

基于改进WGAN考虑特征分布相似性的小样本负荷预测方法

卢俊波¹, 刘俊峰¹, 罗燕¹, 曾君^{2†}

(1. 华南理工大学 自动化科学与工程学院, 广东 广州 510640; 2. 华南理工大学 电力学院, 广东 广州 510640)

摘要: 对于综合能源系统中新接入用户, 其往往由于历史数据匮乏而难以建立精准的短期负荷预测模型. 本文基于迁移学习理论, 提出了一种基于改进Wasserstein生成对抗网络(WGAN)的小样本负荷预测方法. 首先, 本文采用最大信息系数法量化各影响特征与负荷的相关性强弱. 接着, 将源域特征序列进行分割, 计算各分割子序列与目标域小样本的实序列编辑距离确定初始源域. 然后, 引入卷积神经网络和长短期记忆模型建立源域预测网络. 通过WGAN对齐目标域和源域负荷特征的空间分布, 并在最优传输代价函数中加入局部特征损失以提高训练的稳定性 and 快速性. 最后, 将对抗训练后网络用于目标域负荷预测. 采用该方法对某地区小样本负荷进行实验, 结果表明, 本文所提算法与其他预测模型相比能达到更高精度.

关键词: 负荷预测; 迁移学习; 小样本; 改进Wasserstein生成对抗网络; 特征分布; 最优传输

引用格式: 卢俊波, 刘俊峰, 罗燕, 等. 基于改进WGAN考虑特征分布相似性的小样本负荷预测方法. 控制理论与应用, 2024, 41(4): 597 – 608

DOI: 10.7641/CTA.2023.20876

Small sample load forecasting method considering characteristic distribution similarity based on improved WGAN

LU Jun-bo¹, LIU Jun-feng¹, LUO Yan¹, ZENG Jun^{2†}

(1. School of Automation Science and Engineering, South China University of Technology, Guangzhou Guangdong 510640, China;
2. School of Electric Power, South China University of Technology, Guangzhou Guangdong 510640, China)

Abstract: For a new user of an integrated energy system, it is much more difficult to develop an accurate load forecasting model due to the lack of historical data. A small sample load forecasting method based on the improved Wasserstein generative adversarial nets(WGAN) is proposed based on the transfer learning theory. First, the maximal information coefficient method is used to quantify the correlation among the impact characteristics and the load. Next, the source domain characteristic sequence is segmented and the edit distance on real sequence between each segmented sequence and the small sample in the target domain is calculated to determine the initial source domain. Then, the convolution neural network and long short-term memory model are introduced to establish the source domain prediction network. The spatial distribution of load characteristics both in target domain and source domain is aligned by WGAN, and the local characteristic loss is added to the optimal transport cost function to improve the stability and rapidity of training process. Finally, the network after adversarial training is used for the target domain load forecasting. The method proposed is used to test a small sample in a certain area and the result shows that the algorithm proposed in this paper turns out to be more accurate compared with other prediction models.

Key words: load forecast; transfer learning; small sample; improved Wasserstein generative adversarial nets; characteristic distribution; optimal transport

Citation: LU Junbo, LIU Junfeng, LUO Yan, et al. Small sample load forecasting method considering characteristic distribution similarity based on improved WGAN. *Control Theory & Applications*, 2024, 41(4): 597 – 608

1 引言

随着“双碳”目标的提出, 综合能源系统(integrated energy system, IES)成为能源转型过程中一种重要的能源利用方式, 是新型配电系统的重要组成. IES

以电能为核心, 风、光、气等多能互补, 通过内部种类众多的能量转换设备和能源储存设备, 实现多能系统间的协同规划、运行与管理. 作为IES需求侧能量预测的重要方面, IES负荷预测已经成为规划、运行控制与

收稿日期: 2022–10–09; 录用日期: 2023–03–06.

†通信作者. E-mail: junzeng@scut.edu.cn.

本文责任编辑: 岳东.

国家自然科学基金项目(62173148, 51877085)资助.

Supported by the National Natural Science Foundation of China (62173148, 51877085).

调度的首要前提^[1].

区别于传统电网, IES负荷侧规模相对较小, 单用户用能行为的变化或者新用户个体的接入, 都会对系统整体负荷产生明显影响. 而且, 系统具有相互耦合及不同时间尺度变化的电、热、冷等多种负荷, IES的电力负荷精准预测面临更多的挑战与困难. 目前, 短期电力负荷预测通常指对未来几小时到几周的用电负荷进行预测, 预测方法主要分为传统统计学方法和基于人工智能的方法. 传统方法主要考虑数据中的时序关系, 并以此建立线性映射模型, 如自回归积分滑动平均(autoregressive integrated moving average model, ARIMA)模型^[2]、灰色预测模型^[3]等, 但其对于数据的非线性关系表达能力有限. 而在IES中存在风机、光伏、天然气、燃煤等多种类型的发电设备相互耦合, 进一步增加了用户侧电力负荷的非线性变化. 一些具有强非线性表达能力的机器学习方法, 如支持向量回归(support vector machine, SVR)^[4]、反向传播神经网络(back propagation neural network, BPNN)模型^[5]等, 虽具有一定的非线性映射能力, 但缺点是存在序列信息缺失的问题. 近年来, 各种深度学习被用于短期负荷预测中. 文献[6]在考虑经济因素的条件下, 建立了基于实时电价的长短期记忆(long short-term memory, LSTM)负荷预测模型. 文献[7-8]分别采用了卷积神经网络(convolutional neural network, CNN)-LSTM混合神经网络和子序列分解组合模型预测了短期负荷变化, 有效提高了预测精度. 但是, 这些方法都建立在大量历史数据基础之上. 而IES的落地进程还处于起步阶段, 新接入IES的用户普遍不具备充足的负荷数据, 此时, 基于海量历史数据的深度学习电力负荷预测方法不再适用. 针对此, 文献[9]提出了一种遗忘因子特征融合迁移机制, 通过门控循环单元(gated recurrent unit, GRU)提取样本负荷隐特征, 并用余弦相似度挑选相似样本对目标小样本进行了日前预测. 但其只考虑了样本自身特征的相似, 而忽略了不同领域间样本特征空间分布的相似性. 对类似数据匮乏场景的时间序列预测问题, 部分学者从迁移学习出发, 引入生成对抗网络(generative adversarial nets, GAN)^[10]的思想进行研究. 文献[11]针对新建光伏电站原始数据少, 光伏发电功率预测精度低的问题, 利用一种梯度惩罚wasserstein GAN(wasserstein GAN with gradient penalty, WGAN-GP)生成新的样本数据, 对原始样本进行了扩充, 有效提高了光伏功率预测精度. 文献[12]提出了一种改进的pix2pix-GAN模型, 让GAN朝着历史数据的方向进行训练, 并用训练后的生成器对火电厂Nox实时排放量进行了软测量. 文献[13]提出了一种循环卷积GAN(recurrent convolutional GAN, RCGAN), 在数据匮乏的场景下生成“II类元胞负荷”, 实现了工程实例下对空间负荷的预测.

基于此, 本文采用迁移学习结合WGAN的方式, 解决新接入IES的用户在数据匮乏场景下的电力负荷日前预测问题. 区别于以往研究中多注重于GAN结构构造和待迁移样本的特征选择, 本文同时考虑了不同领域间样本特征分布的相似性差异, 提出了一种基于改进WGAN与特征分布相似性的小样本负荷预测方法. 该方法使用最大信息系数^[14](maximal information coefficient, MIC)来表示各特征影响因子对负荷的权重概率; 同时, 根据与目标域小样本负荷的实序列编辑距离(edit distance on real sequence, EDR)距离确定初始相似源领域. 接着, 采用CNN提取初筛源域的高维特征, 并利用WGAN^[15]度量高维特征分布的空间距离, 同时, 加入势能函数局部约束进一步提高其特征内部对齐程度; 最后, 将目标域特征提取器与预训练网络连接, 预测目标域小样本负荷变化. 本文在实际仿真算例中与其它传统预测算法进行了比较, 验证了本文所提算法具有更高预测精度.

2 区域级IES分析与源领域选取

2.1 典型区域级IES能量网络分析

典型区域级IES结构如图1所示. 系统能源输入侧包括电能、气能、风能和太阳能. 能量转换侧包括燃料电池、冷热电联供(combined cooling heating and power, CCHP)、电热泵、制冷机等多种能量转换装置, 以满足用户侧多种形式的用能需求, 同时, 通过各储能装置(如储气、储热、储冷装置)对能源的输入输出进行调节.

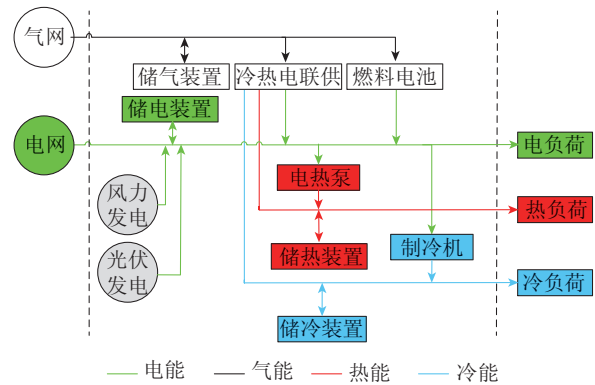


图1 区域级综合能源系统结构

Fig. 1 Regional integrated energy system architecture

从图1可以看出, IES电力负荷的特点是: IES是一个相互耦合、开放动态的系统. IES系统整体体积小, 较小的负荷波动也会对系统产生大的影响, 系统负荷侧随机性和波动性特征显著, 加之系统中分布式可再生能源的存在, 源侧也存在间歇性和波动性, 更增加了系统整体负荷预测的难度; 另一方面, 电热冷不同形式负荷存在耦合, 且动态时间尺度不一. 常规单一的物理建模的方式很难获得准确的负荷预测. 从数据驱动的角度出发, 是解决IES系统负荷预测的有效途

径. 但是, IES系统目前存在可利用历史数据不足, 尤其是面对新接入用户时, 数据不足更是负荷预测中难点. 有鉴于此, 本文引入迁移学习理论, 将源领域已有知识迁移到待预测小样本用户当中去, 解决IES中新接入用户数据匮乏时深度学习方法不适用的问题.

2.2 影响因子相关性分析

IES中新接入用户由于运行时间短、可用数据少, 在本文中定义为目标域小样本用户. 目标域小样本用户电力负荷变化会受到多种因素影响(如气象因素、日期因素等). 为降低模型敏感性, 需量化分析多维影响特征与待预测负荷之间的相关性强弱. 本文所考虑的特征影响因子如表1所示, 采样间隔为30 min.

表1 特征影响因子

Table 1 Characteristic influence factor

影响因子	特征类型	特征描述
日期因子	工作日	工作日用0表示
	节假日	节假日用1表示
气象因子	能见度	目标轮廓能被辨识的最大距离, m
	光照强度	1-9分别表示光照强度由弱到强
	湿度	当日平均湿度值, RH%
	气压	当日平均气压值, Pa
	温度	当日平均温度值, °C
气象因子	风向	当日的风向, 8个风向分别用1-8表示
	风速	当日的平均风速, m/s
负荷因子	电负荷	待预测日前 n 天的电力负荷, 采样间隔为30 min, MW

气象和日期因子统称为辅助影响因子, 负荷因子为直接影响因子. 对于辅助影响因子采用MIC法来捕捉其与待预测负荷之间的深层变化关系. 同时, 为了精准反映每一维特征对目标域小样本负荷影响程度, 利用softmax函数将各MIC计算值转换为对应特征权重概率分布^[14], 即

$$I(x, y) = \sum_{x, y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}, \quad (1)$$

其中: $I(x, y)$ 为辅助特征影响因子 x 和对应负荷序列 y 的互信息; $p(x)$, $p(y)$ 分别为其边缘分布函数; $p(x, y)$ 为联合分布函数. 将二维数据 (x, y) 进行 $\text{count1} \times \text{count2}$ 大小的网格划分, 在不同的划分下分别求出最大的互信息值, 最后, 对互信息值归一化, 即可得

$$\text{MIC}(x, y) = \max_{\text{count1} \times \text{count2} < B} \frac{I(x, y)}{\log(\min(\text{count1}, \text{count2}))}, \quad (2)$$

其中 B 表示最大划分的网格数, 一般为样本量的0.6次方. 权重概率计算如下:

$$d_i = \frac{e^{x_i}}{\sum_{i=1}^k e^{x_i}}, \quad (3)$$

其中: $d_i (i = 1, 2, \dots, k)$ 表示第 i 维辅助影响因子的权重概率; k 为辅助影响因子维度. 具体地, 对于初始源域中各典型日的多维特征矩阵 $X_{\text{sou}} = [x_{\text{sou}}^1 \ x_{\text{sou}}^{i+1} \ \dots \ x_{\text{sou}}^k]$ ($x_{\text{sou}}^i = [x_{\text{sou}}^{(t)} \ x_{\text{sou}}^{(t+1)} \ \dots \ x_{\text{sou}}^{(t+48)}]$), 计算每一个 x_{sou}^i 特征序列与对应负荷序列 y_{sou} 的MIC值(本文变量中下标含sou表示源领域变量, 含tar表示目标域变量), 同时, 根据式(3)转换为相应权重概率 d_i . 训练时, 将待预测日所属典型日的权重概率 d_i 与对应的特征影响因子 x_{sou}^i 做乘积, 作为后续特征提取模型的输入. 对于历史负荷数据, 由于这部分特征对待预测日有直接影响, 故采用灰色关联分析^[16]来寻找与典型日负荷具有密切关联的日期, 并根据关联度确定与待预测日相关的滞后序列长度, 具体细节见文献[16].

2.3 相似源领域选择

为了避免源领域选取不合理给目标域小样本负荷预测时带来的负迁移, 需要寻找相似源域进行领域间知识迁移. 且传统的固定滑动窗口分割方法无法充分提取时间序列数据的局部信息, 故采用基于Gamma分布的多元时间序列分割模型^[17]对源域负荷特征时间序列进行分割. 该算法通过主成分分析法(principal component analysis, PCA)对负荷特征进行降维, 并根据累计贡献率取出前 K 个主成分; 利用不同参数的Gamma分布拟合不同片段的特征序列, 并由赤池信息准则(Akaike information criterion, AIC)确定最优分割边界, 最大程度根据时间序列不同片段差异进行分割, 细节见文献[17].

考虑到目标域小样本负荷与源域负荷具有不同的长度, 因此, 采用实序列编辑距离^[18](edit distance on real sequence, EDR)来衡量两个非等长序列的相似性, 其数学表示如式(4)所示:

$$\text{EDR}(V_{\text{seq1}}, V_{\text{seq2}}) = \begin{cases} \text{seq1}, \text{seq2} = 0, \\ \text{seq2}, \text{seq1} = 0, \\ \min\{\text{EDR}(V_{\text{seq1}-1}, V_{\text{seq2}-1}) + \rho, \\ \text{seq1} \neq 0, \text{且seq2} \neq 0, \\ \text{EDR}(V_{\text{seq1}-1}, V_{\text{seq2}}) + 1, \\ \text{EDR}(V_{\text{seq1}}, V_{\text{seq2}-1}) + 1\}, \end{cases} \quad (4)$$

其中: V_{seq1} 和 V_{seq2} 表示进行比较的多维负荷特征向量; seq1和seq2分别为对应的长度; $V_{\text{seq1}-1}$, $V_{\text{seq2}-1}$ 表示不含已比较元素的剩余序列; ρ 表示处罚因子.

$$d(v_{\text{seq1}}, v_{\text{seq2}}) = \sqrt{\sum_{i, j=1}^{\text{dim}} (v_{\text{seq1}}(i) - v_{\text{seq2}}(j))^2}, \quad (5)$$

$$\begin{cases} \rho = 0, d(v_{\text{seq1}}, v_{\text{seq2}}) \leq \varepsilon, \\ \rho = 1, d(v_{\text{seq1}}, v_{\text{seq2}}) > \varepsilon, \end{cases} \quad (6)$$

$$\gamma(V_{\text{seq1}}, V_{\text{seq2}}) = \frac{1}{\text{EDR}(V_{\text{seq1}}, V_{\text{seq2}}) + 1}, \quad (7)$$

式(5)中: dim 表示特征维数; $d(v_{\text{seq1}}, v_{\text{seq2}})$ 为两负荷序列上任意两个元素之间的距离, 当 $d(v_{\text{seq1}}, v_{\text{seq2}})$ 小于设定阈值 ε 时, ρ 取0, 反之, 取1, 实验中, ε 设为0.5. 其求解过程是一个动态规划问题, 通过子负荷特征向量EDR的累加可得到完整的EDR($V_{\text{seq1}}, V_{\text{seq2}}$). 当 V_{seq1} 和 V_{seq2} 之间的距离EDR($V_{\text{seq1}}, V_{\text{seq2}}$)越小时, 表明源域与目标小样本负荷相似性越大. $\gamma(V_{\text{seq1}}, V_{\text{seq2}})$ 表示两个向量的相似程度, 且 $\gamma(V_{\text{seq1}}, V_{\text{seq2}}) \in (0, 1]$. 本文通过 $\gamma(V_{\text{seq1}}, V_{\text{seq2}})$ 寻找与目标域小样本负荷相似的源域, 可以对源域数据起到“初筛”的作用, 提高源域质量.

3 小样本负荷预测算法模型

3.1 卷积神经网络特征提取

为充分挖掘源域负荷特征影响因子中隐含的高维信息, 采用卷积神经网络进行数据抽象和特征提取^[19].

由于负荷特征影响因子是多维数据, 且二维卷积处理多维数据的能力比一维卷积更强, 本文将二维卷积作为负荷特征提取器, 仿照图像处理中卷积神经网络的输入均为 $N \times N$ 大小的图像尺寸, 依次将输入数据按照负荷因子、日期因子、气象因子的顺序进行排列, 构成输入大小为 $N \times N$ 大小的特征矩阵, 具体排列方式如式(8)所示, 其维度大小为 $[u \times T, k + 1]$ (T 为日采样周期48), 将其变换为方阵, 若数据变换后不为

方阵, 则对缺省的数据进行填0处理, 即

$$\begin{bmatrix} x_{v,1}^1 & x_{v,2}^1 & \cdots & \cdots & x_{v,k}^1 & x_{v,k+1}^1 \\ x_{v,1}^2 & x_{v,2}^2 & \cdots & \cdots & x_{v,k}^2 & x_{v,k+1}^2 \\ \vdots & \vdots & & & \vdots & \vdots \\ x_{v,1}^T & x_{v,2}^T & \cdots & \cdots & x_{v,k}^T & x_{v,k+1}^T \\ \vdots & \vdots & & & \vdots & \vdots \\ x_{u,1}^1 & x_{u,2}^1 & \cdots & \cdots & x_{u,k}^1 & x_{u,k+1}^1 \\ x_{u,1}^2 & x_{u,2}^2 & \cdots & \cdots & x_{u,k}^2 & x_{u,k+1}^2 \\ \vdots & \vdots & & & \vdots & \vdots \\ x_{u,1}^T & x_{u,2}^T & \cdots & \cdots & x_{u,k}^T & x_{u,k+1}^T \end{bmatrix}, \quad (8)$$

在式(8)所示的特征矩阵中, v 和 u 分别表示待预测小样本之前的第 v 天和第 u 天; $x_{v,1}^t$ 表示第 v 天 t 时刻的第1维特征(其中 $t = 1, 2, \dots, T-1, T$, 特征共9维); $x_{v,1}^T, x_{v,2}^T, x_{v,3}^T, \dots, x_{v,8}^T, x_{v,9}^T$ 依次表示电负荷、日期、能见度、光照强度、湿度、气压、温度、风向和风速变量.

本文所采用的特征提取器结构如图2所示. 图中输入为包括负荷因子在内的9维时间序列特征矩阵. 特征提取器模型总体上由3层卷积层构成, 其中: Conv2d, MaxPool, Dense分别表示卷积层、最大池化层、全连接层; BatchNorm表示批标准化层, 数据在输入Dense层之前进行打平操作. 为了防止出现过拟合, 在Dense层之后添加随机失活层Dropout, 输出为负荷的最终提取特征, 作为LSTM预测网络的输入.

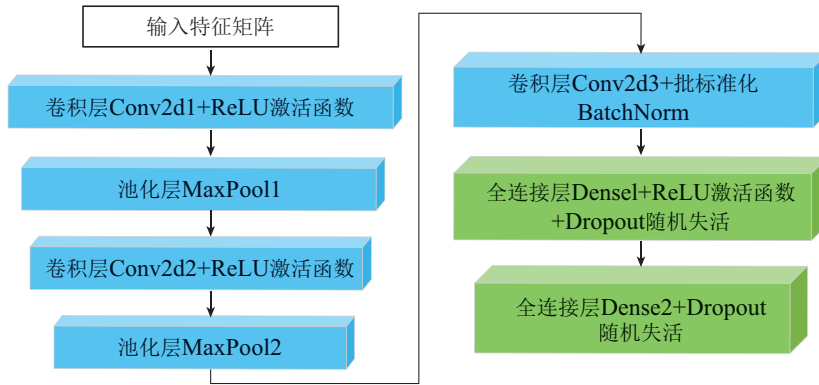


图2 特征提取器结构

Fig. 2 Structure of feature extractor

3.2 LSTM预测网络预训练

在提取源领域高维负荷特征后, 需要构建初始预测网络, 本文采用LSTM模型来构建预训练预测网络. LSTM^[20]是一种改进的用于处理时间序列问题的循环神经网络, 它独特的门结构单元避免了传统循环神经网络中的梯度消失问题.

3.3 基于wasserstein距离的生成对抗网络

经过以上步骤已经获得了源域和目标域的高维空间负荷特征集合, 接着, 从特征迁移的角度出发构建

生成对抗网络(GAN), 将目标域小样本负荷特征与源域负荷特征进行空间对齐, 进而将目标域的小样本迁移到源域的大样本.

传统GAN多采用JS(Jensen-Shannon)距离^[21]来度量两个数据分布之间的距离, 但是当两个分布完全不相关时, JS距离存在度量死区. 本文采用wasserstein距离来度量特征分布之间的距离. wasserstein距离可以用来衡量任意两个数据分布之间的距离, 即使这两个分布完全不相关. 其数学表示如下:

$$W(P_{\text{sou}}, P_{\text{tar}}) = \inf_{T_{\text{dis}} \in \Pi(P_{\text{sou}}, P_{\text{tar}})} E_{(x_{\text{sou}}, x_{\text{tar}}) \sim T_{\text{dis}}} [\|x_{\text{sou}} - x_{\text{tar}}\|], \quad (9)$$

假设 P_{sou} 和 P_{tar} 分别表示源域负荷特征集合和目标域小样本负荷特征集合; $\Pi(P_{\text{sou}}, P_{\text{tar}})$ 表示 P_{sou} 和 P_{tar} 所有可能的联合分布; 从联合分布 $\Pi(P_{\text{sou}}, P_{\text{tar}})$ 中每采样一对样本 $(x_{\text{sou}}, x_{\text{tar}})$, 计算它们之间的距离; 在所有可能的 $\Pi(P_{\text{sou}}, P_{\text{tar}})$ 中, 样本距离的期望下界, 即为 P_{sou} 和 P_{tar} 之间的wasserstein距离. 这里采用首先将其转换为最优传输问题, 再加入正则项进行Kantorovich松弛, 求其近似解^[22]. 其描述如下:

$$W(P_{\text{sou}}, P_{\text{tar}}) = \min_{T_{\text{r}} \in \Pi(P_{\text{sou}}, P_{\text{tar}})} \sum_{i=1}^m \sum_{j=1}^n C_{ij} T_{rij} + \mu H(T_{\text{r}}), \quad (10)$$

$$C_{ij} = (x_i - y_j)^2, \quad (11)$$

其中: $C \in \mathbb{R}_{m \times n}^+$ 表示传输代价(\mathbb{R}^+ 为正实数集); C_{ij} 表示将 P_{sou} 中的第 i 个负荷特征传输到目标域 P_{tar} 中的第 j 个负荷特征所花费的代价, 本文采用欧式距离的平方计算所付出代价; T_{r} 表示传输计划; T_{rij} 表示从 P_{sou} 中的第 i 个负荷特征传输到目标域 P_{tar} 中的第 j 个负荷特征的数据量; $W(P_{\text{sou}}, P_{\text{tar}})$ 表示领域间总的传输距离; $\mu H(T_{\text{r}})$ 为正则项; 同时, 传输计划 T_{r} 还需要满足如下约束条件:

$$\begin{cases} \sum_{j=1}^n T_{rij} \leq P_{\text{sou}}, \\ \sum_{i=1}^m T_{rij} \leq P_{\text{tar}}, \\ T_{rij} \geq 0, \end{cases} \quad (12)$$

直观地说, 将 P_{sou} 和 P_{tar} 看成两个土堆, 需要寻找出一种传输方案使得将 P_{sou} 中的土搬到 P_{tar} 去时, 总的传输代价 $\sum_{i=1}^m \sum_{j=1}^n C_{ij} T_{rij}$ 最小. 要求解上述的最优化问题, 关键是求出最优传输计划 $T_{\text{r}}^{\text{best}}$, 可用sinkhorn迭代^[23]求解, 但sinkhorn迭代很容易因为 μ 的大小设置而使模型变得不稳定. 本文采用近似近端点迭代法^[24-25](inexact proximal point algorithm, IPOT)来求解传输计划 T_{r} . 结合本文研究的负荷特征分布距离度量问题和IPOT方法, 将 T_{r} 的求解转换为式(13)的形式, 即

$$T_{\text{r}}^{(t+1)} = \arg \min_{T_{\text{r}} \in \Pi(P_{\text{sou}}, P_{\text{tar}})} \langle C, T_{\text{r}} \rangle + \beta d_{\text{h}}(T_{\text{r}}, T_{\text{r}}^{(t)}), \quad (13)$$

$$T_{\text{r}}^{(t+1)} = \arg \min_{T_{\text{r}} \in \Pi(P_{\text{sou}}, P_{\text{tar}})} \langle C - \beta \log T_{\text{r}}^{(t)}, T_{\text{r}} \rangle + \beta h(T_{\text{r}}), \quad (14)$$

$$h(T_{\text{r}}) = \sum T_{rij} \log T_{rij}, \quad (15)$$

在式(13)中: $\langle C, T_{\text{r}} \rangle$ 表示传输代价和传输计划的点积; β 表示变化步长, 本文取 $\beta = 1$; $d_{\text{h}}(T_{\text{r}}, T_{\text{r}}^{(t)})$ 表示近端算子, 这里采用Bregman散度^[25]; 进一步令 h 为熵函数 $h(T_{\text{r}})$, 并将式(15)代入式(13)中, 经过变形可以

得到式(14)的形式. IPOT法无需人工选择正则项系数 β , 且对不同的 β 值均表现出较好的稳定性.

3.4 基于改进WGAN的对抗训练

应用最优传输解法, 固定源域特征提取器 G_1 参数, 同时, 构建未训练的具有相同结构的目标域负荷特征提取器 G_2 , 以 G_1 和 G_2 为生成器, 并用wasserstein距离衡量源域和目标域小样本高维负荷特征的相似度, 以此建立WGAN进行特征空间对齐的网络对抗训练. 为了更好地捕捉负荷特征的高维信息, 并优化其局部的细节特征, 增强高维特征内部的对齐, 受文献^[26]启发, 在传输代价 $\sum_{i=1}^m \sum_{j=1}^n C_{ij} T_{rij}$ 中引入马尔可夫随机场约束, 并定义势能函数如式(16)所示:

$$E_{\text{L}}(g_1(x_{\text{sou}}), g_2(x_{\text{tar}})) = \frac{1}{M} \sum_{i=1}^M \|g_{1i}(x_{\text{sou}}) - g_{2i}(x_{\text{tar}})\|^2, \quad (16)$$

其中: $E_{\text{L}}(g_1(x_{\text{sou}}), g_2(x_{\text{tar}}))$ 为势能函数; M 表示卷积层层数; $g_{1i}(x_{\text{sou}})$ 表示源域特征提取器 G_1 第 i 层的输出, 其输出大小为 $H \times W \times \text{Ch}$, 其中 H, W, Ch 分别为输出序列的高度、宽度和通道数, 将其与目标域相同层输出作差后求平方和, 并求 M 层输出的均值. 将 $E_{\text{L}}(g_1(x_{\text{sou}}), g_2(x_{\text{tar}}))$ 加入最优传输WGAN的损失函数中, 在不同特征层进行对齐, 可以对浅层特征进行约束. 新的的损失函数为 $\text{Loss}_{\text{total}}$, 即

$$\begin{aligned} \text{Loss}_{\text{total}} &= L_{\text{transfer}} + L_{\text{lfea}} = \\ &\alpha_1 \min_{T_{\text{r}} \in \Pi(P_{\text{sou}}, P_{\text{tar}})} \sum_{i=1}^m \sum_{j=1}^n C_{ij} T_{rij} + \\ &\alpha_2 E_{\text{L}}(g_1(x_{\text{sou}}), g_2(x_{\text{tar}})), \end{aligned} \quad (17)$$

如式(17)所示: L_{transfer} 为传输代价损失; 由于引入势能函数是为了更好地对齐高维空间的细节特征, 则称 L_{lfea} 为局部特征损失; 传输代价 C_{ij} 为源域和目标域特征提取器最后一层输出 f_1 和 f_2 的 L_2 范数; 最优传输计划 $T_{\text{r}}^{\text{best}}$ 通过第2.2节中的快速近端点迭代法进行求解; α_1 和 α_2 均为权重系数; 根据总损失 $\text{Loss}_{\text{total}}$ 反向优化WGAN中网络参数. 经多次实验, 本文设置 α_1, α_2 分别取值为0.6和0.4.

此外, 所构建的改进的WGAN需要判别器对提取的高维负荷特征进行鉴别. 本文中的判别器 D 由3层全连接层构成, 神经元数量分别为60, 80, 10; 第1层和第2层分别采用ReLU(\cdot)函数进行激活. 对于生成器来说希望能够欺骗判别器, 因此, 生成器损失函数为最小化 $\text{Loss}_{\text{total}}$; 而对于判别器网络来说, 希望能尽可能地区分出真实负荷特征分布和生成负荷特征分布, 因此, 其损失函数为 $-\text{Loss}_{\text{total}}$, 即最大化 $\text{Loss}_{\text{total}}$.

在完成对抗训练后, G_2 已经学习到源域和目标域之间的域不变特征^[27], 因此, 可直接将其与预训练的LSTM预测网络进行拼接, 对目标域小样本进行预测. 图3为预测模型流程图.

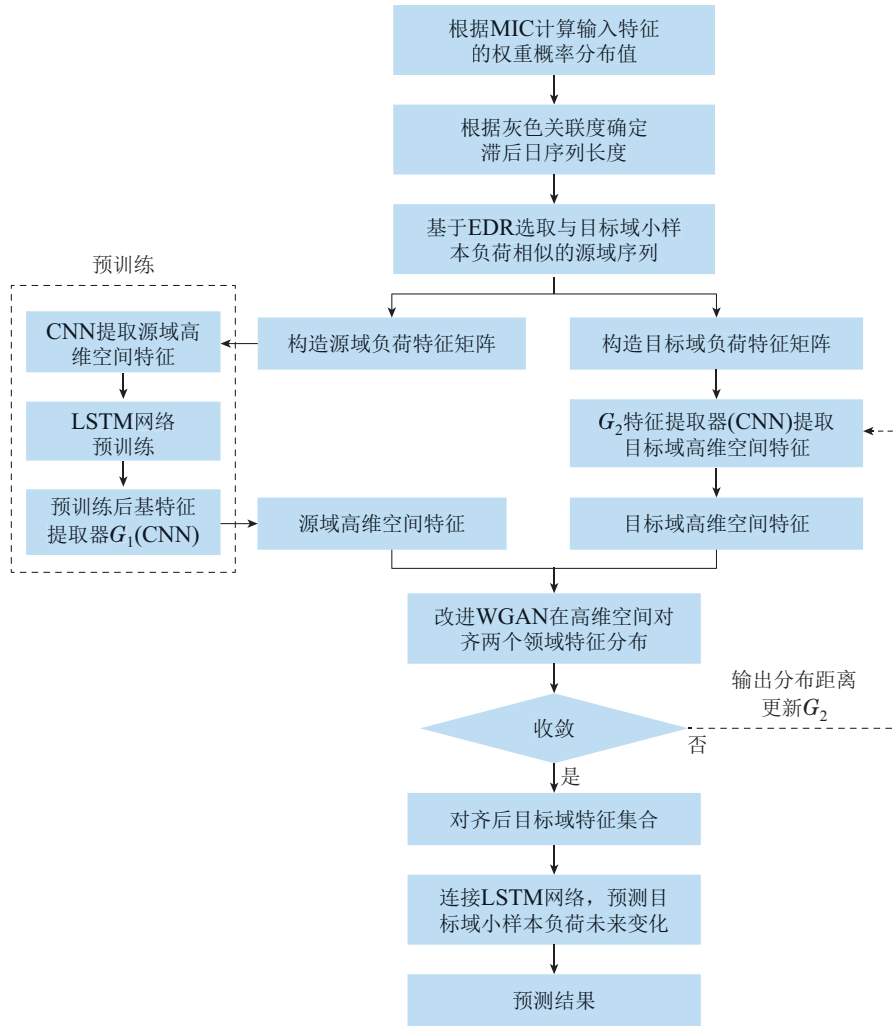


图3 预测模型流程图

Fig. 3 Flow chart of prediction model

整个负荷预测流程如下:

1) 确定权重概率和源域数据: 依次计算典型日辅助影响因子与待预测电负荷的MIC值、历史负荷序列与典型日负荷的灰色关联度, 分别得到各特征权重概率分布和时间序列滞后长度; 根据目标域小样本负荷数据 X_{tar} 与相关源域 X_{sou} 的EDR距离, 确定初始源域样本范围;

2) 网络预训练: 设计CNN特征提取器 G_1 和 G_2 提取源域和目标域小样本负荷的高维特征; 并构造预测LSTM网络, 利用源域样本训练 G_1 和预测网络直至模型收敛;

3) 改进的WGAN对抗训练: 固定 G_1 , 根据 G_1 和 G_2 输出特征向量之间的传输代价损失 $L_{transfer}$ 与局部特征损失 L_{fea} 之和更新 G_2 , 直至总损失 $Loss_{total}$ 收敛;

4) 预测目标域: 连接 G_2 和预训练LSTM预测网络, 预测目标域小样本的负荷变化。

3.5 负荷预测评价指标

为了对模型的预测性能进行评估, 本文选取预测精度 y_{FA} 、平均绝对百分误差 y_{MAPE} 、均方根误差

y_{RMSE} 作为评价指标, 其定义如式(18)–(20)所示:

$$y_{FA} = \left(1 - \frac{|x_{real}(i) - x_{pred}(i)|}{x_{real}(i)}\right) \times 100\%, \quad (18)$$

$$y_{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|x_{real}(i) - x_{pred}(i)|}{x_{real}(i)}, \quad (19)$$

$$y_{RMSE} = \sqrt{\frac{\sum_{i=1}^n (x_{real}(i) - x_{pred}(i))^2}{n}}, \quad (20)$$

其中: $x_{real}(i)$ 表示第 i 个时刻负荷的实际值; $x_{pred}(i)$ 表示第 i 个时刻负荷的预测值; n 为预测次数; 当 y_{FA} 越大, y_{MAPE} 和 y_{RMSE} 越小时, 说明模型的预测性能越好。

4 算例分析

本文实验选取某IES中4个工业园区的真实电力负荷数据和气象数据作为原始数据集。该IES以电能、天然气、风能和太阳能作为能源输入, 能量转换侧含燃料电池、电热泵、制冷机等能量转换装置, 同时, 还具有各储气、储热、储冷装置以备用。符合本文所研究区域级IES架构要求。各园区分别记作园区A, B, C和

园区D. 4个园区均记录了2012年1月1日至2012年12月31日的电力负荷; 为模拟小样本负荷预测情形, 将园区A, B, C一年的数据作为待选择源域, 在园区D随机选择3个连续的15天作为小样本目标域. 根据第1.2节所提出方法计算辅助影响因子的权重概率和直接影响因子的滞后长度; 然后, 进行源领域初步筛选, 将CNN特征提取网络捕捉的高维特征进行对齐, 并利用预测网络预测小样本目标域负荷后一天的负荷变化曲线.

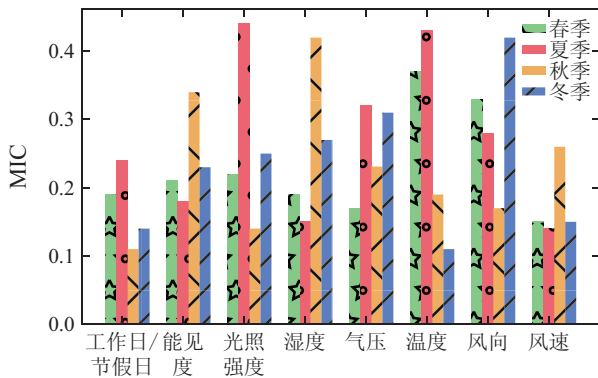
4.1 数据预处理

在实际中, 由于人为因素和通信设备故障经常导致历史负荷的缺失和损坏, 因此, 本文依据 3σ 准则^[28]剔除原始负荷数据集中的异常值, 对于缺失数据用邻近的前后3个共6个采样点平均值进行替代. 同时, 为了消除不同的物理量纲带来的影响, 对变量进行归一化, 消除量纲后数据范围变为 $[0, 1)$, 实际处理时为了避免归一化后出现为0的数据, 加上一个无限小的正常数.

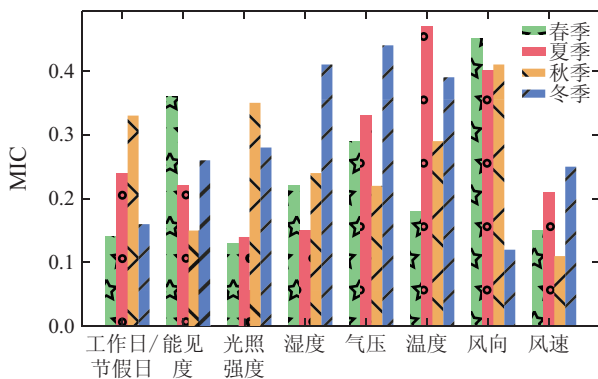
$$x_g = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (21)$$

其中: x_g 表示归一化后数据; x 表示原始负荷数据; x_{\max} 表示原始负荷数据中的最大值; x_{\min} 表示原始负荷数据中的最小值.

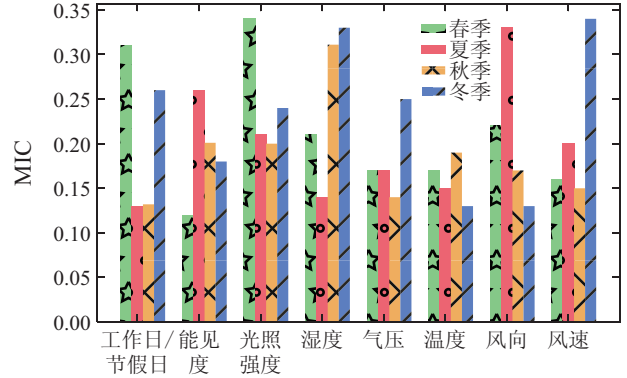
首先, 分别计算园区A, B和C中的各典型日辅助影响因子的MIC值(典型日选取见文献[29]), 并将其转换为对应权重概率 d_i , 结果如图4所示.



(a) 园区A典型日MIC值



(b) 园区B典型日MIC值



(c) 园区C典型日MIC值

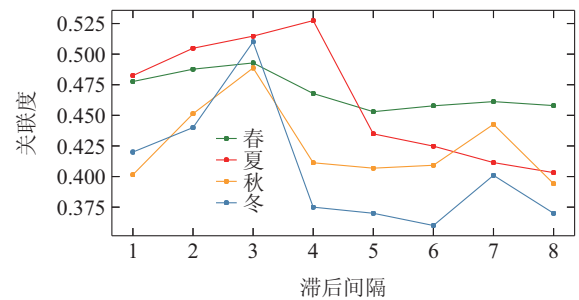
图4 源领域各典型日MIC值计算结果

Fig. 4 The calculation results of typical daily MIC value of source domain

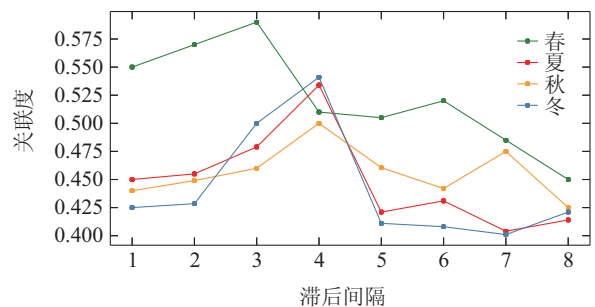
在预训练时采用滑动窗口的方式建立输入特征矩阵(如式(8)所示), 为了保证每个窗口下的特征矩阵能够充分反映待预测日的特征信息, 并结合GRA计算结果可知(见图5), 当滞后宽度大于3天和4天时, 关联度会出现下降, 因此, 设定滑动窗宽为3天. 园区D中目标域随机选择结果为: 2012年8月4日-2012年8月18日、3月11日-3月25日、12月12日-26日, 记作目标域I, II, III. 通过分割后的源域负荷特征子序列与目标域小样本之间的EDR, 得到初始源域样本数据集, 其中EDR划分阈值为0.5.

4.2 实验设置

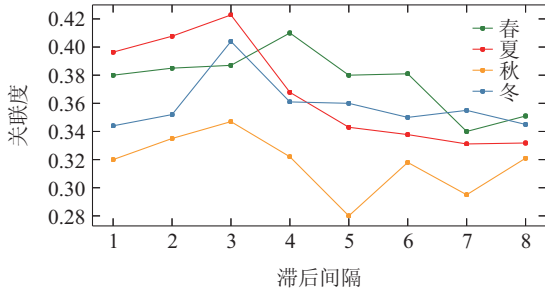
本实验平台的CPU采用Intel(R) Xeon(R) @ 2.20 GHz, GPU为Tesla P100, 内存大小为16 G. 深度学习框架选择Pytorch 1.5, 算法模型的编程语言基于Python 3.5.



(a) 园区A典型日滞后关联度



(b) 园区B典型日滞后关联度



(c) 园区C典型日滞后关联度

图5 源领域各典型日关联度计算结果

Fig. 5 The calculation results of typical daily correlation degree of source domain

4.3 模型训练与分析

由于本实验中采用前3天的历史负荷影响因子预测其后一天的负荷变化,则特征输入维度大小为 $\text{batch} \times 144 \times 9 \times 1$ (batch为批次大小,预训练中取8),特征输入依次排列为负荷因子、气象因子、日期因子,并按第2.1节的方法转换成 $\text{batch} \times n \times n \times 1$ 的形式.此处可将输入大小变为 $\text{batch} \times 36 \times 36 \times 1$,并恰好无需填充.接着,将输入时间序列特征与对应的权重概率 $d_i^{(t)}$ 相乘后送入特征提取器 G_1 ; 经过 G_1 输出大小变为 $\text{batch} \times 4 \times 4 \times 64$; 然后,将高维特征输入LSTM预测网络中.训练时将90%的源域特征数据作为训练集,剩余10%作为测试集,直至预测网络收敛,其训练误差和测试误差曲线如图6所示.

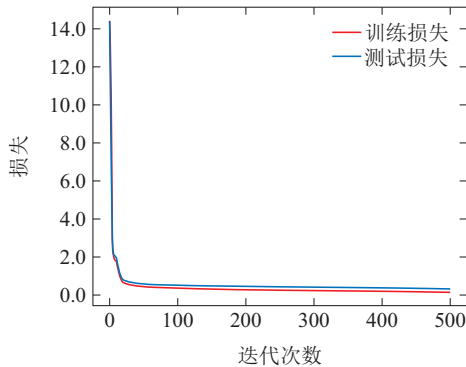


图6 预训练预测网络的损失变化曲线

Fig. 6 Loss curves of the pre-trained prediction network

训练好 G_1 和源域LSTM预测网络后,固定 G_1 网络参数,并利用改进的WGAN训练目标域小样本负荷数据集的特征提取器 G_2 (对抗训练时batch取1). 对抗训练总损失 $\text{Loss}_{\text{total}}$ 的变化如图7所示.

从图7可以看出,生成对抗网络的训练是不稳定的,且未改进的WGAN(不加入局部特征损失的WGAN)在训练初期会产生较长时间的振荡,虽然最终也能够稳定收敛到某一个值,但损失值始终处于小范围的波动状态;同时,由于没有考虑到负荷的浅层细节特征,最终的收敛损失值会偏大;而本文所提出的

改进的WGAN由于考虑了局部损失 $E_L(g_1(x_{\text{sou}}), g_2(x_{\text{tar}}))$, 不仅有效缩短了模型对抗训练初期的振荡时间,同时,使得总损失值进一步下降,并以更小的幅度快速稳定收敛到最优解,且后续能够保持稳定.图6为目标域I及其对应源域经未改进WGAN和改进WGAN训练后的输出的高维负荷特征经t-sne流形^[3]降维后的分布图.由图8可得:在加入了 $E_L(g_1(x_{\text{sou}}), g_2(x_{\text{tar}}))$ 之前,目标域小样本负荷的高维特征不能很好地对齐到源域高维特征的某一种表征当中,会存在少量特征分散状态;在加入了 $E_L(g_1(x_{\text{sou}}), g_2(x_{\text{tar}}))$ 之后,由于考虑了领域间的细节特征,它们在高维空间的分布会更加接近,此时,能够从包含大量特征分布模型的源域样本中明确地找出属于目标域小样本负荷的那一类特征分布模式,将对齐后的目标域小样本高维特征输入到预测网络中去,会具有更高的预测精度,这也与WGAN训练时损失值的变化情况一致.

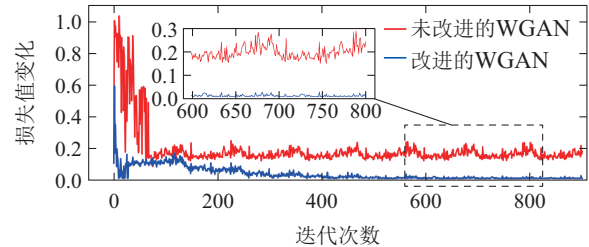


图7 对抗训练总损失变化

Fig. 7 Change of total loss of adversarial training

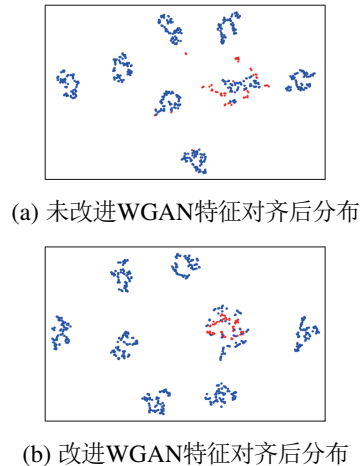


图8 训练后特征对齐分布图

Fig. 8 Diagram of feature alignment distribution after training

4.4 预测结果分析

经过改进WGAN的对抗训练后,源域与目标域小样本负荷的特征分布已经对齐,此时的特征提取器已经学习到两领域间的域不变特征,可将预训练的LSTM预测网络与目标域小样本负荷特征直接进行连接,对目标域I后一天即园区D中2012年8月19日的负荷进行预测.同时,将改进的WGAN预测结果与未改

进的WGAN模型、无特征对齐(有相似源域筛选)CNN-LSTM模型^[7]、无特征对齐(有相似源域筛选)LSTM模型^[6]、无特征对齐(有相似源域筛选)SVR模型^[3]预测结果进行对比。超参数的设置会对预测性能产生很大影响, 本文采用网格搜索法^[31]对超参数进行寻优, 特征提取层初始学习率为0.0001, 衰减率0.92; 预测层学习率0.003, 衰减率0.85; 无对齐LSTM模型学习率为0.001, 均采用Adam优化器; SVR模型核函数为径向基函数。

无对齐LSTM模型和无对齐SVR模型使用同样的源域I数据进行训练。最后, 通过 y_{MAPE} , y_{RMSE} , y_{FA} 这3个指标对不同模型预测性能进行评价。其中, 表2, 表3和表4中 \bar{y}_{FA} 为 y_{FA} 的平均值。

表2 目标域I的预测误差和平均精度

Table 2 Forecasting error and average accuracy of target domain I

模型	y_{MAPE}/MW	y_{RMSE}/MW	$\bar{y}_{FA}/\%$
改进的WGAN	0.0708	1.5996	92.92
未改进的WGAN	0.1114	2.4807	88.86
无对齐CNN-LSTM	0.1548	4.1636	84.52
无对齐LSTM	0.1760	3.8421	82.40
无对齐SVR	0.1927	3.6894	80.73

图9为采用5种预测方法对目标域I其后一天的负荷预测结果对比图。由图可得, 本文所提出的基于改进WGAN方法得到的预测结果最为准确。在目标域小样本负荷数据缺乏的情况下(仅含15日), 由于没有进行领域间的特征对齐, 无特征对齐LSTM和无特征对齐SVR两者均只能捕捉到负荷变化的总体轮廓, 预测负荷的变化趋势, 而对于负荷内部特征变化缺乏强表征能力; 而采用未改进的WGAN的小样本负荷预测方法, 虽然将目标领域高维负荷特征 $g_2(x_{tar})$ 对齐到了源领域空间当中, 但是没有引入马尔可夫随机场约束

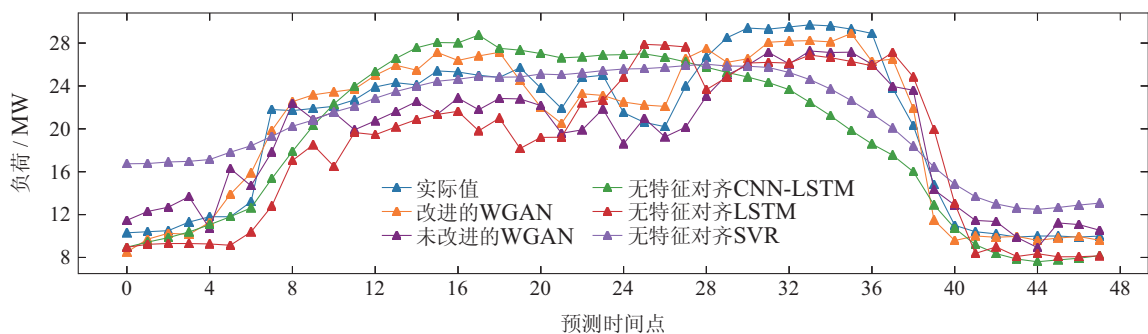


图9 目标域I不同模型预测结果对比图

Fig. 9 Comparison diagram of target domain I prediction results of different models

图10-11分别为目标域II和目标域III其后一天5种不同预测方法的预测结果。由图可知, 本文所提出的改进的WGAN更善于处理短时间内的非剧烈负荷变

来限定局部细节特征在对抗训练过程中的变化, 因此, 其预测精度不如加入约束的改进WGAN。而未采用深度学习的SVR预测模型则无法表达负荷的非线性变化特征。

表3 目标域II和III的预测误差和平均精度

Table 3 Prediction error and average accuracy of target domain II and III

数据集	模型	y_{MAPE}/MW	y_{RMSE}/MW	$\bar{y}_{FA}/\%$
目标域II	改进的WGAN	0.0696	1.6917	93.04
	未改进的WGAN	0.1175	2.8352	88.25
	无对齐CNN-LSTM	0.1612	4.4332	83.88
	无对齐LSTM	0.1793	4.9897	82.07
	无对齐SVR	0.1995	4.0886	80.05
目标域III	改进的WGAN	0.0702	2.3876	92.98
	未改进的WGAN	0.0995	3.2262	90.05
	无对齐CNN-LSTM	0.1329	3.5138	86.71
	无对齐LSTM	0.1608	3.6406	83.92
	无对齐SVR	0.2115	4.6496	78.85

表4 不同阈值下预测指标对比

Table 4 Comparison of prediction indices under different thresholds

EDR匹配阈值	y_{MAPE}/MW	y_{RMSE}/MW	$\bar{y}_{FA}/\%$
0.5	0.0708	1.5996	92.92
0.7	0.1644	3.7234	83.56
0.9	0.2384	4.4388	76.16
0.3	0.1519	3.0787	84.81
0.1	0.2871	6.7122	71.29

为进一步说明本文所提预测方法的有效性, 再对目标域小样本II和III进行验证, 同时, 将其他4种模型与本文方法进行对比。

化。例如, 在目标域II中, 其负荷在第12~25个采样时刻内存在明显的下降和回弹现象, 而改进WGAN能够很好地跟随这种反复的波动。在目标域III上, 改进

WGAN在傍晚时段的第2次负荷上升变化过程中出现了明显的偏差,且变化趋势刚好相反,这可能是由于待预测日负荷在短时间内变化幅度过小,导致在特征

对抗训练时丢失了这部分“缓慢上升”信息.尽管如此,本文方法的 y_{RMSE} 也仅有2.3876 MW, \bar{y}_{FA} 也达到了5种预测方法种最高的92.98%,如表3所示.

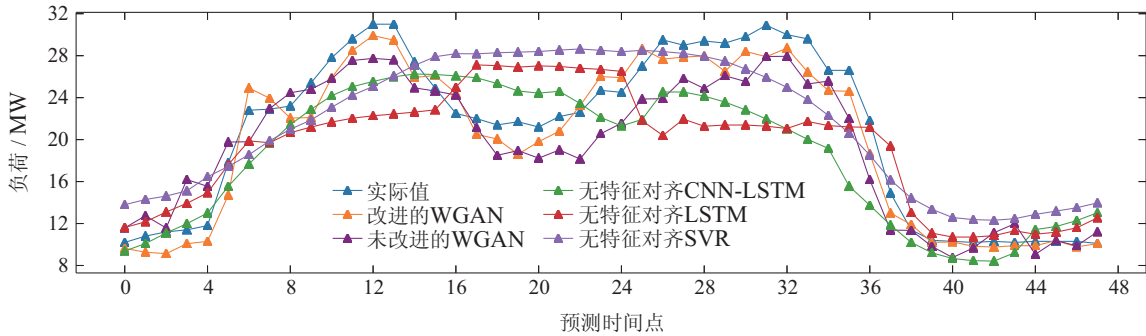


图 10 目标域II不同模型预测结果对比图

Fig. 10 Comparison diagram of target domain II prediction results of different models

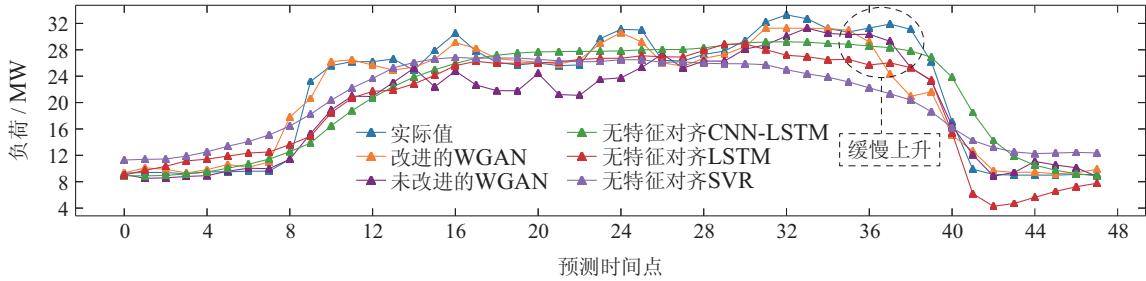


图 11 目标域III不同模型预测结果对比图

Fig. 11 Comparison diagram of target domain III prediction results of different models

为了比较加入局部特征损失 L_{lfea} 前后对抗训练网络训练速度的不同,表5分别记录了3个领域分别使用未改进的WGAN和改进的WGAN进行训练所用时间.由表可得,加入 L_{lfea} 后训练速度更快.

4.5 相似源域匹配阈值分析

本节将讨论进行相似源域选取时,匹配阈值对预测性能的影响.在目标域I上设置不同EDR阈值,获得相应源领域负荷样本并采用本文所提出方法进行迁移学习模型训练,各种EDR阈值下的预测结果与评价指标如图12和表4所示.

从预测结果可以看出,当相似源域匹配阈值选取偏大时,即阈值设置为0.7时,由于相似样本选择的门槛过高,会导致源域无法提供足够的样本进行迁移,

导致特征对齐训练的效果不佳,预测模型精度降低.当阈值设置过大(设置为0.9),预测效果会进一步下降.反之,当阈值设为0.3和0.1时,负迁移程度明显上升,源域样本中会出现较多的不相关样本,对迁移过程产生干扰.因此,需要根据实际目标域小样本情况对迁移阈值进行合理设置.

表 5 加入局部特征损失前后训练所用时间

Table 5 Training time before and after adding local feature loss

训练时间/s	域I	域II	域III
未改进的WGAN	50.13	46.16	60.52
改进的WGAN	38.70	32.11	43.23

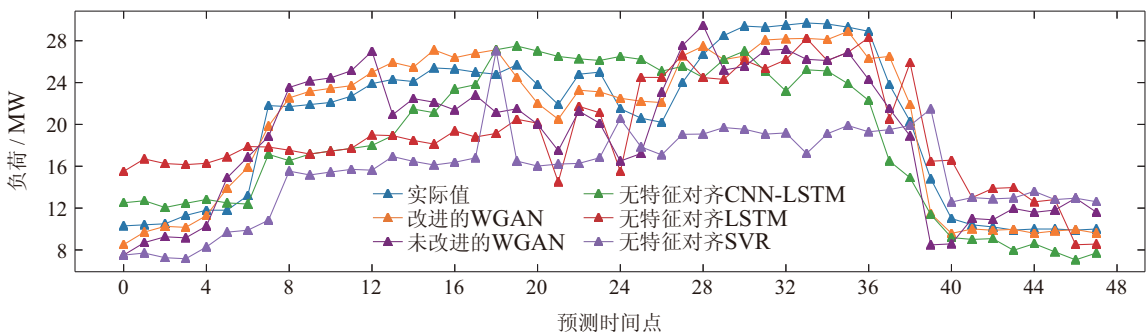


图 12 不同匹配阈值下预测结果对比图

Fig. 12 Comparison diagram of prediction results under different matching thresholds

5 结语

综上, 针对目前IES中新增用户数据缺乏, 难以建立精确负荷预测模型的问题, 本文提出了一种考虑特征分布相似性的小样本负荷预测方法, 所做出的主要贡献在于: 1) 在小样本负荷预测问题中引入迁移学习理论, 并基于特征分布相似性的思想, 用生成对抗网络对齐两个领域的高维特征, 使得预训练网络能够直接用于目标域小样本负荷预测; 2) 应用最优传输的思想求解WGAN的距离分布问题, 并在损失函数中加入马尔可夫随机场约束, 加强了对负荷高维细节特征的空间限制, 提高了网络对抗训练过程的稳定性. 本文通过实际算例将所提小样本负荷预测模型与未改进的WGAN、无特征对齐CNN-LSTM、无特征对齐LSTM、无特征对齐SVR预测模型进行对比, 验证了本文方法在预测精度上的优势. 由于在进行相似领域匹配过程时容易出现负迁移现象, 进而影响预测精度, 后续工作中将对领域间的趋势性变化加以考虑以减少负迁移, 同时, 将电价等经济因素考虑进去进一步提高模型的通用程度和预测精度.

参考文献:

- [1] ZHU Jizhong, DONG Hanjiang, LI Shenglin, et al. Review of data-driven load forecasting for integrated energy system. *Proceedings of the CSEE*, 2021, 41(23): 7905 – 7924.
(朱继忠, 董瀚江, 李盛林, 等. 数据驱动的综合能源系统负荷预测综述. 中国电机工程学报, 2021, 41(23): 7905 – 7924.)
- [2] GUO Jianpeng, SHE Yingling, WEN Buying. Regional monthly load forecast based on EEMD-ARIMA model. *Electric Power Engineering Technology*, 2018, 37(6): 28 – 32, 74.
(郭建鹏, 余颖铃, 温步瀛. 基于EEMD-ARIMA模型的地区月负荷量预测. 电力工程技术, 2018, 37(6): 28 – 32, 74.)
- [3] LIU Dandan, ZHU Jiaming, HUANG Tingting. Short-term power load forecasting based on time series and grey model. *Journal of Qiqihar University (Natural Science Edition)*, 2017, 33(3): 7 – 12, 20.
(刘丹丹, 朱家明, 黄婷婷. 基于时间序列和灰色模型的短期电力负荷预测. 齐齐哈尔大学学报(自然科学版), 2017, 33(3): 7 – 12, 20.)
- [4] CAI Shuping, YAN Jing, LIU Guohai, et al. A dynamic modeling methodology based on fisher information and on-line SVR for smart grids weather sensitive load forecasting. *Proceedings of the CSEE*, 2020, 40(11): 3441 – 3452.
(蔡舒平, 闫静, 刘国海, 等. 基于Fisher信息和在线SVR的智能电网气象敏感负荷预测动态建模技术. 中国电机工程学报, 2020, 40(11): 3441 – 3452.)
- [5] WU Yun, LEI Jianwen, BAO Lishan, et al. Short-term load forecasting based on improved grey relational analysis and neural network optimized by bat algorithm. *Automation of Electric Power System*, 2018, 42(20): 67 – 72.
(吴云, 雷建文, 鲍丽山, 等. 基于改进灰色关联分析与蝙蝠优化神经网络的短期负荷预测. 电力系统自动化, 2018, 42(20): 67 – 72.)
- [6] LI Peng, HE Shuai, HAN Pengfei, et al. Short-term load forecasting of smart grid based on long-short-term memory recurrent neural networks in condition of real-time electricity price. *Power System Technology*, 2018, 42(12): 4045 – 4052.
(李鹏, 何帅, 韩鹏飞, 等. 基于长短期记忆的实时电价条件下智能电网短期负荷预测. 电网技术, 2018, 42(12): 4045 – 4052.)
- [7] LU Jixiang, ZHANG Qiwei, YANG Zhihong, et al. Short-term load forecasting method based on CNN-LSTM hybrid neural network model. *Automation of Electric Power Systems*, 2019, 43(8): 131 – 137.
(陆继翔, 张琪培, 杨志宏, 等. 基于CNN-LSTM混合神经网络模型的短期负荷预测方法. 电力系统自动化, 2019, 43(8): 131 – 137.)
- [8] DU Hanxiao, TANG Min'an. Short-term load forecasting for user-side micro-grid based on CEEMD-RSVPSO-KELM model. *Electrical Measurement & Instrumentation*, 2020, 57(18): 69 – 76.
(杜涵潇, 汤旻安. 基于CEEMD-RSVPSO-KELM的用户侧微电网短期负荷预测. 电测与仪表, 2020, 57(18): 69 – 76.)
- [9] SUN Xiaoyan, LI Jiazhao, ZENG Bo, et al. Small-sample day-ahead power load forecasting of integrated energy system based on feature transfer learning. *Control Theory & Applications*, 2021, 38(1): 63 – 72.
(孙晓燕, 李家钊, 曾博, 等. 基于特征迁移学习的综合能源系统小样本日前电力负荷预测. 控制理论与应用, 2021, 38(1): 63 – 72.)
- [10] ZHAO Jinquan, XIA Xue, XU Chunlei, et al. Review on application of new generation artificial intelligence technology in power system dispatching and operation. *Automation of Electric Power Systems*, 2020, 44(24): 1 – 10.
(赵晋泉, 夏雪, 徐春雷, 等. 新一代人工智能技术在电力系统调度运行中的应用评述. 电力系统自动化, 2020, 44(24): 1 – 10.)
- [11] YIN Hao, ZHANG Zheng, DING Weifeng, et al. Short term prediction of small sample photovoltaic power based on generative adversarial network and LSTM-CSO. *High Voltage Engineering*: 2022, 48(11): 4342 – 4351.
(殷豪, 张铮, 丁伟锋, 等. 基于生成对抗网络和LSTM-CSO的少样本光伏功率短期预测. 高电压技术: 2022, 48(11): 4342 – 4351.)
- [12] ZHANG Zhiyu, LI Xinli, YANG Guotian. NOx emission prediction of thermal power plant based on improved generative adversarial network pix2pix. *Proceedings of the CSEE*: 2023, 43(12): 4678 – 4688.
(张志宇, 李新利, 杨国田. 基于改进生成对抗网络pix2pix的火电厂NOx排放预测. 中国电机工程学报: 2023, 43(12): 4678 – 4688.)
- [13] XIAO Bai, HUANG Yuru, JIANG Zhuo, et al. The method of spatial load forecasting based on the generative adversarial network for data scarcity scenarios. *Proceedings of the CSEE*, 2020, 40(24): 7990 – 8001, 8236.
(肖白, 黄钰茹, 姜卓, 等. 数据匮乏场景下采用生成对抗网络的空间负荷预测方法. 中国电机工程学报, 2020, 40(24): 7990 – 8001, 8236.)
- [14] LIANG Tie, ZHANG Qingyu, HONG Lei, et al. An improved maximal information coefficient algorithm applied in the analysis of functional corticomuscular coupling for stroke patients. *Journal of Biomedical Engineering*, 2021, 38(6): 1154 – 1162.
(梁 tie, 张清瑜, 洪磊, 等. 一种改进的最大信息系数算法在脑卒中患者的皮质肌功能耦合分析中的应用. 生物医学工程杂志, 2021, 38(6): 1154 – 1162.)
- [15] MARTIN ARJOVSKY S C, BOTTOU L. Wasserstein generative adversarial networks. *Proceedings of the 34th International Conference on Machine Learning*. Sydney, Australia: PMLR, 2017: 214 – 223.
- [16] LIU Jinjin, ZHOU Ping, WEN Liang. Root mean square error probability weighted integrated learning based modeling for molten iron quality in blast furnace ironmaking. *Control Theory & Applications*, 2020, 37(5): 987 – 998.
(刘进进, 周平, 温亮. 高炉铁水质量均方根误差概率加权集成学习建模. 控制理论与应用, 2020, 37(5): 987 – 998.)
- [17] WANG Benchao, LI Dan, QIN Pan, et al. Gamma distribution based traffic flow time series segmenting model. *Journal of Dalian University of Technology*, 2020, 60(3): 293 – 299.
(王本超, 李丹, 秦攀, 等. 基于Gamma分布的交通流时间序列分割模型. 大连理工大学学报, 2020, 60(3): 293 – 299.)
- [18] PAN Xiao, MA Ang, GUO Jingfeng, et al. Research and application survey of similarity measurement methods on trajectory data based on time series. *Journal of Yanshan University*, 2019, 43(6): 531 – 545.

- (潘晓, 马昂, 郭景峰, 等. 基于时间序列的轨迹数据相似性度量方法研究及应用综述. 燕山大学学报, 2019, 43(6): 531 – 545.)
- [19] XU Yanlu, WU Zhikai, ZHU Heyan, et al. Short-term load forecasting based on multi-scale convolutional neural network. *Journal of Shenyang University of Technology*, 2020, 42(6): 618 – 623. (许言路, 武志楷, 朱赫炎, 等. 基于多尺度卷积神经网络的短期电力负荷预测. 沈阳工业大学学报, 2020, 42(6): 618 – 623.)
- [20] GE Leijiao, ZHAO Kang, SUN Yonghui, et al. Automation of electric power systems, short-term load forecasting of distribution network based on combination of siamese network and long short-term memory network. 2021, 45(23): 41 – 50. (葛磊蛟, 赵康, 孙永辉, 等. 基于孪生网络和长短时记忆网络结合的配电网短期负荷预测. 电力系统自动化, 2021, 45(23): 41 – 50.)
- [21] TENG Shaohua, KONG Lengrui. Chinese fonts style transfer based on generative adversarial networks. *Application Research of Computers*. 2019, 36(10): 3164 – 3167. (滕少华, 孔棱睿. 基于生成式对抗网络的中文字体风格迁移. 计算机应用研究, 2019, 36(10): 3164 – 3167.)
- [22] CHE Lingfu, TIAN Yukun, ZHU Haiping, et al. Optimal transport based transfer learning. *Pattern Recognition and Artificial Intelligence*, 2019, 32(6): 481 – 493. (车令夫, 田宇坤, 朱海平, 等. 基于最优运输的迁移学习. 模式识别与人工智能, 2019, 32(6): 481 – 493.)
- [23] CUTURI M. Sinkhorn distances: Lightspeed computation of optimal transport. *The 27th Annual Conference on Neural Information Processing Systems*. NV, United states: Neural Information Processing Systems Foundation, 2013: 2292 – 2300.
- [24] XIE Y, WANG X, WANG R, et al. A fast proximal point method for Wasserstein distance. *ArXiv preprint*, 2018: arXiv: 1802. 04307.
- [25] HU Maohan. *Research on text generation based on generative adversarial network*. Chengdu: University of Electronic Science and Technology of China, 2020. (胡懋晗. 基于生成对抗网络的文本生成的研究. 成都: 电子科技大学, 2020.)
- [26] XU Zhigang, YIN Wenyu, ZHU Xufeng. Color restoration of mural image based on double constrained convolutional neural network. *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, 2020, 48(6): 6 – 12. (徐志刚, 尹文玉, 朱旭锋. 基于双重约束CNN的壁画图像色彩还原. 华中科技大学学报(自然科学版), 2020, 48(6): 6 – 12.)
- [27] ZHU J, CHEN N, SHEN C. A new deep transfer learning method for bearing fault diagnosis under different working conditions. *IEEE Sensors Journal*, 2020, 20(15): 8394 – 8402.
- [28] ZHENG W, WANG C, YANG Y J, et al. Multi-objective combustion optimization based on data-driven hybrid strategy. *Energy*, 2020, 191: 116478 – 116491.
- [29] WANG Rui, CHEN Shiwen, LU Jing. Time-division refined short-term load forecasting based on BOA-SVR and fuzzy clustering. *Engineering Journal of Wuhan University*, 2021, 54(12): 1140 – 1149. (王瑞, 陈诗雯, 逮静. 基于模糊聚类的BOA-SVR分时段精细化短期负荷预测. 武汉大学学报(工学版), 2021, 54(12): 1140 – 1149.)
- [30] ZHANG Xiaochen, ZHU Jinda, YANG Dongmei, et al. Photovoltaic inverter fault prediction technology based on t-sne manifold learning and fast clustering algorithm. *Electric Power*, 2020, 53(6): 41 – 47. (张筱辰, 朱金大, 杨冬梅, 等. 基于t-SNE流形学习与快速聚类算法的光伏逆变器故障预测技术. 中国电力, 2020, 53(6): 41 – 47.)
- [31] BERGSTRA J, BENGIO Y. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 2012, 13: 281 – 305.

作者简介:

卢俊波 硕士研究生, 主要研究方向为负荷预测、机器学习在电力系统中的应用, E-mail: 532969300@qq.com;

刘俊峰 博士, 教授, 博士生导师, 主要研究方向为高频微电网关键技术、电力电子及控制技术, E-mail: aujfliu@scut.edu.cn;

罗燕 硕士研究生, 主要研究方向为微电网能量管理及优化, E-mail: 2905718508@qq.com;

曾君 博士, 教授, 主要研究方向为微电网能量管理及优化、可再生能源发电系统中的电力电子及控制技术, E-mail: junzeng@scut.edu.cn.