

DOI: 10.19783/j.cnki.pspc.211238

基于改进 GMM-CNN-GRU 混合的非侵入式 负荷监测方法研究

杨秀¹, 李安¹, 孙改平¹, 田英杰², 刘方¹, 潘瑞媛¹, 吴吉海¹

(1. 上海电力大学电气工程学院, 上海 200090; 2. 华东电力科学研究院, 上海 200437)

摘要: 为挖掘用户侧节能减排潜力, 对用户用电行为进行精细化分析和管理工作, 提升电能利用效率, 提出了一种基于高斯混合模型聚类 and 深度神经网络相结合的非侵入式负荷监测方法。首先, 针对同一电器常出现功率相近但运行状态不一致问题, 利用高斯混合模型聚类算法中“软分类”和类簇灵活的优势, 对负荷工作状态进行精细分类, 形成负荷用电设备实际运行情况的负荷状态特征库。其次, 针对常见的应用于非侵入式负荷监测模型的深度神经网络在多标签分类时存在识别精度低等问题, 提出卷积神经网络与门控循环单元混合的深度神经网络模型。最后, 综合考虑外部环境数据对家庭用户用电习惯的影响, 在 AMPds2 数据集上开展验证分析, 并与其他模型进行对比。结果表明, 所提的非侵入式负荷监测模型具有较高的准确性。

关键词: 非侵入式负荷监测与分解; 高斯混合模型聚类; 卷积神经网络; 门控循环单元; 深度学习

Non-invasive load monitoring based on an improved GMM-CNN-GRU combination

YANG Xiu¹, LI An¹, SUN Gaiping¹, TIAN Yingjie², LIU Fang¹, PAN Ruiyuan¹, WU Jihai¹

(1. School of Electric Power Engineering, Shanghai University of Electric Power, Shanghai 200090, China;

2. East China Electric Power Research Institute, Shanghai 200437, China)

Abstract: A non-intrusive load monitoring method based on Gaussian mixture model clustering combined with a deep neural network is proposed to explore the potential of energy saving and emission reduction at the customer side. It also fine-tunes the analysis and management of customer electricity consumption behavior, and improves the efficiency of electricity use. First, we tackle the problem that the same electrical appliance often has similar power but inconsistent operating status. In order to classify the load working status in fine manner, the advantages of "soft classification" and flexible clustering in the Gaussian mixture model clustering algorithm can be used to form a load status feature library that conforms to the actual operating conditions of electrical equipment. Secondly, note that in the common deep neural networks applied to non-invasive load monitoring models, there are problems such as low recognition accuracy in multi-label classification. Thus a deep neural network model with a mixture of convolutional neural networks and gated recurrent units is proposed. Finally, the validation analysis is carried out on the AMPds2 dataset by considering the influence of external environmental data on the energy consumption habits of household users, and the results are compared with other models. The results show that the proposed non-invasive load monitoring model has high accuracy.

This work is supported by the National Natural Science Foundation of China (No. 61872230).

Key words: non-invasive load monitoring and decomposition; Gaussian mixture model clustering; convolutional neural networks; gated recurrent unit; deep learning

0 引言

随着全球气候变暖、能源争夺加剧以及环境污

染日益严重, 为了降低能源消耗, 改善电力系统效率, 为用户提供优质服务, 高级量测体系(Advanced Metering Infrastructure, AMI)等应运而生^[1]。其中负荷监测是 AMI 最重要的应用之一, 其主要任务是对负荷进行分解、归类, 有针对性地管理用户用电行为, 实现用户精细化用能管理, 提升能源利用效率,

基金项目: 国家自然科学基金项目资助(61872230); 上海电力人工智能工程技术研究中心研究项目资助(19D72252800)

降低用能成本, 对于降低碳排放具有重要意义^[2-5]。

上世纪 80 年代 Hart 教授首次提出非侵入式负荷监测(Non-intrusive Load Monitoring, NILM)概念^[6], 并基于事件监测^[7]方法对用户电力入口处记录的总功率信息进行分解, 进而获得用户用电设备能耗情况及用电规律, 但受限于当时技术问题精度难以提高。随着人工智能及高级量测技术的发展, 学者开始将新技术应用于新能源出力预测^[8-9]、负荷预测^[10-13]与监测^[14-17]等研究中。2015 年 Kelly 教授首次将深度学习(Deep Learning)算法应用于 NILM, 相比于传统算法具有高精度、易于扩展等特点^[18]。目前许多学者将神经网络方法应用于 NILM^[19-21]。

将神经网络应用于 NILM 问题的关键在于, 如何构建负荷工作状态特征库作为神经网络的监督数据, 多数学者采用聚类算法进行构建。文献^[22]提出基于改进 K-Means 聚类的电器状态分类方法, 通过挖掘低频数据中电器功率特征消除冗余信息。但用电设备在不同的工作状态下常常会出现功率大小相近的情况, 而 K-Means 算法难以解决该问题。这是由于 K-Means 属于“硬分配”, 不同类别的划分严格依靠距离关系, 故无法精准划分不同工作状态。为解决“硬分类”的局限, 提高负荷状态分类精细度, 本文基于高斯混合模型(Gaussian Mixture Model, GMM)聚类构建负荷运行特征库的方法。GMM 聚类算法按群集概率分配群集成员, 属于“软分类”, 其灵活的类簇可有效避免“硬分类”的局限性, 面向大规模数据集能同时保证聚类效率以及聚类精度^[23-25]。

另一方面, 更精细化的负荷状态特征库对深度神经网络也提出更高的性能要求。常见应用于非侵入式负荷监测模型的深度神经网络有多层感知器(Multilayer Perceptron, MLP)、卷积神经网络(Convolutional Neural Networks, CNN)等。虽 CNN 神经网络能够挖掘非连续性特征在高维空间中的联系, 其对图片的分类性能在计算机视觉领域得到充分认可, 但在识别具有时序性的多标签数据时精度较低^[26]。长短时记忆网络(Long Short-term Memory Networks, LSTM)是在序列传递方向上进行递归处理的一种神经网络, 非常适合处理具有时序性的功率数据。门控循环单元(Gated Recurrent Unit, GRU)是一种基于 LSTM 的优化网络, 其简化了 LSTM 的内部单元结构, 在拥有 LSTM 网络优点的同时可减少梯度的消失, 有效缩短模型的训练时间, 因此更适用于负荷辨识^[27]。本文设计 CNN 与 GRU 分别作为子模块的改进 CNN-GRU 深度神经网络结构。为满足负荷分解需求, 将一个神经网络模型分为两个输出, 其一为负荷工作状态, 其二为负荷工作功率。

最终通过 GMM 聚类算法建立的设备运行状态特征库作为工作状态输出的监督, 分项计量的负荷功率作为负荷功率输出的监督, 形成一个基于改进 GMM-CNN-GRU 混合的非侵入式负荷监测模型。在 AMPds2 数据集^[28]上进行试验, 并与 CNN、LSTM 和常规 CNN-GRU 神经网络比较, 验证了该方法负荷状态监测和功率分解的精度。

1 NILM 模型

1.1 GMM-CNN-GRU 混合的 NILM 模型流程图

本文采用基于改进 GMM-CNN-GRU 混合的 NILM 模型流程图, 如图 1 所示。首先通过 GMM 聚类算法对各电器运行历史数据进行状态划分, 形成负荷状态特征库; 然后将外部环境数据和总功率数据输入基于改进 CNN-GRU 深度神经网络模型, 负荷状态特征和负荷详细分项计量数据作为模型训练监督量, 进行模型训练, 最终达到各类负荷的精准监测和分解。

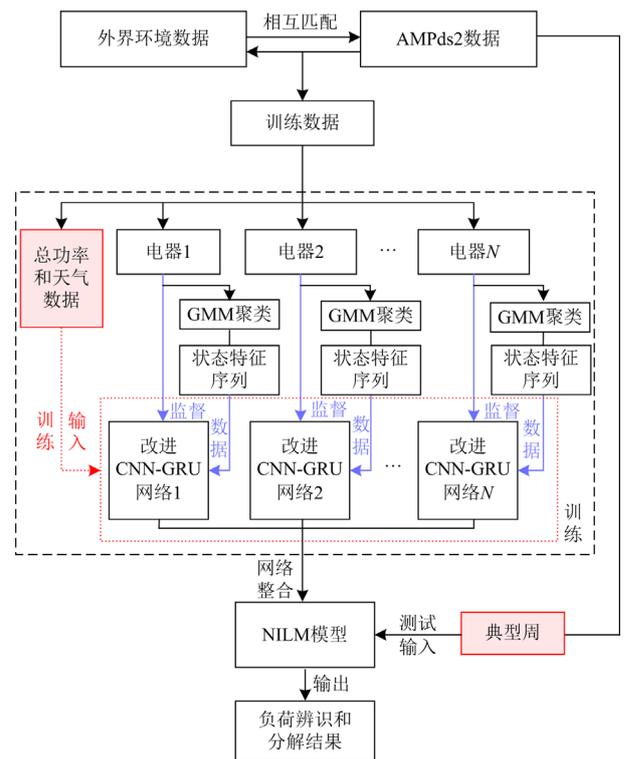


图 1 NILM 框架图

Fig. 1 NILM framework diagram

1.2 高斯混合模型聚类

负荷状态特征库的形成依赖于 GMM 聚类, 其中高斯混合模型是用若干个高斯概率密度函数(正态分布曲线)来描述一个事物的模型。对于一个样本来说, 混合高斯分布得到的是其属于各个类的概率。

求解高斯混合模型参数一般采用期望极大算^[29] (Expectation Maximization algorithm, EM), 其步骤如下。

Step1: 已知样本集 $D = \{x_1, x_2, \dots, x_m\}$, 高斯混合成分个数 k , 即聚类中心数, 可定义高斯混合分布函数为

$$P(x) = \sum_{i=1}^k \alpha_i \cdot p(x | \mu_i, \Sigma_i) \quad (1)$$

式中: $P(\cdot)$ 为概率密度函数; $p(x | \mu_i, \Sigma_i)$ 为高斯混合模型的第 i 个概率分布模型; μ_i 和 Σ_i 分别为第 i 个高斯混合成分的均值和协方差; α_i 为混合系数, 满足 $\alpha_i \geq 0$ 且 $\sum_{i=1}^k \alpha_i = 1$, $i = 1, 2, \dots, k$ 。

Step2: 根据贝叶斯定理, 求出后验分布概率, 如式(2)。

$$\gamma_{ji} = \frac{\alpha_i \cdot p(x_j | \mu_i, \Sigma_i)}{\sum_{n=1}^k \alpha_n \cdot p(x_j | \mu_n, \Sigma_n)} \quad (2)$$

式中: $j = 1, 2, \dots, m$; $i = 1, 2, \dots, k$ 。

Step3: 计算新一轮迭代的模型参数。

$$\mu'_i = \frac{\sum_{j=1}^m \gamma_{ji} x_j}{\sum_{j=1}^m \gamma_{ji}} \quad (3)$$

$$\Sigma'_i = \frac{\sum_{j=1}^m \gamma_{ji} (x_j - \mu'_i)(x_j - \mu'_i)^2}{\sum_{j=1}^m \gamma_{ji}} \quad (4)$$

$$\alpha'_i = \frac{\sum_{j=1}^m \gamma_{ji}}{m} \quad (5)$$

Step4: 重复 Step2 和 Step3 直至收敛或达到最大迭代轮数。

Step5: 将样本集 D 划分 k 个簇, 每个样本簇标记 λ_j 由式(6)确定。

$$\lambda_j = \arg \max_{i \in \{1, 2, \dots, k\}} \gamma_{ji} \quad (6)$$

高斯混合模型聚类算法流程图如图 2 所示。

GMM 聚类算法对每个用电设备进行聚类分析, 以有功功率作为工作状态划分的依据, 最终形成对应应用电设备的负荷状态特征库。将负荷状态特征进行 One-hot 编码, 作为改进 CNN-GRU 神经网络学习训练的监督数据。

1.3 改进 CNN-GRU 网络模型

深度神经网络凭借其强大的表达能力在分类方面有巨大优势。直接将常规 CNN-GRU 网络应用

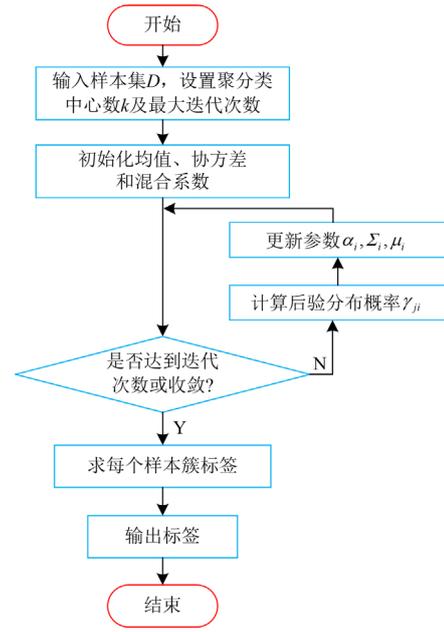


图 2 高斯混合模型聚类算法流程图

Fig. 2 Gaussian mixture model clustering flow chart

于 NILM 问题不能取得良好的效果, 为了提高深度神经网络的性能, 可增加网络层数和每层神经元数目, 但这并不能有效提升监测精度, 反而会增加训练复杂度。

本文将多个不同的神经元模块横向“并联”排布, 分别获得非连续性输入数据的非线性特征联系、高维空间特征联系以及时序特征联系, 最终将各个网络获得的特征张量进行叠加后, 传入逻辑回归单元, 形成一个改进 CNN-GRU 深度神经网络。该结构中包含以下内容。

1) Dense 模块

神经元采用全连接方式, 其中每个神经元传播公式为

$$\mathbf{H} = \mathbf{W}_h \cdot \mathbf{X} + \mathbf{b}_h \quad (7)$$

式中: \mathbf{H} 为隐藏层输出; \mathbf{W}_h 为权重矩阵; \mathbf{X} 为前一层的输出; \mathbf{b}_h 为偏差参数。

2) 一维卷积模块

卷积神经网络(CNN)在图像处理中的应用广泛。由于图像中有色彩、尺寸等多维特征, 需要挖掘图像中的空间特征, 故采用卷积网络结构, 卷积核多为 2 维。而在 NILM 问题中, 功率数据为时间序列类型数据, 所以本文采用一维卷积方式, 计算公式为

$$\mathbf{S} = \mathbf{X} * \mathbf{g} = \sum_{n=1}^{N+M-1} \sum_{i=1}^N \mathbf{X}(i) \mathbf{g}(n-i) \quad (8)$$

式中： X 为卷积层的输入序列； g 为卷积核； N 为输入序列 X 的长度； M 为卷积核的长度； S 为卷积结果序列，长度为 $N + M - 1$ 。

3) GRU 模块

门控循环单元 GRU 是对长短时记忆网络的一种改进，参数更少，训练速度更快。GRU 中包含重置门 r_t 和更新门 z_t ，其计算公式为

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (9)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (10)$$

式中： h_{t-1} 为前一时刻隐藏状态； x_t 为当前时刻的输入； σ 为激活函数 Sigmoid。GRU 单元的输出为当前时刻的隐藏状态 h_t ，其计算公式为

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (11)$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [r_t * h_{t-1}, x_t]) \quad (12)$$

本文搭建的神经网络分为两大部分，一部分实现用电设备的状态监测，另一部分实现用电设备的功率分解，其结构设计如图 3 所示，超参数设置如下。

(1) 总输入层：将输入数据设置成维度 $N \times 1 \times 7$ 的张量(N 为输入序列的长度)，时间序列长度为 1，

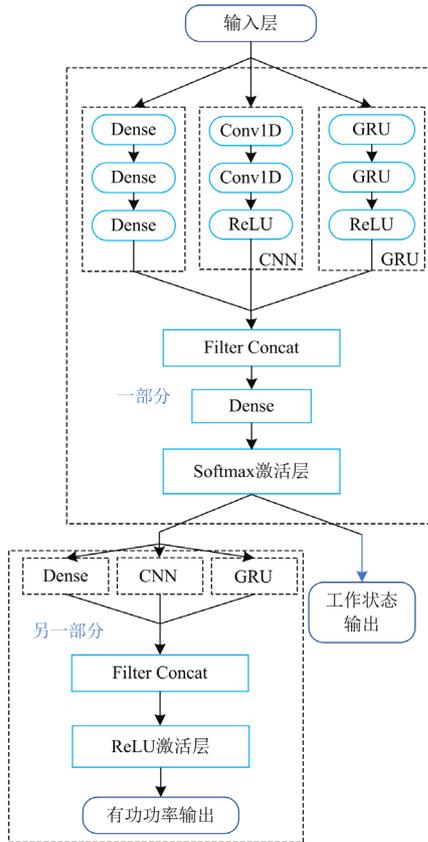


图 3 网络结构

Fig. 3 Network structure

输入特征维度为 7。

(2) 改进 CNN-GRU 层，总计 352 个神经元。

(3) 全连接层：64 个神经元。

(4) 激活层：激活函数为 Softmax。

(5) 第一部分输出层：输出维度为 I (I 为用电器的状态数，即聚类中心数 k)，得到该用电器所处的状态(One-hot 编码形式)。

(6) 第二部分输入层：由于输入数据为第一部分的输出结果，所以输入层维度设置为 $N \times 1 \times I$ 。

(7) 改进 CNN-GRU 模块，总计 352 个神经元。

(8) 全连接层：64 个神经元。

(9) 激活层：激活函数为 ReLU。

(10) 第二部分输出层：输出层维度为 1 维，为用电器的有功功率。

2 模型训练与评估

2.1 数据集选取

文献[30]采用国外公开数据集 REDD 以及实验室测量私人数据集(国内社区采集)对所设计的模型进行验证分析，其实验结果表明应用于非侵入式负荷监控的模型在国内外家庭用户层面的差异影响较小。由于国内有关家庭用电设备详细公开数据集很少，目前国内绝大多数学者采用国外公开数据集作为模型研究验证的对象，故本文选取国外公开数据集。

本文选取 2016 年加拿大学者发布的 AMPds2 数据集，该数据采集自加拿大温哥华市区内一个家庭用户，居住总面积 199 m²，地下室面积 100 m²，有功功率的采样间隔为 1 min，外部环境数据采样间隔为 1 h，数据集的采集时间跨度从 2012 年 4 月 1 日至 2014 年 4 月 1 日，总计 1 051 200 条记录。

2.2 训练策略

本文搭建的神经网络结构分为两大部分：状态监测和功率分解。一部分实现用电设备的状态监测。由于多状态用电设备的各个状态工作时长存在差异，导致采样到的数据量存在差异，此问题属于不均衡分类问题，故对训练数据进行随机过采样方式扩充数据集。在扩充数据集中选取用电设备总有功功率数据 $P = \{P_{t-4}, P_{t-3}, P_{t-2}, P_{t-1}, P_t\}$ 以及外界环境数据 $E = \{E_{t-4}, E_{t-3}, E_{t-2}, E_{t-1}, E_t\}$ ，作为时刻 t 的网络输入，网络的输出则为 t 时刻用电设备的运行状态 S_t (One-hot 形式)，选择处理多分类问题的分类交叉熵作为损失函数。另一部分实现功率的分解，将第一部分监测出的用电设备状态作为输入，输出为用电设备功率，损失函数选用均方误差 MSE。两部分均采用 Adam 优化器，学习率为 0.001。

2.3 评价指标

NILM 有很多评价指标, 本文选择针对状态监测的指标: 准确度(precision)、召回率(recall)和 F1 分数(F1-score)。选择针对功率分解的指标: 平均绝对误差(Mean Absolute Error, MAE)。

1) 状态监测指标

首先分别计算三个统计量 TP、FP、FN, 其中 TP 为真阳性: 状态监测为正, 实际为正的个数; FP 为假阳性: 状态监测为正, 实际为负的个数; FN 为假阴性: 状态监测为负、实际为正的个数。

然后计算 $e_{precision}$ 、 e_{recall} 、 $e_{F1-score}$, 可表示为^[31]

$$e_{precision} = \frac{TP}{TP + FP} \quad (13)$$

$$e_{recall} = \frac{TP}{TP + FN} \quad (14)$$

$$e_{F1-score} = 2 \times \frac{e_{precision} \cdot e_{recall}}{e_{precision} + e_{recall}} \quad (15)$$

式中: $e_{precision}$ 为准确度; e_{recall} 为召回率; $e_{F1-score}$ 为 F1 分数。这三个指标值越接近 1, 说明模型精度越高。

2) 功率分解指标

平均绝对误差(MAE)计算公式为^[31]

$$e_{MAE} = \frac{1}{T} \sum_{t=1}^T |\hat{y}_t - y_t| \quad (16)$$

式中: \hat{y}_t 为 t 时刻的功率分解值; y_t 为 t 时刻的功率实际值; T 为时长。 e_{MAE} 越小, 说明模型精度越高。

3 实例分析

3.1 家庭用电设备相关因素分析

为接近我国国情, 从 AMPds2 数据集中选择 7 种典型家庭用电设备: 热泵(暖通空调)、干衣机、冰箱、电视、地下照明、洗衣机和洗碗机, 基本包含国内家庭大功率用电设备种类。日常生活中外界的环境温度与室内电器的使用存在很大相关性, 故本文通过计算 Pearson 相关系数^[32]分析 AMPds2 数据集中用电设备的总功率以及 7 种典型设备有功功率与外界环境(气温、气压、风速、能见度)之间的相关性, 此处环境数据与电器功率数据在时间及颗粒度上保持一致, 如图 4 所示。

由于加拿大全年温度较低, 而夏季温度较为适宜, 所以室内温控系统(热泵)与气温呈负相关。同样该家庭电视、冰箱、地下照明和洗衣机的使用与外界环境存在一定的相关性。表 1 为天气数据对本文所设计的 NILM 模型精度的影响情况。

表 1 中可看出部分与外部环境数据相关性较大的电器(热泵、冰箱、电视、地下照明和洗衣机),

在加入天气因素后其监测精度都得到明显提升, 可见天气数据为负荷精度的重要影响因素之一, 环境数据的获取并不困难, 具有可行性。

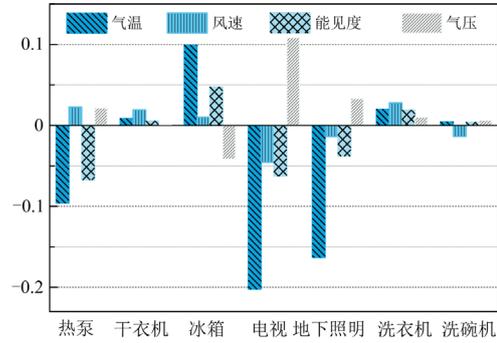


图 4 用电设备与环境特征的相关性

Fig. 4 Correlation between power-using equipment and environmental characteristics

表 1 天气数据对精度的影响对比表

Table 1 Comparison table of influence of weather

	data on accuracy						
e_{MAE}	热泵	干衣机	冰箱	电视	地下照明	洗衣机	洗碗机
不含天气	12.88	2.14	16.51	16.00	12.08	6.57	5.71
包含天气	7.31	1.35	10.16	9.57	8.74	4.44	4.75
精度提升	5.57	0.79	6.35	6.43	3.34	2.13	0.96

3.2 用电设备运行状态提取

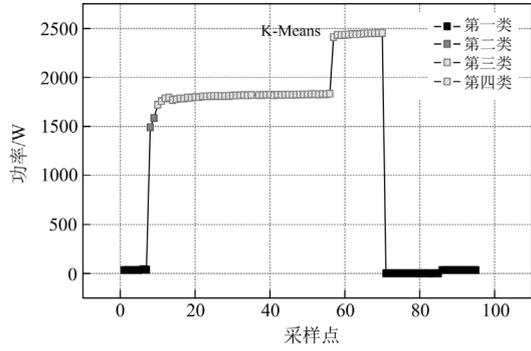
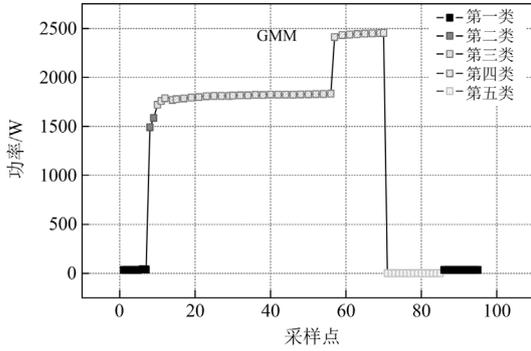
利用高斯混合模型聚类算法对用电设备有功功率进行挖掘, 获取用电设备运行状态的特征信息, 并形成用电设备运行状态特征库。由于 7 种典型用电设备都属于多工作状态设备, 所以依据实际功率数据以及戴维森堡丁指数确定聚类数目 k 。将 GMM 聚类效果与文献[22]中所用 K-Means 聚类效果相比较。

设置高斯混合模型聚类最大迭代次数为 100 次, 收敛阈值为 6×10^{-5} , 初始化中心点次数为 5。图 5 为 GMM 与 K-Means 的聚类结果比较, 分别为热泵、洗碗机、电视、冰箱和洗衣机的有功功率聚类结果, 其中左侧图为 GMM 聚类结果, 右侧为 K-Means 聚类结果。

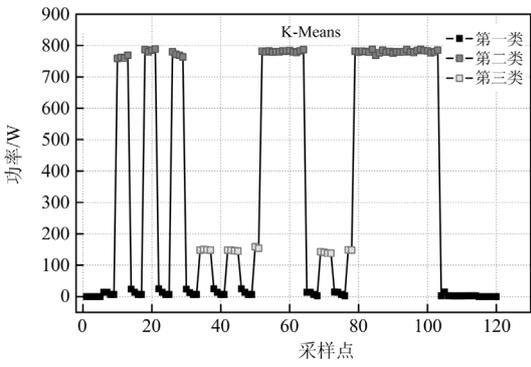
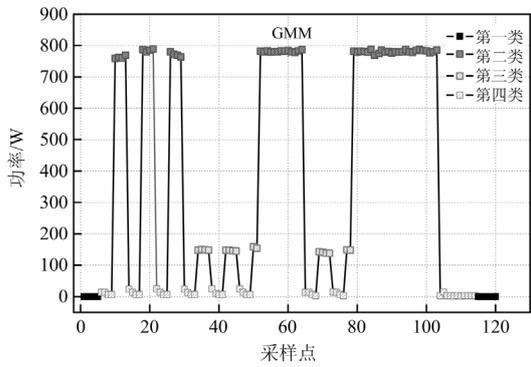
在图 5(a)中, GMM 聚类将热泵工作过程分为五类, 即待机状态功率约为 37 W、启动状态功率由 37 W 突变到 1800 W、平稳运行状态约为 1800 W、大功率运行状态功率约为 2400 W 及关闭状态 0 W, 而 K-Means 聚类将关闭状态 0 W 与待机状态 37 W 归为一类, 明显不符合实际; 图 5(b)中洗碗机被分为四类, 即关闭状态功率 0 W、待机状态功率约为 11 W、大功率工作状态功率约为 775 W 和低功率工作状态功率约为 140 W, 同样 K-Means 聚类算法将

洗碗机待机状态与关闭状态归为一类。K-Means 聚类在电视、冰箱和洗衣机上也存在类似现象，GMM

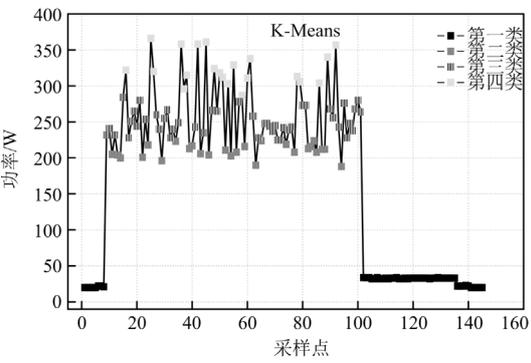
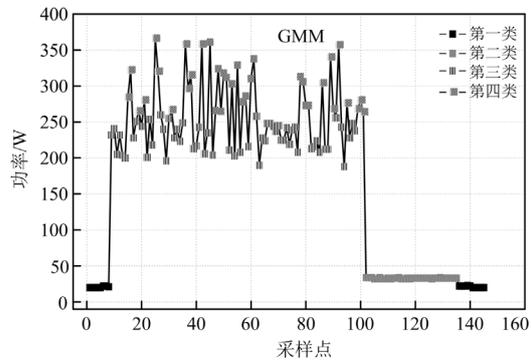
聚类算法结果精细度更高，更符合实际，聚类结果如表 2 所示。



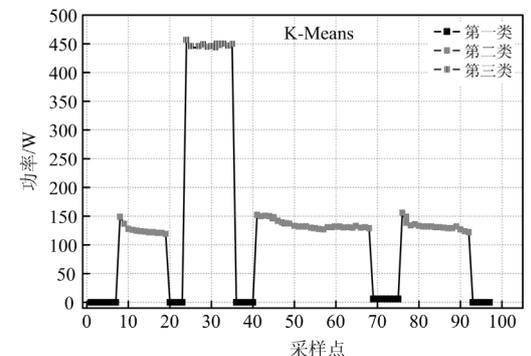
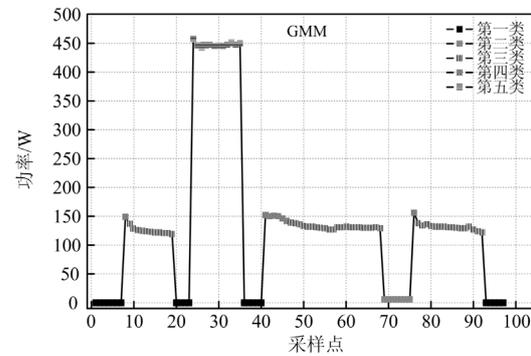
(a) 热泵



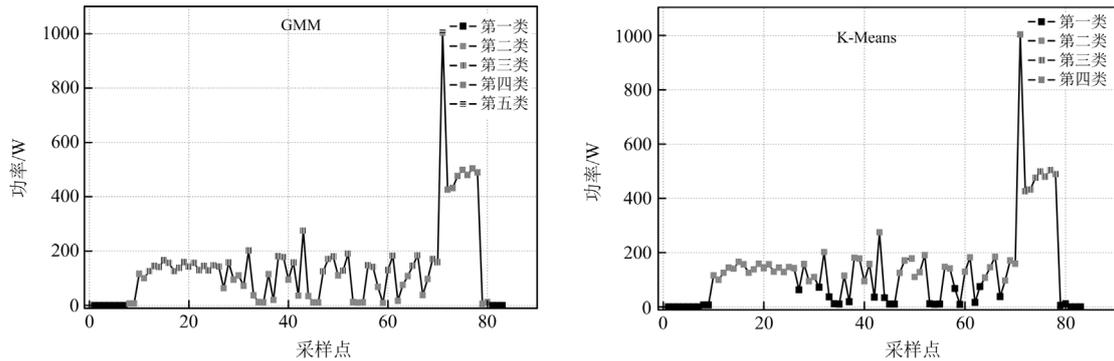
(b) 洗碗机



(c) 电视



(d) 冰箱



(e) 洗衣机

图 5 聚类结果比较

Fig. 5 Comparison of clustering results

表 2 电器运行状态编号以及相关特征

Table 2 Electrical operating status number and related characteristics

用电器名称	聚类中心数 k	状态编号	采集数据量
干衣机	3	0	980 555
		1	10 882
		2	6738
电视	5	0	15 260
		1	879 323
		2	16 333
		3	70 809
		4	16 351
洗碗机	4	0	961 166
		1	14 450
		2	18 405
		3	4154
地下照明	3	0	853 948
		1	42 801
		2	101 426
洗衣机	5	0	970 229
		1	12 930
		2	11 642
		3	3036
		4	338
热泵	5	0	46 966
		1	877 182
		2	6772
		3	60 816
		4	6439
冰箱	5	0	546 948
		1	84 877
		2	327 469
		3	34 717
		4	4164

将 GMM 聚类算法与 K-Means 聚类算法所得负荷状态特征库, 分别送入神经网络进行训练, 最终负荷功率分解结果对比如表 3 所示。

表 3 不同聚类算法对比

Table 3 Comparison of different clustering algorithms

e_{MAE}	洗衣 机	洗碗 机	地下 照明	电视	冰箱	干衣 机	热泵
K-Means	7.75	9.46	10.43	19.67	18.96	2.39	11.00
GMM	4.44	4.75	8.74	9.57	10.16	1.35	7.31
精度提升	3.31	4.71	1.69	10.1	8.8	1.04	3.69

从表 3 中可看出, GMM 聚类形成的负荷状态特征库在最终负荷监测精度上有明显优势, 电视、冰箱的负荷监测精度提升较高, 可见更精细化的负荷状态特征有利于提升负荷监测的性能。

3.3 负荷监测与分解结果

将训练数据输入所搭建的改进 CNN-GRU 深度神经网络中, 按训练策略进行训练, 负荷功率分解结果如图 6 所示。

从图 6 可以看出, 本文所提的负荷监测模型对用电设备的启停过程辨识精准, 功率分解效果好, 可满足 NILM 任务要求。将本文所提模型与卷积神经网络模型(CNN)^[8]、长短时记忆网络模型(LSTM)^[8]和常规 CNN-GRU 神经网络模型^[8]进行对比。比较过程中训练数据一致, 均选自基于 GMM 聚类得到的负荷状态特征库, 比较结果见表 4。

由表 4 可知, 本文所设计神经网络相比 CNN 和 LSTM 网络精度明显提升。由于热泵设备在该家庭的能耗占比为 23.23%, 工作功率在 1500 W 以上且稳定, 故改进 CNN-GRU 模型在功率分解上精度提升较小, 其他用电设备提升明显。图 7 中详细比较了改进 CNN-GRU 与常规 CNN-GRU 模型的状态监测和功率分解性能。

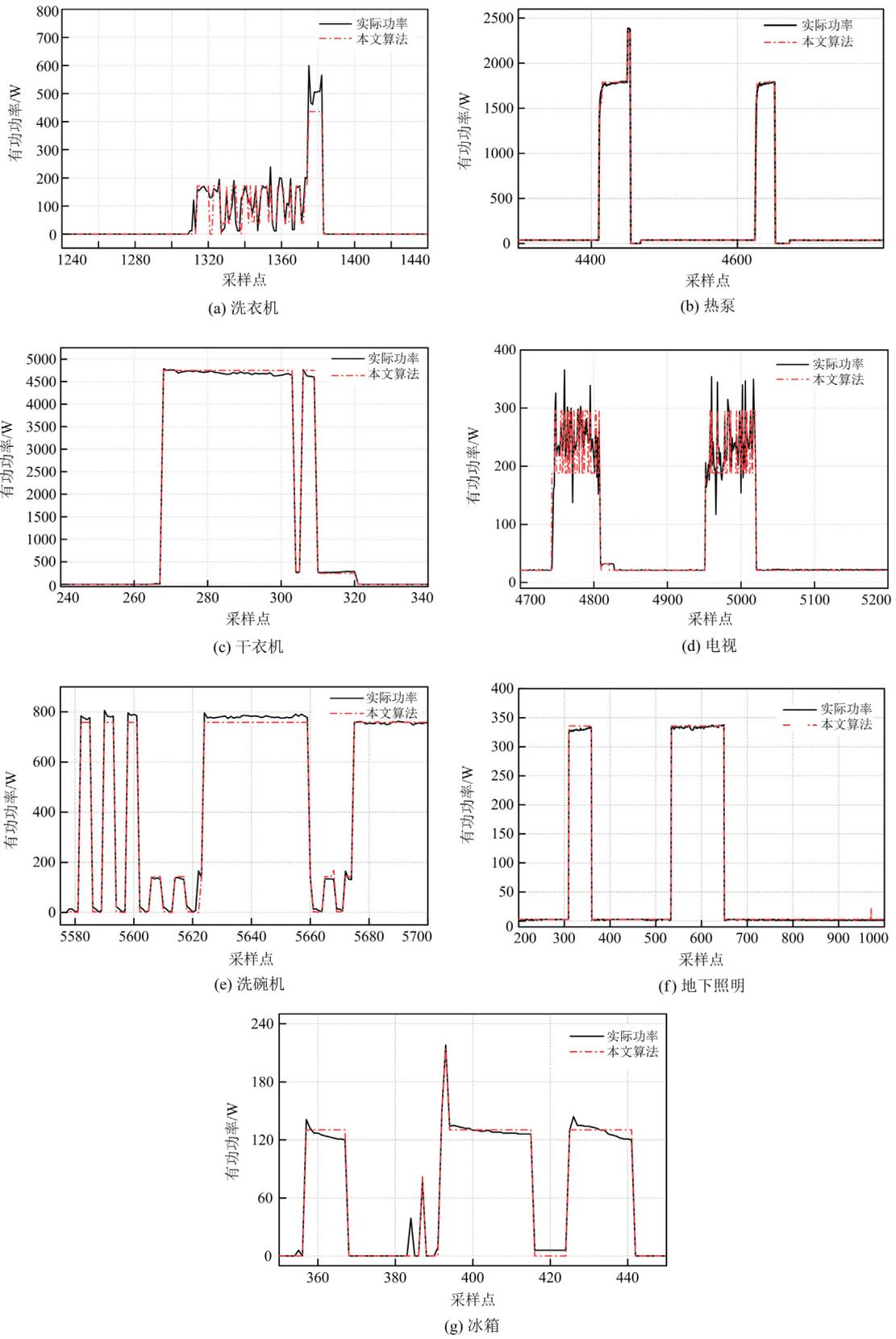


图 6 负荷功率分解结果

Fig. 6 Load power decomposition results

表 4 深度学习算法对比

Table 4 Deep learning algorithm comparison

e_{MAE}	热泵	干衣机	冰箱	电视	地下照明	洗衣机	洗碗机
CNN	10.41	1.93	18.89	16.30	9.95	7.81	7.66
LSTM	8.98	1.74	17.71	14.03	8.88	8.32	6.70
CNN-GRU	7.31	1.74	16.15	14.65	9.26	5.01	6.70
本文算法	7.02	1.35	10.16	9.57	8.74	4.44	4.75

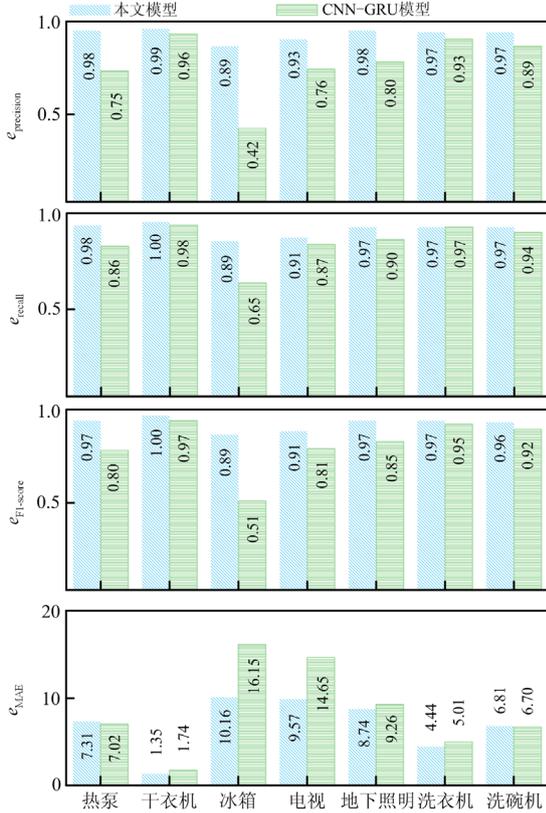


图 7 本文方法与 CNN-GRU 结果对比

Fig. 7 Comparison results of CNN-GRU and our algorithm

从图 7 可以看出, 针对功率较大、状态少且特征差异较大的设备, 本文所提算法比常规 CNN-GRU 算法识别负荷工作状态精度提升 5%~20%, 负荷工作功率分解精度接近, 提升 1%~2%。但针对状态较多的用电设备时, 本文算法精度更高, 工作状态监测精度提升 30%, 功率分解精度提升 6%。通过结果对比可知, 本文神经网络结构在处理多分类问题时优势显著。

3.4 模型应用能力验证

考虑在一年中不同的外界环境下, 家庭用户会有不同的用电场景, 各场景下各电器的使用习惯也会出现相应变化, 故为保证本文模型在全年时间段不同电器使用场景和不同用电习惯下都有较高的分解精度, 选择 AMPds2 数据集中春夏秋冬的典型周以及度假周的真实数据(不包含在训练集中)作为测

试集, 其中夏季周选择 2013 年 7 月 21 日—27 日, 期间热泵使用频率降低, 冰箱使用频率增大; 冬季周选择 2013 年 1 月 20 日—26 日, 期间热泵使用频率增大, 冰箱使用频率降低; 春秋两季电器使用较为均衡, 选择 2013 年 4 月 14 日—20 日; 度假周选择 2013 年 8 月 1 日—7 日, 期间只有热泵和冰箱维持基本运行。表 5 和表 6 分别展示了模型应对用户不同行为习惯时的状态监测与功率分解精度。

表 5 状态监测结果($e_{F1-score}$)

Table 5 Status recognition results ($e_{F1-score}$)

周类型	热泵	干衣机	冰箱	电视	地下照明	洗衣机	洗碗机
冬季周	0.991	0.998	0.912	0.934	0.976	0.964	0.975
春季周	0.990	0.998	0.898	0.939	0.984	0.973	0.971
夏季周	0.993	0.998	0.897	0.952	0.961	0.965	0.975
秋季周	0.990	0.999	0.920	0.920	0.977	0.966	0.974
度假周	0.998	1.000	0.914	0.971	1.000	1.000	1.000

表 6 功率分解结果(e_{MAE})

Table 6 Power decomposition accuracy (e_{MAE})

周类型	热泵	干衣机	冰箱	电视	地下照明	洗衣机	洗碗机
冬季周	4.91	1.59	11.60	8.76	7.76	4.31	4.43
春季周	6.92	1.43	12.87	8.01	7.39	3.69	5.39
夏季周	5.05	1.47	10.75	7.06	11.84	4.52	2.55
秋季周	4.62	0.84	9.61	11.08	9.70	3.77	3.95
度假周	5.70	0.19	4.88	1.51	2.41	0.05	0.00

由表 5 和表 6 可以看出, 本文模型在面对不同的用电场景时, 负荷工作状态监测误差均在 10%以内, 只有冰箱误差在 10%左右。该模型功率分解的精度维持在较高的水平。

4 结论

本文提出一种基于改进 GMM-CNN-GRU 混合的 NILM 方法。首先对大规模负荷数据采用 GMM 聚类算法获取负荷状态特征库, 然后输入改进 CNN-GRU 神经网络模型以实现非侵入式负荷状态监测与功率分解, 所提方法具有以下优势。

1) 相比于 K-Means 聚类结果, 本文所提 GMM 聚类方法的类簇更加灵活, 面对同一电器功率接近但状态迥异的情况, 可实现更加精准的分类, 且与电器运行状态贴合度更高。

2) 在相同的训练数据下, 本文所提改进 CNN-GRU 神经网络与 CNN 神经网络、LSTM 神经网络和常规 CNN-LSTM 神经网络相比, 在负荷状态监测和功率分解等方面具备更高的精度。

3) 本文主要选取了 AMPds2 数据集中符合我国国情的典型用电设备, 其中洗碗机和干衣机在国内

需求越来越高, 本文模型将其考虑在内, 符合未来家庭发展需要。热泵(暖通空调)在国内主要应用于公商业楼宇等设施, 本文所设计模型在应用于工商业用户时也有一定的竞争力。

本文模型基于大量侵入式测量数据训练得到, 而在实际应用中会面临侵入式用电设备功率数据不足的问题, 如何利用少量的数据达到 NILM 的精度要求以及针对国内家庭用电情况进一步验证模型将成为下一步的研究目标。

参考文献

[1] 张永生, 巢清尘, 陈迎, 等. 中国碳中和: 引领全球气候治理和绿色转型[J]. 国际经济评论, 2021(3): 1-18.
ZHANG Yongsheng, CHAO Qingchen, CHEN Ying, et al. China's carbon neutrality: leading in global climate governance and green transition[J]. International Economic Review, 2021(3): 1-18.

[2] 程祥, 李林芝, 吴浩, 等. 非侵入式负荷监测与分解研究综述[J]. 电网技术, 2016, 40(10): 3108-3117.
CHENG Xiang, LI Linzhi, WU Hao, et al. A survey of the research on non-intrusive load monitoring and disaggregation[J]. Power System Technology, 2016, 40(10): 3108-3117.

[3] 李叶茂, 李雨桐, 郝斌, 等. 低碳发展背景下的建筑“光储直柔”配用电系统关键技术分析[J]. 供用电, 2021, 38(1): 32-38.
LI Yemao, LI Yutong, HAO Bin, et al. Key technologies of building power supply and distribution system towards carbon neutral development[J]. Distribution & Utilization, 2021, 38(1): 32-38.

[4] 盛四清, 吴昊, 顾清, 等. 含碳捕集装置的电气综合能源系统低碳经济运行[J]. 电测与仪表, 2021, 58(6): 82-90.
SHENG Siqing, WU Hao, GU Qing, et al. Low-carbon economic operation of integrated electricity and natural gas system with carbon capture devices[J]. Electrical Measurement & Instrumentation, 2021, 58(6): 82-90.

[5] 白云, 贾燕冰, 陈浩, 等. 计及供气充裕性的电-气互联综合能源系统低碳经济调度[J]. 电测与仪表, 2021, 58(11): 32-38.
BAI Yun, JIA Yanbing, CHEN Hao, et al. Low-carbon economic dispatch of electric-gas interconnection systems considering gas supply adequacy[J]. Electrical Measurement & Instrumentation, 2021, 58(11): 32-38.

[6] HART G W. Nonintrusive appliance load monitoring[J]. Proceedings of the IEEE, 1992, 80(12): 1870-1891.

[7] 周东国, 张恒, 周洪, 等. 基于状态特征聚类的非侵入式负荷事件检测方法[J]. 电工技术学报, 2020, 35(21): 4565-4575.

ZHOU Dongguo, ZHANG Heng, ZHOU Hong, et al. Non-intrusive load event detection method based on state feature clustering[J]. Transactions of China Electrotechnical Society, 2020, 35(21): 4565-4575.

[8] BOZORG M, BRACALE A, CARAMIA P, et al. Bayesian bootstrap quantile regression for probabilistic photovoltaic power forecasting[J]. Protection and Control of Modern Power Systems, 2020, 5(3): 218-229.

[9] MADHIARASAN M. Accurate prediction of different forecast horizons wind speed using a recursive radial basis function neural network[J]. Protection and Control of Modern Power Systems, 2020, 5(3): 230-238.

[10] 陈弘川, 蔡旭, 孙国歧, 等. 基于智能优化方法的相似日短期负荷预测[J]. 电力系统保护与控制, 2021, 49(13): 121-127.
CHEN Hongchuan, CAI Xu, SUN Guoqi, et al. Similar day short-term load forecasting based on intelligent optimization method[J]. Power System Protection and Control, 2021, 49(13): 121-127.

[11] 刘云, 吴家宏. 计及电动汽车渗透率的台区负荷预测研究[J]. 电网与清洁能源, 2020, 36(1): 72-78.
LIU Yun, WU Jiahong. Research on forecasting of transformer area load considering the permeability of electric vehicles[J]. Power System and Clean Energy, 2020, 36(1): 72-78.

[12] 全新宇, 张宇泽, 张长生, 等. 基于广义回归神经网络的有源配电网网供负荷预测方法[J]. 供用电, 2020, 37(12): 40-45.
TONG Xinyu, ZHANG Yuze, ZHANG Changsheng, et al. Load forecasting method of active distribution network based on generalized regression neural network[J]. Distribution & Utilization, 2020, 37(12): 40-45.

[13] 李强, 赵健, 王磊, 等. 配电网馈线负荷预测及风险评估预警分析方法[J]. 供用电, 2020, 37(5): 50-55.
LI Qiang, ZHAO Jian, WANG Lei, et al. Distribution network feeder load forecasting and risk assessment warning analysis method[J]. Distribution & Utilization, 2020, 37(5): 50-55.

[14] 杨立余, 陈昊, 黎明, 等. 非侵入式电力负荷多目标分解框架[J]. 电力系统保护与控制, 2020, 48(6): 100-107.
YANG Liyu, CHEN Hao, LI Ming, et al. A framework for non-intrusive load monitoring using multi-objective evolutionary algorithms[J]. Power System Protection and Control, 2020, 48(6): 100-107.

[15] 张安安, 庄景泰, 郭红鼎, 等. 结合图半监督与广义回归神经网络的非侵入式海洋平台负荷监测[J]. 电力系统保护与控制, 2020, 48(7): 85-91.

- ZHANG An'an, ZHUANG Jingtai, GUO Hongding, et al. Non-intrusive offshore platform load monitoring based on graph-based semi-supervised learning and generalized regression neural networks[J]. *Power System Protection and Control*, 2020, 48(7): 85-91.
- [16] 朱永灿, 张鹏, 田毅, 等. 变电设备在线监测技术中的同步采集触发方法对比分析[J]. *中国电力*, 2022, 55(3): 64-73.
- HU Yongcan, ZHANG Peng, TIAN Yi, et al. Comparison of synchronous acquisition trigger methods in on-line monitoring technology of transformer equipment[J]. *Electric Power*, 2022, 55(3): 64-73.
- [17] 许易经, 韩学山, 杨明, 等. 基于设备在线监测的电网状态检修决策模型[J]. *电力系统自动化*, 2020, 44(23): 72-81.
- XU Yijing, HAN Xueshan, YANG Ming, et al. Decision-making model of condition-based maintenance for power grid with equipment on-line monitoring[J]. *Automation of Electric Power Systems*, 2020, 44(23): 72-81.
- [18] KELLY J, KNOTTENBELT W J. *Neural NILM: deep neural networks applied to energy disaggregation*[M]. Springer International Publishing, 2015.
- [19] VERMA S, SINGH S, MAJUMDAR A. Multi-label LSTM autoencoder for non-intrusive appliance load monitoring[J]. *Electric Power Systems Research*, 2021, 199.
- [20] 王轲, 钟海旺, 余南鹏, 等. 基于 seq2seq 和 Attention 机制的居民用户非侵入式负荷分解[J]. *中国电机工程学报*, 2019, 39(1): 75-83.
- WANG Ke, ZHONG Haiwang, YU Nanpeng, et al. Nonintrusive load monitoring based on sequence-to-sequence model with Attention mechanism[J]. *Proceedings of the CSEE*, 2019, 39(1): 75-83.
- [21] ZHOU X, FENG J, LI Y. Non-intrusive load decomposition based on CNN-LSTM hybrid deep learning model[J]. *Energy Reports*, 2021, 7: 5762-5771.
- [22] 燕续峰, 翟少鹏, 王治华, 等. 深度神经网络在非侵入式负荷分解中的应用[J]. *电力系统自动化*, 2019, 43(1): 126-132.
- YAN Xufeng, ZHAI Shaopeng, WANG Zhihua, et al. Application of deep neural network in non-intrusive load disaggregation[J]. *Automation of Electric Power Systems*, 2019, 43(1): 126-132.
- [23] LI Kehua, MA Zhenjun, ROBINSON D, et al. Identification of typical building daily electricity usage profiles using Gaussian mixture model-based clustering and hierarchical clustering[J]. *Applied Energy*, 2018, 231: 331-342.
- [24] Model-based clustering and Gaussian mixture model [EB/OL]. [2017-01-02]. <https://en.proft.me/2017/02/1/model-based-clustering-r/>.
- [25] 张美霞, 李丽, 杨秀, 等. 基于高斯混合模型聚类和多维尺度分析的负荷分类方法[J]. *电网技术*, 2020, 44(11): 4283-4296.
- ZHANG Meixia, LI Li, YANG Xiu, et al. A load classification method based on gaussian mixture model clustering and multi-dimensional scaling analysis[J]. *Power System Technology*, 2020, 44(11): 4283-4296.
- [26] KIM Y. *Convolutional neural networks for sentence classification*[Z]. 2014.
- [27] LIU S, YOU S, ZENG C, et al. Data source authentication of synchrophasor measurement devices based on 1D-CNN and GRU[J]. *Electric Power Systems Research*, 2021, 196: 107-207.
- [28] MAKONIN S, ELLERT B, BAJIC I V, et al. Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014[J]. *Scientific Data*, 2016, 3: 1-12.
- [29] 周志华. *机器学习*[M]. 北京: 清华大学出版社, 2016.
- [30] 何敏瑶. 非侵入式家庭用户负荷监测算法研究[D]. 西安: 西安理工大学, 2021.
- HE Minyao. *Research on household non-invasive load monitoring algorithm*[D]. Xi'an: Xi'an University of Technology, 2021.
- [31] MENGISTU M A, GIRMAY A A, CAMARDA C, et al. A cloud-based on-line disaggregation algorithm for home appliance loads[J]. *IEEE Transactions on Smart Grid*, 2019, 10(3): 3430-3439.
- [32] MOON T K. The expectation-maximization algorithm[J]. *IEEE Signal Processing Magazine*, 1996, 13(6): 47-60.
-
- 收稿日期: 2021-09-08; 修回日期: 2021-12-09
作者简介:
杨秀(1972—), 男, 教授, 博士生导师, 从事分布式发电与微电网技术; Email: yangxiu721102@126.com
李安(1997—), 男, 硕士研究生, 从事非侵入式负荷监测与分解; E-mail: spiderla97@163.com
孙改平(1984—), 女, 讲师, 从事电力系统优化调度。E-mail: sunfrog2002@163.com
- (编辑 许威)