energy Disaggregation using Variational Autoencoders[J]. Energy and Buildings, 2022, (254): 1 – 10.
- [29] XU X H, ZHAO S H, CUI K B. Non-intrusive load disaggregate algorithm based on convolutional block attention module[J]. Power System Technology, 2021, 45(9): 3700 – 3705.
- [30] 杨建, 常学军, 姚帅, 等. 基于WT-CNN-BiLSTM模型的日常光伏功率预测[J]. 南方电网技术, 2024, 18(8): 61 – 69.
YANG Jian, CHANG Xuejun, YAO Shuai, et al. Day-ahead photovoltaic power forecasting based on WT-CNN-BiLSTM model[J]. Southern Power System Technology, 2024, 18(8): 61 – 69.