

# 控制与决策

Control and Decision

## 迁移学习引导的变源域长短时记忆网络建筑负荷预测

张勇, 陶一凡, 巩敦卫

引用本文:

张勇, 陶一凡, 巩敦卫. 迁移学习引导的变源域长短时记忆网络建筑负荷预测[J]. *控制与决策*, 2021, 36(10): 2328–2338.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2020.0215>

---

### 您可能感兴趣的其他文章

Articles you may be interested in

#### [基于深度时序特征迁移的轴承剩余寿命预测方法](#)

Remaining useful life prediction of bearing based on deep temporal feature transfer

*控制与决策*. 2021, 36(7): 1699–1706 <https://doi.org/10.13195/j.kzyjc.2019.1809>

#### [基于近端强化学习的股价预测方法](#)

Method of stock prices forecast based on proximal reinforcement learning

*控制与决策*. 2021, 36(4): 967–973 <https://doi.org/10.13195/j.kzyjc.2019.1245>

#### [基于近端强化学习的股价预测方法](#)

Method of stock prices forecast based on proximal reinforcement learning

*控制与决策*. 2021, 36(4): 967–973 <https://doi.org/10.13195/j.kzyjc.2019.1245>

#### [基于FWADE-ELM的短时交通流预测方法](#)

Short-term traffic flow forecasting based on hybrid FWADE-ELM

*控制与决策*. 2021, 36(4): 925–932 <https://doi.org/10.13195/j.kzyjc.2019.1103>

#### [结合注意力机制的循环神经网络复述识别模型](#)

Recurrent neural networks based paraphrase identification model combined with attention mechanism

*控制与决策*. 2021, 36(1): 152–158 <https://doi.org/10.13195/j.kzyjc.2019.0638>

# 迁移学习引导的变源域长短时记忆网络建筑负荷预测

张 勇, 陶一凡<sup>†</sup>, 巩敦卫

(中国矿业大学 信息与控制工程学院, 江苏 徐州 221116)

**摘 要:** 历史数据不足会严重影响到长短时记忆网络(long short-term memory, LSTM)预测建筑负荷的精度. 迁移并利用源域中其他相似建筑的用能数据, 可以提高 LSTM 处理目标域中建筑的预测精度, 但现有方法没有考虑预测过程中数据增加所导致的源域匹配关系变化问题. 鉴于此, 提出迁移学习引导的变源域 LSTM 建筑负荷预测方法. 在执行过程中, 根据源域建筑负荷与目标建筑负荷在最新窗口的关联度, 实时调整要选择的源域建筑及其能耗数据, 以确保源域与目标域始终保持较高的相似程度. 在多个典型实例上的应用表明, 相比传统的定源域迁移学习方法, 所提变源域 LSTM 负荷预测方法可以始终保持较高的预测精度.

**关键词:** 负荷预测; 迁移学习; 长短时记忆网络; 变源域; 灰色关联分析

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2020.0215

开放科学(资源服务)标识码(OSID):



**引用格式:** 张勇, 陶一凡, 巩敦卫. 迁移学习引导的变源域长短时记忆网络建筑负荷预测[J]. 控制与决策, 2021, 36(10): 2328-2338.

## Load forecasting of buildings using LSTM based on transfer learning with variable source domain

ZHANG Yong, TAO Yi-fan<sup>†</sup>, GONG Dun-wei

(School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China)

**Abstract:** Insufficient historical data severely affects the accuracy of the long short-term memory (LSTM) in predicting building loads. Transferring and using the energy consumption data of other similar buildings in the source domain can improve the prediction accuracy of LSTM processing of the buildings in the target domain. However, the existing methods do not take into account the change of the source domain matching relationship caused by the increase in data during the prediction process. In view of this, an LSTM-based building load prediction method guided by transfer learning with variable source domain is proposed. During the execution process, according to the correlation between the load of the source domain building and that of the target building in new windows, the source domain building and its energy consumption data to be selected are adjusted in real time to ensure that the source and target domains always remain high similarity. Finally, the application on several typical examples shows that compared with the traditional fixed source domain transfer learning method, the proposed LSTM-based load prediction method guided by transfer learning with variable source domain can always maintain a higher prediction accuracy.

**Keywords:** load forecasting; transfer learning; LSTM; variable source domain; grey relational analysis

## 0 引 言

负荷预测是电力系统高效稳定运行的基础. 准确预测电力负荷可以大大降低运营成本, 提高其稳定性<sup>[1-2]</sup>. 目前, 对电力负荷的预测已有很多有效方法, 包括自回归移动平均(ARMA)<sup>[3]</sup>、灰色模型<sup>[4]</sup>、卡尔曼滤波器<sup>[5]</sup>等. 这些方法使用数学模型将特征映射到负荷消耗, 从而进行预测. 然而, 如何构建准确的数学模型仍是一个开放性难题. 近些年, 基于人工智能

的预测方法逐渐得到学者们的重视, 诸如神经网络、小波分析法和支持向量机等<sup>[6-10]</sup>被成功应用于负荷预测领域.

循环神经网络(recurrent neural network, RNN), 因其具有保留最近事件记忆的能力, 已被广泛运用于自然语言处理<sup>[11]</sup>、计算机视觉<sup>[12]</sup>等热门领域. 针对 RNN 无法记忆久远信息的缺点, 长短时记忆网络(long short-term memory, LSTM)在隐含层内部添

收稿日期: 2020-03-01; 修回日期: 2020-06-29.

基金项目: 国家自然科学基金项目(61876185).

责任编委: 阳春华.

<sup>†</sup>通讯作者. E-mail: 839861804@qq.com.

加遗忘门、输入门和输出门,并增加一条代表长期记忆的信息流<sup>[13]</sup>.目前,LSTM已被用于负荷预测领域.Muzaffar等<sup>[14]</sup>指出,使用LSTM对负荷进行预测的效果要优于ARMA等传统方法;李鹏等<sup>[15]</sup>研究发现,在考虑实时电价情况下的短期负荷预测时,使用LSTM的效果要显著好于BP神经网络和支持向量回归方法;Shi等<sup>[16]</sup>利用LSTM对家庭用电负荷进行预测,进一步证明了它的有效性.然而,利用上述方法对用电负荷进行准确预测时都离不开充足的数据.正如大家所知,在对使用新能源系统的园区或者是新建筑进行负荷预测时,难免遇到数据缺失的情况.如何针对仅具有少量数据的建筑进行较为精确的负荷预测,需要提出新的解决方法.

迁移学习(transfer learning)是一种使用以前解决相关任务或问题时获得的信息来帮助解决新问题的学习机制<sup>[17]</sup>.目前迁移学习已被用于图形识别和网页分类等诸多问题<sup>[18-22]</sup>,但是将迁移学习应用于负荷预测中的相关研究成果依然很少.Hu等<sup>[23]</sup>将迁移学习与深度神经网络相结合,成功解决了数据不足情况下的新建风电场短期风速预测问题;Ma等<sup>[24]</sup>在预测空气质量时,利用迁移学习成功解决了加大时间分辨率后数据的稀疏问题;Zhang等<sup>[25]</sup>的研究表明,利用迁移的源域知识可以显著提升短期负荷预测的准确度;Zeng等<sup>[26]</sup>通过迁移其他城市的数据对当前城市的节假日负荷进行预测,显著提高了城市节假日的负荷预测精度.尽管这些方法都涉及到了源域的选择问题,但是在确定源域后则一直固定不变,并没有考虑预测过程中源域的变化问题.当目标域中数据更加丰富时,这些方法无法确保源域与目标域一直保持较高的相似程度.如果能在预测过程中动态监测目标域与源域数据之间的匹配程度,时刻选择最为匹配的源域数据,势必会持续保证负荷预测的精度.

鉴于此,本文提出一种迁移学习引导的变源域LSTM建筑负荷预测方法,用来解决目标域中数据不充足、源域与目标域中数据匹配程度动态变化的问题.首先,引入一种基于灰色关联分析的关联度指标,用来评价目标域和源域中建筑负荷数据的匹配程度;接着,提出一种源域自主更新策略.在算法执行过程中,根据源域建筑的负荷与目标建筑的负荷在当前窗口的关联度,实时调整要选择的源域建筑及其能耗数据,以确保源域与目标域保持较高的相似程度.当找到相匹配的源域建筑后,迁移其负荷数据到目标域,用于训练面向当前建筑的LSTM负荷预测模型.

## 1 基础工作

### 1.1 LSTM网络

LSTM网络是Hochreiter等<sup>[27]</sup>提出的一种递归神经网络体系结构.LSTM由存储信息的若干存储单元组成,每个存储单元通过3个特殊门(输入门、遗忘门和输出门)实现信息的保护和控制.LSTM单元结构如图1所示.其中: $h_{t-1}$ 表示LSTM在前一时刻的输出, $x_t$ 表示输入, $h_t$ 表示当前单元输出, $c_t$ 和 $c_{t-1}$ 表示存储单元当前和上一时刻的状态.

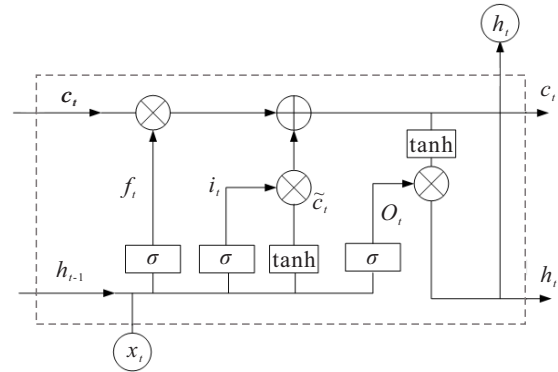


图1 LSTM的基本单元结构

遗忘门 $f_t$ 决定了当前单元保留多少上一单元的信息,其值如下式所示:

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f). \quad (1)$$

其中: $w_f$ 是权重矩阵, $b_f$ 是偏差, $\sigma$ 是sigmoid函数.

输入门 $i_t$ 控制当前输入数据到存储单元状态值的更新,其值如下式所示:

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i). \quad (2)$$

其中: $w_i$ 是权重矩阵, $b_i$ 是偏置.同时,新状态信息 $\tilde{c}_t$ 可由下式得到:

$$\tilde{c}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c). \quad (3)$$

其中: $w_c$ 是权重矩阵, $b_c$ 是偏置.由此,可由下式得到当前状态:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t. \quad (4)$$

输出门 $o_t$ 控制存储单元状态值的输出值,其值如下式所示:

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o), \quad (5)$$

$$h_t = o_t \cdot \tanh(c_t). \quad (6)$$

其中: $w_o$ 是权重矩阵, $b_o$ 是偏置.

借助3个控制门和存储单元,LSTM可以读取、重置和更新长时间的信息.

### 1.2 迁移学习

迁移学习主要包含域和任务两个概念<sup>[28]</sup>.通常,域由两个部分组成,可以表示为 $D = \{Z, P(Z)\}$ ,

即特征空间  $Z$  和特征空间的边缘分布  $P(Z)$ . 其中特征空间  $Z$  可以表示为  $Z = \{z | z_i \in Z, i = 1, 2, \dots, n\}$ . 在给定一个域  $D = \{Z, P(Z)\}$  的情况下, 一个任务可以表示为  $T = \{Y, f(\cdot)\}$ , 即标签空间  $Y$  和一个目标预测函数  $f(\cdot)$ .

现有迁移学习大多只考虑存在一个源域  $D_s$  和一个目标域  $D_t$  的情况, 称为单源域迁移学习<sup>[29]</sup>. 不妨设源域  $D_s = \{(z_{s,1}, y_{s,1}), (z_{s,2}, y_{s,2}), \dots, (z_{s,n}, y_{s,n})\}$ ,  $z_{(s,j)} \in Z_s$  表示源域的观测样本,  $y_{(s,j)} \in Y_s$  表示源域观测样本  $z_{(s,j)}$  对应的标签. 目标域  $D_t = \{(z_{t,1}, y_{t,1}), (z_{t,2}, y_{t,2}), \dots, (z_{t,n}, y_{t,n})\}$ ,  $z_{(t,j)} \in Z_t$  表示目标域观测样本,  $y_{(t,j)} \in Y_t$  表示目标域样本  $z_{(t,j)}$  对应的输出. 基于以上的符号定义, 迁移学习的目的是: 在给定源域  $D_s$  和源域学习任务  $T_s$ 、目标域  $D_t$  和目标域任务  $T_t$ , 且满足  $D_t \neq D_s$  和  $T_t \neq T_s$  的情况下, 通过使用源域  $D_s$  和  $T_s$  中的知识, 提升目标域  $D_t$  中目标预测函数  $f(\cdot)$  的学习效果<sup>[17]</sup>.

## 2 迁移学习引导的变源域LSTM负荷预测

### 2.1 所提负荷预测方法的思想及框架

本节提出一种迁移学习引导的变源域LSTM建筑负荷预测方法. 图2展示了所提变源域LSTM负荷预测方法的框架.

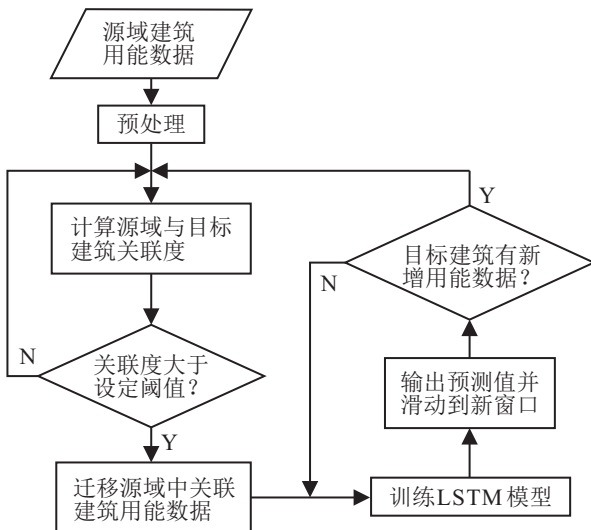


图2 迁移学习引导的变源域LSTM建筑负荷预测方法

首先, 对源域中的历史用能数据进行预处理, 删除异常数据点, 填充缺失数据; 其次, 计算源域中各建筑与目标建筑用能数据之间的关联度, 若两者之间关联度大于设定阈值, 则将当前源域建筑设置为迁移对象; 再次, 将源域中所选迁移对象的用能数据迁移到目标域中, 与目标建筑历史用能数据相融合后, 用于训练LSTM网络; 最后, 输出当前窗口的预测结果, 并转入下一轮窗口. 如果目标建筑有新增用能数据, 则

重新计算源域中各建筑与目标建筑用能数据之间的关联度, 自主切换源域中的迁移对象; 否则, 继续使用上一轮迁移对象中的数据训练LSTM网络, 并输出预测结果.

### 2.2 数据的预处理

对于源域或目标域中缺失的数据, 本文采用常用的后向学习策略, 使用前一时刻数据进行填充处理. 拉伊达准则指出, 当数据服从正态分布时, 数值分布在  $(\mu - 3\sigma, \mu + 3\sigma)$  中的概率为99.74%. 其中:  $\mu$  代表数据平均值,  $\sigma$  代表标准差. 观察部分建筑物的能耗数据发现, 由于用户用能行为具有规律性, 其在某一时刻的能耗数据极大可能服从正态分布. 鉴于此, 依据文献[30]所提方法, 使用拉伊达准则判断数据是否异常. 确定异常数据后, 同样采用后向学习策略, 使用前一时刻数据进行填充处理. 除缺失与异常数据需要处理, 还需归一化相关数据, 将不同建筑的用能数据置于同一量纲, 以提高模型的预测精度. 本文利用下式对数据进行归一化处理:

$$\tilde{y} = \frac{y - \mu}{\sigma}. \quad (7)$$

其中:  $\tilde{y}$  表示归一化后的数据,  $y$  表示归一化之前的数据,  $\mu$  表示所有数据的平均值,  $\sigma$  表示所有数据的标准差.

本文预测过程使用滑动窗口法, 即利用  $[t - \Delta t, t)$  时刻数据来预测  $[t, t + \Delta t)$  时刻目标建筑能耗, 其中  $\Delta t$  为时间步长. 随之, 目标域和源域中的时序数据会被改造成滑动窗口的形式, 具体表现为, 将  $[t - \Delta t, t)$  时刻的数据作为特征,  $[t, t + \Delta t)$  时刻的值作为对应的标签. 以  $\Delta t$  等于7天为例, 第1个滑动窗口中以第1至7天的数据作为特征, 对应第8至14天的值为标签; 第2个滑动窗口以第8至14天的数据作为特征, 对应第15至21天的值为标签; 直到最后, 以  $[t - 7, t)$  的数据作为特征, 对应第  $[t, t + 7)$  的值为标签.

### 2.3 源域选择与自主更新策略

随着时间的变化, 源域建筑和目标建筑间的关联程度会发生变化. 如果源域建筑选取的不合适, 则可能会引起负迁移, 降低模型的预测性能. 鉴于此, 提出一种源域自主更新策略. 在算法执行过程中, 根据源域建筑负荷与目标建筑负荷在当前窗口的关联度, 实时调整要选择的源域建筑及其能耗数据, 以确保源域与目标域始终保持较高的相似程度.

#### 2.3.1 源域建筑的选择

考虑待预测的目标建筑  $T_0$ , 源域中保存的建筑集合为  $T = \{T_k | k = 1, 2, \dots, n\}$ , 如何从  $T$  中选取合适的建筑十分关键. 鉴于此, Zeng等<sup>[26]</sup>在处理城市节

假日负荷预测时,给出了一种使用基于灰色关联分析的数据关联度计算方法。

本文采用灰色关联分析法来衡量建筑间用能数据的相似度. 对于任意建筑  $T_i$ , 其历史用能数据集可表示为  $\{(x_i^{(t)}, y_i^{(t)}), t = 1, 2, \dots, m\}$ . 其中:  $x_i^{(t)}$  为建筑  $T_i$  在时刻  $t$  的特征部分,  $y_i^{(t)}$  为对应的标签部分. 以目标建筑  $T_0$  中的能耗数据  $y_0$  为参考序列, 数据库中建筑  $T_i$  的能耗数据  $y_i$  为比较序列. 下式给出了参考序列与观察序列间某关联样本点的关联系数计算公式:

$$\xi_i(t) = \frac{\min_i \min_t |y_0^{(t)} - y_i^{(t)}| + \rho \max_i \max_t |y_0^{(t)} - y_i^{(t)}|}{|y_0^{(t)} - y_i^{(t)}| + \rho \max_i \max_t |y_0^{(t)} - y_i^{(t)}|}. \quad (8)$$

其中:  $y_0^{(t)}$  表示参考序列在  $t$  时刻的负荷值,  $y_i^{(t)}$  表示观察序列在  $t$  时刻的负荷值;  $\rho$  为0到1内取值的分辨系数, 一般取0.5.

进一步, 计算关联系数的平均值, 该值即为建筑  $T_0$  和  $T_i$  的关联度值:

$$r(i) = \frac{1}{N} \sum_{t=1}^N \xi_i(t). \quad (9)$$

如第2.2节所述, 在很多情况下只利用单个源域建筑的数据很难明显改善预测模型的精度. 在目标域数据极度缺乏的情况下, 尽可能多地使用更多源域建筑的知识, 将有助于持续改善预测模型. 鉴于此, 本文采用多源域迁移策略, 从多个相匹配的源域建筑提取训练数据. 一方面, 选取源域中所有建筑作为学习目标, 不仅会大大增加计算成本, 而且可能出现负迁移现象. 另一方面, 如果选择的源域建筑太少, 则可能得不到足够多的训练数据, 进而影响模型的预测精度. 为此, 本文设定阈值  $\varepsilon$ , 选择所有与目标建筑关联度高于  $\varepsilon$  的源域建筑作为迁移对象. 进一步, 本文第3.1节讨论了  $\varepsilon$  的取值.

### 2.3.2 源域的自主更新策略

本文采用滑动窗口预测不同窗口下目标建筑的能耗. 随着窗口的不断推进, 目标建筑将会持续得到新的能耗数据; 随着目标建筑数据的增加, 前期源域建筑和目标建筑的相似程度势必发生变化. 以7天为一个窗口, 图3展示了源域中某一建筑与目标建筑关联度在42周内的变化情况. 可以看出, 随着目标建筑数据的不断丰富, 两个建筑之间的关联度变化产生了明显波动, 最大关联度值为0.774, 最低值仅为0.575. 在预测初期, 选择该建筑作为迁移对象可提供高质量的训练数据, 但是随着时间的推移, 该建筑引

起负迁移的概率逐渐变大.

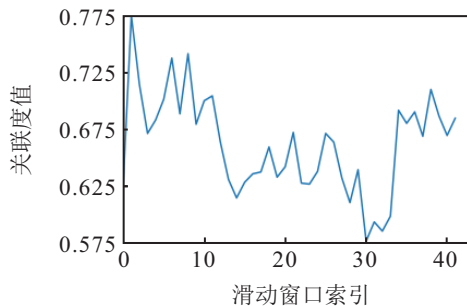


图3 源域建筑与目标建筑关联度的变化曲线

在判断目标问题和源域问题的关联程度时, 传统迁移学习通常计算整个时序数据的相近程度, 称之为整体关联度. 相对滑动关联度, 整体关联度主要用来评价目标建筑与其他建筑在较长时间内的相关程度, 在相应时间段内其值通常静态不变. 滑动关联度则用来评价目标建筑与其他建筑在较短周期内的相关程度, 随着窗口的滑动, 其值会动态变化. 在一个较长时间段(包含多个预测周期)内, 可能出现某个周期时段两个建筑的关联度较高、但另一个周期时段关联度较低的情况. 此时, 两个建筑的整体关联度可能仍然较高.

鉴于此, 本文给出一种源域自主更新策略. 采用滑动关联度周期性地监测两个建筑耗能行为的相似程度, 并根据源域建筑与目标建筑关联程度的变化, 在每个窗口中自主确定合适的迁移对象. 因为可以及时反映两个建筑近期关联度的变化, 滑动关联度可以保证迁移数据的时效性. 这其中, 源域建筑与目标建筑关联度的变化分为两种: 1) 某源域建筑与目标建筑的关联度在新滑动窗口中低于阈值  $\varepsilon$ ; 2) 某建筑与目标建筑的关联度在先前窗口中低于阈值  $\varepsilon$ , 但在新的滑动窗口高于阈值  $\varepsilon$ . 对于第1)种情况, 当关联度在最近一个窗口低于阈值时, 将该建筑直接从迁移集合中剔除, 以防其影响训练模型的准确度. 对于第2)种情况, 将该新建筑加入到迁移对象集合中, 但需要从该建筑数据中去除关联度低于  $\varepsilon$  的窗口数据. 随后, 利用目标域中高相似度建筑的能耗数据, 填充缺失数据. 缺失数据填充公式如下:

$$\tilde{y}_i(t + \Delta t) = \frac{\bar{y}_i(t + \Delta t)}{\bar{y}_0(t + \Delta t)} y_0(t + \Delta t). \quad (10)$$

其中:  $\tilde{y}_i(t + \Delta t)$  代表数据填充后建筑  $i$  的负荷数据,  $\bar{y}_i(t + \Delta t)$  与  $\bar{y}_0(t + \Delta t)$  分别代表源域建筑  $i$  与目标建筑在当前时间窗口内负荷数据的平均值,  $y_0(t + \Delta t)$  代表目标建筑在时间窗口内的负荷数据. 这样进行填充的原因如下: 1) 通过第2.3.1节方法选出的源域建筑与目标建筑具有较高相似性的用能规律, 这在

一定程度上意味着,用目标建筑能耗数据填充源域建筑数据是合理的;2)为保证被填充数据与源域数据具有相同的量纲,式(10)采用系数 $\frac{\bar{y}_i(t+\Delta t)}{\bar{y}_0(t+\Delta t)}$ 对其进行了统一化;3)填充后的源域数据将被用来训练目标域中使用的预测模型.因此,相对于基于源域自身数据进行填充的传统方法,如均值填充、特殊值填充等,利用目标域数据产生填充数据更为合适.

## 2.4 隐藏层共享的LSTM模型

选取了合适的源域建筑后,利用目标建筑和源域建筑的能耗数据共同训练模型.所用LSTM训练模型如图4所示,该模型使用两层LSTM网络.本文利用前 $m$ 天的数据预测下一个周期的负荷,故模型的输入是目标建筑与源域建筑 $m$ 天的负荷数据.正如第2.3节所述,随着窗口的滑动,选取的源域建筑也随之变化,故此模型的输入也会随之改变.

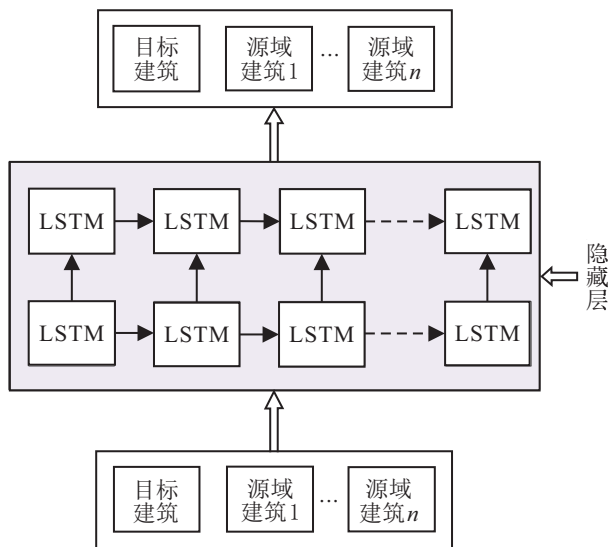


图4 隐藏层共享的LSTM训练模型

模型隐藏层是共享的,这意味着源域数据同样会被选作特征来使用,这是一种特征的转换过程<sup>[31]</sup>.基于此,Tan等<sup>[32]</sup>利用LSSVM对一个综合能源系统的负荷进行预测;同时,Shih等<sup>[33]</sup>将共享隐藏层思想与LSTM相结合进行多变量负荷预测.这些研究表明,对于具有很多源域数据的预测问题,共享隐藏层可以提高模型的预测精度.鉴于此,本文给出隐藏层共享的LSTM预测模型,用以解决建筑能耗预测问题.

## 3 实验部分

本节选择Carbon Culture平台(<https://platform.carbonculture.net>)提供的10所学校作为测试对象,以验证所提算法的有效性.实验中,数据以15 min为间隔统计,从而将一天分为48个时刻点;利用前一周的负荷数据来预测后一周的负荷,即利用前面336个

时刻预测后面的336个时刻;滑动窗口为7天.实验将10所学校中的8所学校视为数据库中已有源域建筑,剩余2所作为目标建筑.

实验采用均方根误差(RMSE)、平均绝对误差(MAE)和平均绝对百分比误差(MAPE)这3种标准来评估预测结果的优劣.RMSE能够直观衡量观测值与真实值之间的偏差;MAE是绝对误差的平均值,相较于RMSE,该测度对小误差更加敏感;MAPE是误差与真实值的比,其值越小表示预测的结果越准确.

### 3.1 阈值的选取

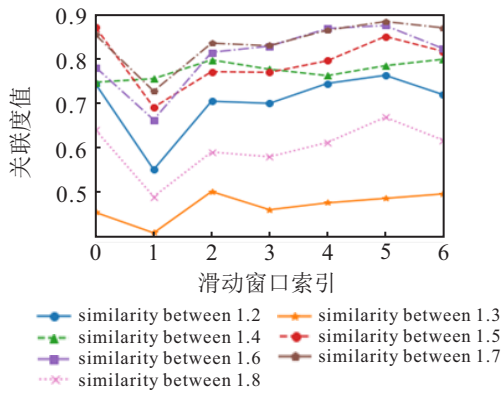
在从源域中选择迁移对象时,阈值 $\epsilon$ 的选取至关重要.若 $\epsilon$ 值过小,则会将目标关联度较低的建筑选入源域中,导致负迁移;若 $\epsilon$ 值过大,则可能漏选可供学习的知识,影响预测模型的准确度.根据式(8),源域建筑与目标建筑相对应负荷之间差值的最大和最小值会影响其关联度,因此,即使两个同样建筑间的关联度在不同数据库中的值也有可能不同.鉴于此,在选取源域时需要提前利用已有建筑的历史负荷数据确定出阈值 $\epsilon$ .

针对源域数据库中已有的8所学校,从中随机选择3所作为待预测的目标学校,如校1、2和8.假设已有2018年4月1日至4月28日共计4周的数据训练模型,需要预测其后4周的数据.图5展示了源域中各学校与目标学校之间的关联度曲线.可以看出,学校之间的关联度值大都高于0.6.当关联度值大于0.8时,源域中相似学校的数目显著减少,在某些窗口甚至出现相似学校数目为0的情况(如图5(a)中窗口1).鉴于此, $\epsilon$ 分别取0.65、0.70和0.75,实验分析不同 $\epsilon$ 取值对算法预测结果的影响.针对目标学校1、2和8,表1列出了当 $\epsilon$ 取不同值时被选中的源域学校.

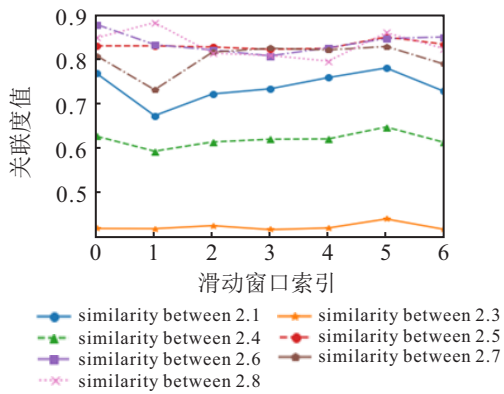
表1 不同阈值情况下各目标学校所选源域学校

| 阈值   | 目标学校1      | 目标学校2         | 目标学校8      |
|------|------------|---------------|------------|
| 0.65 | 4, 5, 6, 7 | 1, 5, 6, 7, 8 | 2, 5, 6, 7 |
| 0.70 | 4, 7       | 5, 6, 7, 8    | 2, 5, 6    |
| 0.75 | 4          | 5, 6, 8       | 2, 6       |

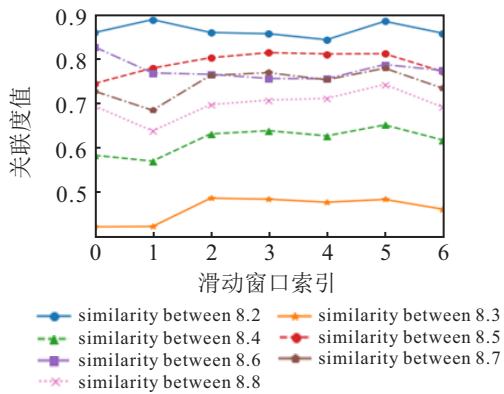
在不同阈值下,表2~表4分别展示了关于校1、2和8的能耗预测结果.可以看出:当阈值取0.7时,算法所得结果的误差大多要比阈值为0.65时小;当阈值分别取0.7和0.75时,算法所得结果的差异不显著.当预测日期位于5月6日至5月12日时,对于目标学校1和2,带迁移算法的预测精度要小于无迁移算法的预测精度.这主要是因为周一(5月7日)变成了节假日,而前面用到的训练数据中周一是工作日.综合表2~表4中的结果, $\epsilon$ 取0.7或0.75是较为合适的.



(a) 学校1与目标学校在各滑动窗口中的关联度值



(b) 学校2与目标学校在各滑动窗口中的关联度值



(c) 学校8与目标学校在各滑动窗口中的关联度值

图5 源域学校与目标学校在各滑动窗口中的关联度值

表2 取不同阈值时算法对目标学校1的预测误差

| 日期                   | 误差   | 阈值     |        |        |        |
|----------------------|------|--------|--------|--------|--------|
|                      |      | 不迁移    | 0.65   | 0.7    | 0.75   |
| 4月29日~5月5日<br>(第1周)  | RMSE | 16.962 | 3.488  | 3.357  | 4.472  |
|                      | MAE  | 9.992  | 2.525  | 2.200  | 3.061  |
|                      | MAPE | 19.611 | 8.573  | 7.004  | 9.494  |
| 5月6日~5月12日<br>(第2周)  | RMSE | 9.704  | 11.990 | 11.922 | 12.068 |
|                      | MAE  | 4.488  | 5.184  | 5.775  | 5.877  |
|                      | MAPE | 20.009 | 21.948 | 24.095 | 23.913 |
| 5月13日~5月19日<br>(第3周) | RMSE | 4.844  | 4.673  | 4.260  | 3.368  |
|                      | MAE  | 3.373  | 3.462  | 3.227  | 2.592  |
|                      | MAPE | 10.447 | 11.062 | 10.831 | 8.864  |
| 5月20日~5月26日<br>(第4周) | RMSE | 4.375  | 4.61   | 3.499  | 10.591 |
|                      | MAE  | 3.207  | 3.323  | 2.677  | 3.909  |
|                      | MAPE | 10.222 | 10.295 | 9.601  | 9.616  |

表3 取不同阈值时算法对目标学校2的预测误差

| 日期                   | 误差   | 阈值     |        |        |        |
|----------------------|------|--------|--------|--------|--------|
|                      |      | 不迁移    | 0.65   | 0.7    | 0.75   |
| 4月29日~5月5日<br>(第1周)  | RMSE | 18.337 | 4.910  | 4.525  | 4.618  |
|                      | MAE  | 10.798 | 2.816  | 2.784  | 2.749  |
|                      | MAPE | 26.517 | 10.202 | 11.063 | 11.453 |
| 5月6日~5月12日<br>(第2周)  | RMSE | 8.631  | 11.302 | 10.380 | 10.411 |
|                      | MAE  | 4.499  | 5.390  | 4.758  | 5.095  |
|                      | MAPE | 27.405 | 30.503 | 28.353 | 29.284 |
| 5月13日~5月19日<br>(第3周) | RMSE | 8.066  | 5.968  | 5.526  | 5.060  |
|                      | MAE  | 33.483 | 3.205  | 2.704  | 2.698  |
|                      | MAPE | 11.631 | 11.411 | 10.136 | 10.485 |
| 5月20日~5月26日<br>(第4周) | RMSE | 5.206  | 3.672  | 3.586  | 6.568  |
|                      | MAE  | 2.834  | 2.144  | 2.088  | 3.675  |
|                      | MAPE | 11.055 | 9.071  | 9.177  | 13.176 |

表4 取不同阈值时算法对目标学校8的预测误差

| 日期                   | 误差   | 阈值     |        |        |        |
|----------------------|------|--------|--------|--------|--------|
|                      |      | 不迁移    | 0.65   | 0.7    | 0.75   |
| 4月29日~5月5日<br>(第1周)  | RMSE | 5.852  | 4.582  | 3.507  | 4.545  |
|                      | MAE  | 2.789  | 2.195  | 1.855  | 2.192  |
|                      | MAPE | 17.855 | 14.167 | 12.733 | 14.268 |
| 5月6日~5月12日<br>(第2周)  | RMSE | 7.660  | 7.293  | 7.385  | 7.558  |
|                      | MAE  | 3.045  | 2.950  | 3.324  | 3.318  |
|                      | MAPE | 25.973 | 24.826 | 25.432 | 25.511 |
| 5月13日~5月19日<br>(第3周) | RMSE | 3.543  | 3.480  | 3.357  | 3.346  |
|                      | MAE  | 2.276  | 2.120  | 2.119  | 2.005  |
|                      | MAPE | 15.055 | 12.504 | 13.714 | 11.794 |
| 5月20日~5月26日<br>(第4周) | RMSE | 4.436  | 4.222  | 3.267  | 3.303  |
|                      | MAE  | 2.769  | 2.691  | 2.136  | 1.989  |
|                      | MAPE | 19.623 | 19.504 | 17.059 | 15.491 |

3.2 滑动关联度与整体关联度的对比

本节通过与传统关联度计算方式(简称整体关联方式)比较,验证所提滑动关联度计算方式的有效性.在取 $\epsilon = 0.7$ 的情况下,利用2018年4月1日至4月28日4周的数据预测后4周,表5展示了分别采用整体关联度与滑动关联度两种方式所选源域学校.可以看出,滑动关联方式所选源域学校数明显少于整体关联方式所选学校数,且前者所选学校都包含于后者所选学校集合中.

表5 采用不同关联度时各目标学校所选源域学校

| 目标学校 | 整体关联          | 滑动关联       |
|------|---------------|------------|
| 1    | 2, 4, 5, 6, 7 | 4, 7       |
| 2    | 1, 5, 6, 7, 8 | 5, 6, 7, 8 |
| 8    | 2, 5, 6, 7, 8 | 2, 5, 6    |

表6~表8展示了校1、2、8分别采用整体关联与滑动关联方式时算法所得预测结果.可以看出,在绝

大多数情况下,滑动关联方式所得结果都要好于整体关联方式所得结果.以目标学校1为例,除去第3周的MAPE以及第4周的MAE与MAPE,滑动关联方式所得结果都要好于整体关联方式所得结果.

表6 校1采用整体关联与滑动关联方式时算法所得预测误差

| 日期                   | 误差   | 整体关联   | 滑动关联   |
|----------------------|------|--------|--------|
| 4月29日~5月5日<br>(第1周)  | RMSE | 3.366  | 3.357  |
|                      | MAE  | 2.561  | 2.200  |
|                      | MAPE | 8.993  | 7.004  |
| 5月6日~5月12日<br>(第2周)  | RMSE | 12.760 | 11.922 |
|                      | MAE  | 6.030  | 5.775  |
|                      | MAPE | 24.287 | 24.095 |
| 5月13日~5月19日<br>(第3周) | RMSE | 4.735  | 4.260  |
|                      | MAE  | 3.449  | 3.227  |
|                      | MAPE | 10.678 | 10.831 |
| 5月20日~5月26日<br>(第4周) | RMSE | 5.100  | 3.499  |
|                      | MAE  | 2.637  | 2.677  |
|                      | MAPE | 7.183  | 9.601  |

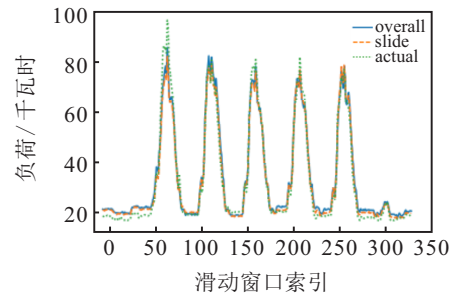
表7 校2采用整体关联与滑动关联方式时算法所得预测误差

| 日期                   | 误差   | 整体关联   | 滑动关联   |
|----------------------|------|--------|--------|
| 4月29日~5月5日<br>(第1周)  | RMSE | 4.910  | 4.525  |
|                      | MAE  | 2.816  | 2.784  |
|                      | MAPE | 10.202 | 11.063 |
| 5月6日~5月12日<br>(第2周)  | RMSE | 11.302 | 10.380 |
|                      | MAE  | 5.390  | 4.758  |
|                      | MAPE | 30.503 | 28.353 |
| 5月13日~5月19日<br>(第3周) | RMSE | 5.968  | 5.526  |
|                      | MAE  | 3.205  | 2.704  |
|                      | MAPE | 11.411 | 10.136 |
| 5月20日~5月26日<br>(第4周) | RMSE | 3.672  | 3.586  |
|                      | MAE  | 2.144  | 2.088  |
|                      | MAPE | 9.071  | 9.177  |

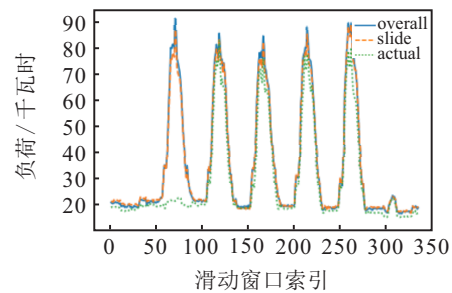
表8 校8采用整体关联与滑动关联方式时算法所得预测误差

| 日期                   | 误差   | 整体关联   | 滑动关联   |
|----------------------|------|--------|--------|
| 4月29日~5月5日<br>(第1周)  | RMSE | 3.840  | 3.507  |
|                      | MAE  | 2.311  | 1.855  |
|                      | MAPE | 13.785 | 12.733 |
| 5月6日~5月12日<br>(第2周)  | RMSE | 7.342  | 7.385  |
|                      | MAE  | 2.969  | 3.324  |
|                      | MAPE | 24.895 | 25.432 |
| 5月13日~5月19日<br>(第3周) | RMSE | 3.504  | 3.357  |
|                      | MAE  | 2.151  | 2.119  |
|                      | MAPE | 13.105 | 13.714 |
| 5月20日~5月26日<br>(第4周) | RMSE | 4.741  | 3.267  |
|                      | MAE  | 2.814  | 2.136  |
|                      | MAPE | 18.888 | 17.059 |

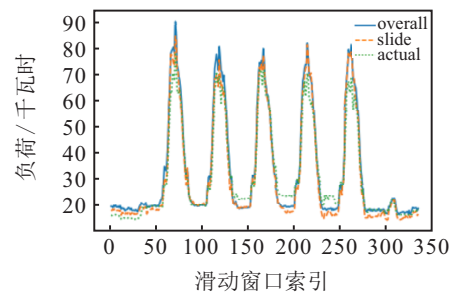
图6~图8展示了3所目标学校4周的预测结果,其中:实线代表使用整体关联方式所得结果,虚线代表滑动关联方式所得结果,点线代表真实值.



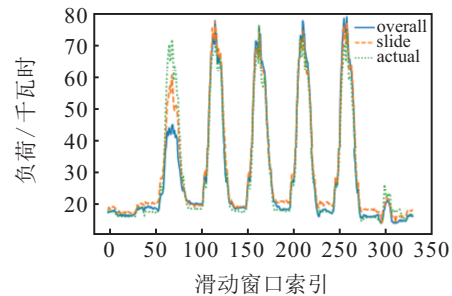
(a) 第1周预测结果



(b) 第2周预测结果



(c) 第3周预测结果

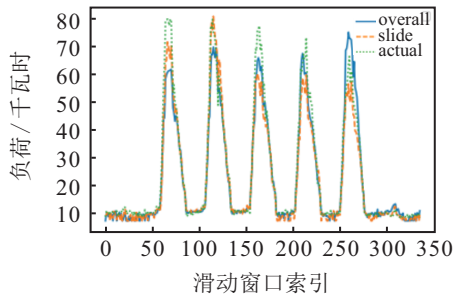


(d) 第4周预测结果

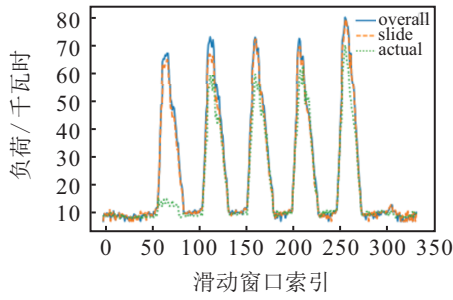
图6 针对目标学校1,整体关联与滑动关联方式下算法所得预测结果

### 3.3 对比结果分析

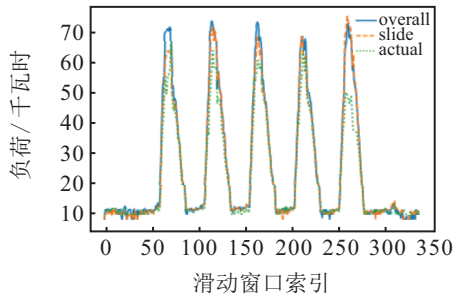
为了验证所提算法的有效性,选择两种对比算法.第一种是不含迁移策略的传统LSTM方法,简称LSTM/R;另一种是定源域的LSTM建筑负荷预测方法,简称Fixed LSTM.与本文所提算法不同,在初期确定源域学校后,Fixed LSTM算法将持续使用这些源域学校的数据.假设新增两所学校,学校9与学校10,表9和表10分别展示了3种算法算法处理目标学校9和学校10所得预测结果.



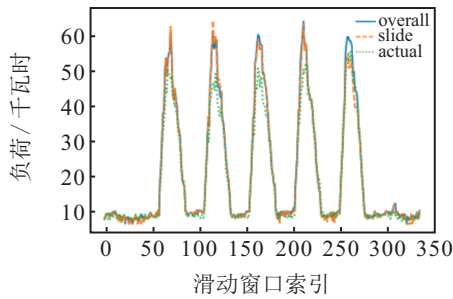
(a) 第1周预测结果



(b) 第2周预测结果

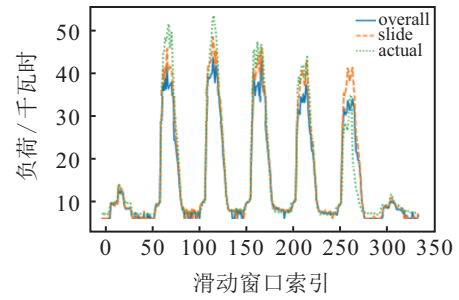


(c) 第3周预测结果

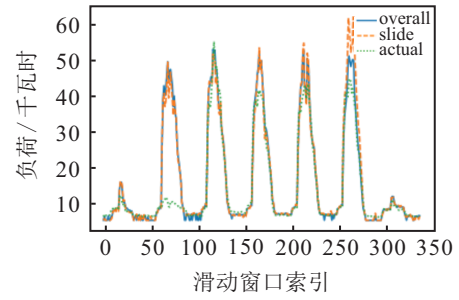


(d) 第4周预测结果

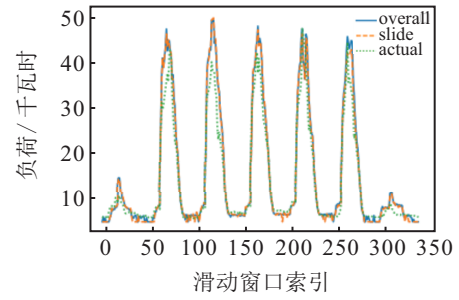
图7 针对目标学校2,整体关联与滑动关联方式下算法所得预测结果



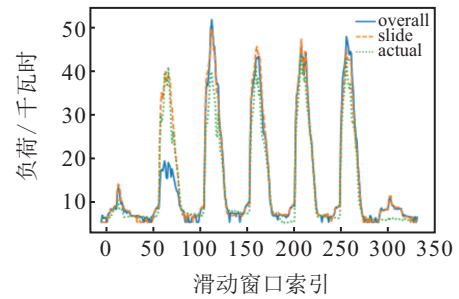
(a) 第1周预测结果



(b) 第2周预测结果



(c) 第3周预测结果



(d) 第4周预测结果

图8 针对目标学校8,整体关联与滑动关联方式下算法所得预测结果

表9 3种算法处理目标学校9所得预测误差

| 日期                   | 误差   | LSTM/R | Fixed LSTM | 所提算法  |
|----------------------|------|--------|------------|-------|
| 6月5日~6月11日<br>(第1周)  | RMSE | 7.756  | 4.077      | 4.077 |
|                      | MAE  | 3.511  | 2.459      | 2.459 |
|                      | MAPE | 12.464 | 9.606      | 9.606 |
| 6月12日~6月18日<br>(第2周) | RMSE | 7.356  | 3.475      | 3.252 |
|                      | MAE  | 3.293  | 2.345      | 1.948 |
|                      | MAPE | 11.002 | 8.791      | 7.049 |
| 6月19日~6月25日<br>(第3周) | RMSE | 4.430  | 4.712      | 3.803 |
|                      | MAE  | 2.576  | 2.609      | 1.889 |
|                      | MAPE | 9.961  | 10.838     | 7.966 |
| 6月26日~7月2日<br>(第4周)  | RMSE | 3.540  | 3.206      | 2.907 |
|                      | MAE  | 2.214  | 2.134      | 1.951 |
|                      | MAPE | 8.271  | 7.978      | 7.722 |

表10 3种算法处理目标学校10所得预测误差

| 日期                   | 误差   | LSTM/R | Fixed LSTM | 所提算法   |
|----------------------|------|--------|------------|--------|
| 6月5日~6月11日<br>(第1周)  | RMSE | 8.885  | 6.895      | 5.442  |
|                      | MAE  | 3.575  | 2.999      | 2.786  |
|                      | MAPE | 11.616 | 10.208     | 9.695  |
| 6月12日~6月18日<br>(第2周) | RMSE | 9.418  | 4.176      | 3.080  |
|                      | MAE  | 4.100  | 2.600      | 2.219  |
|                      | MAPE | 13.560 | 9.021      | 8.311  |
| 6月19日~6月25日<br>(第3周) | RMSE | 4.004  | 4.006      | 3.356  |
|                      | MAE  | 3.269  | 3.123      | 2.817  |
|                      | MAPE | 12.825 | 11.728     | 11.466 |
| 6月26日~7月2日<br>(第4周)  | RMSE | 6.128  | 3.892      | 3.741  |
|                      | MAE  | 3.339  | 2.689      | 2.601  |
|                      | MAPE | 11.017 | 9.507      | 9.006  |

由表9、表10可以看出:1)相比于传统LSTM/R算法,在大多数情况下基于迁移学习的Fixed LSTM和本文算法都得到了较小的预测误差;少数情况下,如预测目标学校9在第3周能耗时,受迁移对象相关性变差等因素影响,Fixed LSTM的预测误差要大于传统LSTM算法。2)随着时间的推进,3种算法的预测精度都呈现不同程度的改善。以表9中本文所提算法为例,预测第1周到第4周能耗时,其RMSE值从4.077逐渐下降到2.907。特别地,由于目标学校自身数据的不断增加,传统LSTM/R算法的预测精度得到

明显改善。以表9中传统LSTM/R算法为例,预测第1周到第4周能耗时,其RMSE值从7.756逐渐下降到3.540。然而,即便是在目标学校数据相对充足的第4周,在迁移学习的帮助下本文所提算法的预测精度依然好于传统LSTM/R算法。3)除第1周Fixed LSTM和本文算法得到相同预测结果外,对于剩余3周,本文算法的预测精度都要好于Fixed LSTM。

图9和图10展示了对于目标学校9、10四周的负荷预测结果。可以看出,多数情况下,本文算法的预测结果比LSTM/R和Fixed LSTM更靠近真实结果。

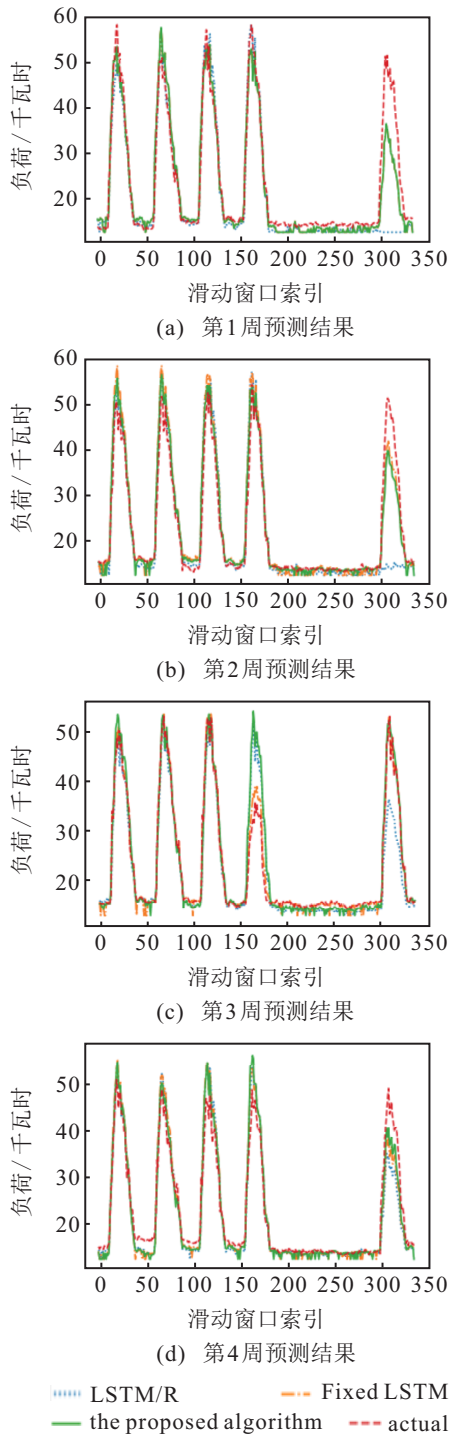


图9 针对目标学校9,3种算法四周的负荷预测结果

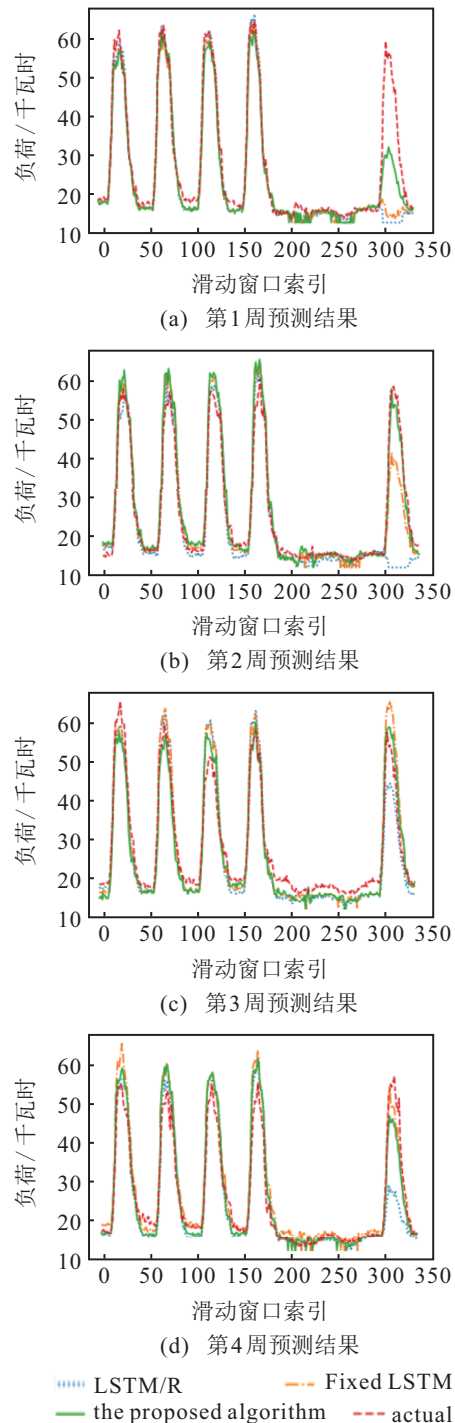


图10 针对目标学校10,3种算法四周的负荷预测结果

由此可见,相比于LSTM/R和Fixed LSTM,本文所提算法不仅可以在数据不足的情况下得出满意的预测结果,而且在数据相对充足的情况下也可以得到更精确的预测结果。

### 3.4 进一步讨论

与常用的BP神经网络预测模型进行对比,以进一步验证本文所提方法的有效性。BP神经网络的参数如下:输入节点 $d$ 与输出节点 $l$ 都为336;根据经验公式,隐藏层个数 $q = 2d + 1 = 773$ ,学习率 $lr$ 取0.01。表11和表12分别展示了本文算法和BP神经网络处理目标学校9和学校10时所得预测误差。可以看到,随着数据的逐渐增多,BP神经网络的预测准确度有所提升;但是,在4周时间内,其预测精度明显差于本文所提算法。这主要是因为,LSTM本质是针对时间序列预测问题提出,相比传统神经网络,它能够更为准确预测时序数据的变化规律。这进一步验证了本文选择LSTM进行能耗预测的合理性。

表 11 本文算法和BP网络处理学校9时所得预测误差

| 日期                   | 误差   | BP     | 所提算法  |
|----------------------|------|--------|-------|
| 6月5日~6月11日<br>(第1周)  | RMSE | 19.088 | 4.077 |
|                      | MAE  | 14.820 | 2.459 |
| 6月12日~6月18日<br>(第2周) | RMSE | 16.660 | 3.252 |
|                      | MAE  | 13.851 | 1.948 |
| 6月19日~6月25日<br>(第3周) | RMSE | 16.131 | 3.803 |
|                      | MAE  | 13.658 | 1.889 |
| 6月26日~7月2日<br>(第4周)  | RMSE | 15.585 | 2.907 |
|                      | MAE  | 13.465 | 1.951 |

表 12 本文算法和BP网络处理学校10时所得预测误差

| 日期                   | 误差   | BP     | 所提算法  |
|----------------------|------|--------|-------|
| 6月5日~6月11日<br>(第1周)  | RMSE | 19.431 | 5.442 |
|                      | MAE  | 16.196 | 2.786 |
| 6月12日~6月18日<br>(第2周) | RMSE | 18.887 | 3.080 |
|                      | MAE  | 15.987 | 2.219 |
| 6月19日~6月25日<br>(第3周) | RMSE | 18.073 | 3.356 |
|                      | MAE  | 14.910 | 2.817 |
| 6月26日~7月2日<br>(第4周)  | RMSE | 16.626 | 3.741 |
|                      | MAE  | 12.883 | 2.601 |

## 4 结论

为有效解决目标建筑历史数据偏少的情况,本文提出了一种基于迁移学习的变源域LSTM负荷预测方法。以Carbon Culture平台上提供的10所学校作为测试对象进行实验,结果表明,相比于不使用迁移学习的传统LSTM/R算法,在大多数情况下基于迁移学习的Fixed LSTM和本文所提算法都得到了高的预测精度;在源域自主更新等策略的帮助下,本文所提算法的预测精度又普遍高于Fixed LSTM。融合日历和

经济等其他常见的负荷预测特征,进一步改善所提算法处理节假日等特殊日期的能力,是需要进一步研究的问题。

### 参考文献(References)

- [1] 李军. 基于贪心核特征提取方法的中期峰值负荷预测[J]. 控制与决策, 2014, 29(9): 1661-1666.  
(Li J. Greedy kernel feature extraction method for medium term electricity peak load forecasting[J]. Control and Decision, 2014, 29(9): 1661-1666.)
- [2] Ranaweera D K, Karady G G, Farmer R G. Economic impact analysis of load forecasting[J]. IEEE Transactions on Power Systems, 1997, 12(3): 1388-1392.
- [3] Li Y Y, Han D, Yan Z. Long-term system load forecasting based on data-driven linear clustering method[J]. Journal of Modern Power Systems and Clean Energy, 2018, 6(2): 306-316.
- [4] Jin M, Zhou X, Zhang Z M, et al. Short-term power load forecasting using grey correlation contest modeling[J]. Expert Systems with Applications, 2012, 39(1): 773-779.
- [5] Infield D G, Hill D C. Optimal smoothing for trend removal in short term electricity demand forecasting[J]. IEEE Transactions on Power Systems, 1998, 13(3): 1115-1120.
- [6] 牛东晓, 王建军, 李莉, 等. 基于粗糙集和决策树的自适应神经网络短期负荷预测方法[J]. 电力自动化设备, 2009, 29(10): 30-34.  
(Niu D X, Wang J J, Li L, et al. Short-term load forecasting using adaptive ANN based on rough set and decision tree[J]. Electric Power Automation Equipment, 2009, 29(10): 30-34.)
- [7] Pandey A S, Singh D, Sinha S K. Intelligent hybrid wavelet models for short-term load forecasting[J]. IEEE Transactions on Power Systems, 2010, 25(3): 1266-1273.
- [8] 曾鸣, 吕春泉, 田廓, 等. 基于细菌群落趋药性优化的最小二乘支持向量机短期负荷预测方法[J]. 中国电机工程学报, 2011, 31(34): 93-99.  
(Zeng M, Lü C Q, Tian K, et al. Least squares-support vector machine load forecasting approach optimized by bacterial colony chemotaxis method[J]. Proceedings of the CSEE, 2011, 31(34): 93-99.)
- [9] 张素香, 赵丙镇, 王风雨, 等. 海量数据下的电力负荷短期预测[J]. 中国电机工程学报, 2015, 35(1): 37-42.  
(Zhang S X, Zhao B Z, Wang F Y, et al. Short-term power load forecasting based on big data[J]. Proceedings of the CSEE, 2015, 35(1): 37-42.)
- [10] Haykin S S. 神经网络原理[M]. 北京: 机械工业出版社, 2004: 1-30.  
(Haykin S S. Neural network principle[M]. Beijing: China Machine Press, 2004: 1-30.)
- [11] Kalchbrenner N, Blunsom P. Recurrent continuous translation models[C]. Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. Seattle: Association for Computational Linguistics, 2013: 1700-1709.

- [12] Donahue J, Hendricks L A, Guadarrama S, et al. Long-term recurrent convolutional networks for visual recognition and description[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Bostan: IEEE, 2015: 2625-2634.
- [13] Cao J, Li Z, Li J. Financial time series forecasting model based on CEEMDAN and LSTM[J]. Physica A: Statistical Mechanics and its Applications, 2019, 519: 127-139.
- [14] Muzaffar S, Afshari A. Short-term load forecasts using LSTM networks[J]. Energy Procedia, 2019, 158: 2922-2927.
- [15] 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[J]. 电网技术, 2018, 42(12): 4045-4052.  
(Li P, He S, Han P F, et al. Short-term load forecasting of smart grids based on long short-term memory under real-time electricity prices[J]. Power System Technology, 2018, 42(12): 4045-4052.)
- [16] Shi H, Xu M H, Li R. Deep learning for household load forecasting — A novel pooling deep RNN[J]. IEEE Transactions on Smart Grid, 2018, 9(5): 5271-5280.
- [17] Pan S J, Yang Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 22(10): 1345-1359.
- [18] Murre J M J. Transfer of learning in backpropagation and in related neural network models[J]. Confluence the Next Generation Information, 2017, 23(12): 60-64.
- [19] Pratt L Y. Discriminability-based transfer between neural networks[C]. Proceedings of the 5th International Conference on Neural Information Processing Systems. San Francisco: Morgan Kaufmann Publishers Inc., 1993: 204-211.
- [20] Mihalkova L, Huynh T, Mooney R J. Mapping and revising Markov logic networks for transfer learning[C]. Proceedings of the 22nd National Conference on Artificial Intelligence. Vancouver: AAAI, 2007: 608-614.
- [21] Mihalkova L, Mooney R. Transfer learning with Markov logic networks[C]. Proceedings of the ICML-06 Workshop on Structural Knowledge Transfer for Machine Learning. Pittsburgh, 2006.
- [22] Gupta R, Ratnoff L A. Text categorization with knowledge transfer from heterogeneous data sources[C]. Proceedings of the 23rd National Conference on Artificial Intelligence. AAAI, 2008: 842-847.
- [23] Hu Q, Zhang R J, Zhou Y. Transfer learning for short-term wind speed prediction with deep neural networks[J]. Renewable Energy, 2016, 85: 83-95.
- [24] Ma J, Cheng J C P, Lin C Q, et al. Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques[J]. Atmospheric Environment, 2019, 214: 116885.
- [25] Zhang Y L, Luo G M. Short term power load prediction with knowledge transfer[J]. Information Systems, 2015, 53: 161-169.
- [26] Zeng P, Sheng C, Jin M. A learning framework based on weighted knowledge transfer for holiday load forecasting[J]. Journal of Modern Power Systems and Clean Energy, 2019, 7(2): 329-339.
- [27] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Comput, 1997, 9(8): 1735-1780.
- [28] 王俊, 李石君, 杨莎, 等. 一种新的用于跨领域推荐的迁移学习模型[J]. 计算机学报, 2017, 40(10): 2367-2380.  
(Wang J, Li S J, Yang S, et al. A new transfer learning model for cross-domain recommendation[J]. Chinese Journal of Computers, 2017, 40(10): 2367-2380.)
- [29] Zhuang F Z, Qi Z Y, Duan K Y, et al. A Comprehensive Survey on Transfer Learning[J]. Proceedings of the IEEE, 2021, 109(1): 43-76.
- [30] 盛昌. 基于节假日加权迁移学习的短期电力负荷预测[D]. 长沙: 湖南大学信息科学与工程学院, 2016: 12-13.  
(Sheng C. Short-term power load forecasting based on holiday weighted transfer learning[D]. Changsha: College of Computer Science and Electronic Engineering, Changsha University, 2016: 12-13.)
- [31] Bengio Y. Learning deep architectures for AI[J]. Foundations and Trends<sup>o</sup> in Machine Learning, 2009, 2(1): 1-127.
- [32] Tan Z F, De G, Li M L, et al. Combined electricity-heat-cooling-gas load forecasting model for integrated energy system based on multi-task learning and least square support vector machine[J]. Journal of Cleaner Production, 2020, 248: 119252.
- [33] Shih S Y, Sun F K, Lee H. Temporal pattern attention for multivariate time series forecasting[J]. Machine Learning, 2019, 108(8): 1421-1441.

### 作者简介

张勇(1979—), 男, 教授, 博士生导师, 从事智能优化和数据挖掘等研究, E-mail: yongzh401@126.com;

陶一凡(1995—), 男, 硕士生, 从事负荷预测的研究, E-mail: 839861804@qq.com;

巩敦卫(1970—), 男, 教授, 博士生导师, 从事智能数据处理与解析等研究, E-mail: dwgong@vip.163.com.

(责任编辑: 齐 霖)