



基于异构数据的电力短期负荷大数据预测方案

赵海波¹, 相志军², 肖林松¹

(1. 威胜信息技术股份有限公司, 湖南 长沙 410205;

2. 中国电力技术装备有限公司, 北京 100052)

摘要: 随着多种可再生能源电力的接入, 电力系统正在向更智能、更灵活、交互性更高的系统过渡。负荷预测, 特别是针对单个电力客户的短期负荷预测在未来电网规划和运行中发挥着越来越重要的作用。提出了一个基于异构数据的电力短期负荷大数据预测方案, 该方案收集来自智能电表和天气预报的数据, 预处理后将其加载到非关系型数据库中进行存储并做进一步的异构数据处理; 设计并实现了一个长短期记忆递归神经网络模型, 用于确定负荷分布并预测未来 24 h 的住宅小区用电量; 最后利用一个住宅小区的智能电表数据集对提出的短期负荷预测框架进行了测试, 并使用均方根误差和平均绝对百分比误差两个指标, 对比了预测模型与两种经典算法的性能, 验证了所提模型的有效性。

关键词: 短期负荷预测; 长短期记忆网络; 递归神经网络; 聚类; 大数据

中图分类号: U416.216

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2022292

A big data framework for short-term power load forecasting using heterogenous data

ZHAO Haibo¹, XIANG Zhijun², XIAO Linsong¹

1. Willfar Information Technology Co., Ltd., Changsha 410205, China

2. China Electronic Power Equipment and Technology Co., Ltd., Beijing 100052, China

Abstract: The power system is in a transition towards a more intelligent, flexible and interactive system with higher penetration of renewable energy generation, load forecasting, especially short-term load forecasting for individual electric customers plays an increasingly essential role in future grid planning and operation. A big data framework for short-term power load forecasting using heterogenous was proposed, which collected the data from smart meters and weather forecast, pre-processed and loaded it into a NoSQL database that was capable to store and further processing large volumes of heterogeneous data. Then, a long short-term memory (LSTM) recurrent neural network was designed and implemented to determine the load profiles and forecast the electricity consumption for the residential community for the next 24 hours. The proposed framework was tested with a publicly available smart meter dataset of a residential community, of which LSTM's performance was compared with two benchmark algorithms in terms of root mean square error and mean absolute percentage error, and its validity has been verified.

Key words: short-term load forecasting, long short-term memory network, recurrent neural network, clustering, big data

收稿日期: 2022-07-01; 修回日期: 2022-12-06

基金项目: 国家重点研发计划项目 (No.2020YFB0906000, No.2020YFB0906002)

Foundation Items: The National Key Research and Development Program of China (No.2020YFB0906000, No.2020YFB0906002)



0 引言

电力是当今社会最重要的能源之一,收集利用历史电力负荷数据,挖掘其潜在价值,对于电网企业和人们的日常生活都具有重要意义。智能电网利用传感和信息通信技术提高电网能源传递的效率、可靠性和安全性,智能计量的普及有利于电力负荷大数据的收集和分析处理,由此产生的一个好处是基于收集的历史负荷海量数据预测下一个时期的电力需求,可以给发电控制和电网调度提供辅助决策依据。

短期负荷预测(short-term load forecasting, STLF)是电力系统需求响应分析的基础,重要性不言而喻,其一直是新能源系统、微电网控制和电力大数据研究者关注的焦点^[1-4]。就统计模型而言,时间序列建模用于捕捉电力负荷的时间序列特征,多元线性回归利用年、周和日的季节性对每小时的能源需求进行建模^[5]。研究表明,相同信息的复杂特征导致预测准确率较低,从而提出用反向传播(back propagation, BP)神经网络对短期负荷进行预测,以负荷和气象信息为输入,输出目标日的峰谷负荷,取得了良好的预测效果^[6]。Oprea 等^[7]提出可扩展大数据框架收集智能电表和天气传感器的数据,实施并对比了几种机器学习算法,以确定负荷曲线并预测未来 24 h 住宅建筑的用电量。王德文等^[8]采用遗传算法对灰色预测模型进行优化,并对中长期电力负荷进行预测,结果表明,改进后的模型具有较高的精度。

本文阐述了一个大数据预测方案,该方案从各种来源收集相关数据,即来自低压用电信息采集系统中智能电表、天气预报的数据,然后将其预处理并加载到非关系型(NoSQL)数据库中,该数据库能够存储各种数据格式的数据,无须像传统关系数据库那样进行耗时耗力的数据规范化;然后,使用聚类算法识别对用户电力需求有

重大影响的一组关键特征,最后设计并训练提出的长短期记忆(long short-term memory, LSTM)预测模型,根据存储在 NoSQL 数据库中的数据预测未来 24 h 的电力负荷预测结果。

1 STLF 的大数据分析框架

本文提出的方案通过 LSTM 建模电力需求和相关特征参量之间的关系来预测未来短期的电力需求。电力短期负荷预测的 LSTM 流程如图 1 所示。

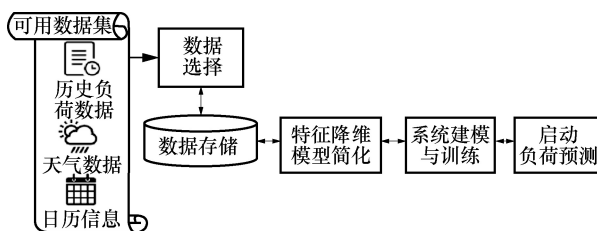


图 1 电力短期负荷预测的 LSTM 流程

依托记录了公寓用电情况的可用数据集合,开发了 3 类与电力需求相关的特征变量,即历史负荷数据、天气数据和日历信息;接着通过特征选择删除不太重要的特征,并降低维数和模型复杂度;最后使用所选择的特征训练本文提出的 LSTM 预测模型。

1.1 数据的收集、处理和存储

大数据分析的挑战之一是处理和分析来自不同来源的海量数据。住宅用户的智能电表通常每隔 15 min 记录和报告电力消耗数据,包括以下信息:时间戳、电表 ID、公寓 ID、有功功率(用户消耗的电能)、无功功率和视在功率等。

天气数据可从气象局网站获得,分辨率通常为 1 h,包含温度、风速、风向、相对湿度、能见度、云量、气压、降水强度和概率等。

智能仪表和气象传感器记录的测量数据需要适当的验证、过滤和转换,以对 STLF 具有价值并保证一致性。因此,对这些原始数据进行预处理,并用插值法将丢失或损坏的值替换为以前时间间隔中最合适的值。从不同来源收集的异构数据为不同格

式和不同时间分辨率的半结构化数据,例如数据格式包括 RAW、JSON、CSV 等。数据处理和存储的解决方案可以是具有强大搜索引擎的 NoSQL 数据库,如 MongoDB、Oracle、NoSQL^[9]。

1.2 聚类 and 降维

选用 3 类数据,即从智能电表收集的历史负荷数据、天气数据和日历信息,在每个类别中,有几个特征对功耗的影响是不均匀的,因此需要分析每个特征的重要性,并选择一组合适的特征在本文的预测模型中使用。

对于历史负荷数据,将消费者分组到相关集群并获得其负载概况,不同的消费模式比总的聚合消费更具可预测性。除了消费数据,还可以从网络调查或问卷中提取和处理有用的见解。此外,负荷曲线通过强调其对峰值和非峰值的贡献,提供了对电力消耗结构更准确的观点。因此,通过估计每个集群的用电量,可以减少预测误差,提高 STLF 的性能。

天气数据有很多属性,需要应用相关分析法选择影响用电量的最重要属性。通过对不相关属性进行噪声滤波,减少了输入维数和后续计算时间,提高了精度。此外,考虑大功率电器的利用率在不同用户间的巨大差异,需要对每个季节的用电量进行聚类分析。

1.3 建模和训练

作为一种深度学习方法,LSTM 通过使用大量的训练样本,挖掘数据中包含的隐藏信息,实现了对高维函数的更多逼近,增强了基于多层非线性变换的建模能力。此外,LSTM 擅长处理时间序列问题。而每个用户的用电量本身就是一种时间序列,因此本文使用 LSTM 预测用电量。

LSTM 的核心思想是一个存储单元可以随时保持由各种门控单元控制的信息^[10-11]。LSTM 的实现结构如图 2 所示。各个门控单元的状态基于当前输入和先前输出来维持,通过引入输入门控单元、输出门控单元和遗忘门控单

元来决定要遗弃的信息以及要传递的信息。输入门控单元用于控制是否允许外部信息覆盖当前单元中的状态,遗忘门控单元允许 LSTM 重置其自身状态,输出门控单元状态决定是否影响其他单元。

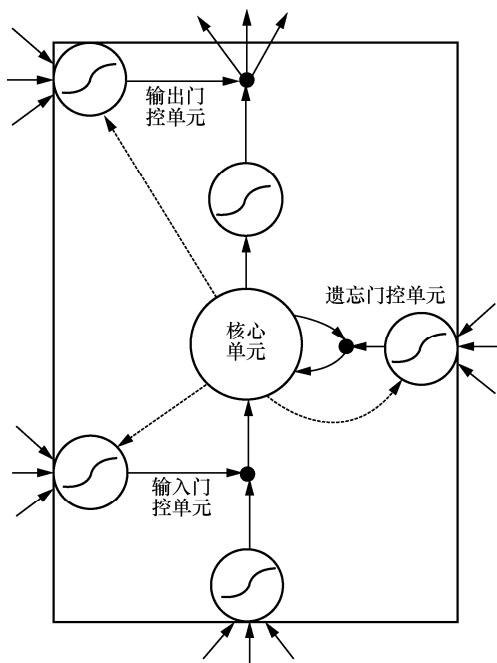


图 2 LSTM 的实现结构

LSTM 可以以多层方式工作,每层由多个单元组成,需要在模型的复杂度和计算效率之间进行折中。模型越复杂,其性能可能越好,然而,这可能会导致过度拟合模型,该模型在训练集中表现得非常好,但无法适应数据集的其他数据。同时,过于复杂的模型将是一个低效的模型。

目前没有明确的规则或标准程序来确定神经元或层的数量,算法的有效性取决于环境属性,如训练集的大小、数据特征的数量、计算性能。一般来说,层的数量与神经元的数量成反比,而大量的层可能会遇到梯度消失和过度拟合等问题。

在核心功能不变的前提下,为了简化 LSTM,引入了门控循环单元(gated recurrent unit, GRU)^[12],GRU 中重置单元 $r_{[t]}$ 、更新单元 $z_{[t]}$ 、激活单元 $k_{[t]}$



和遗忘单元 $h_{[t]}$ 的计算方式分别如式 (1) ~ 式 (4) 所示。

$$r_{[t]} = \sigma(W_{xr}x_{[t]} + b_{xr} + W_{hr}h_{[t-1]} + b_{hr}) \quad (1)$$

$$z_{[t]} = \sigma(W_{xz}x_{[t]} + b_{xz} + W_{hz}h_{[t-1]} + b_{hz}) \quad (2)$$

$$k_{[t]} = \tanh(W_{xk}x_{[t]} + b_{xk} + r_{[t]} \odot (W_{hk}h_{[t-1]} + b_{hk})) \quad (3)$$

$$h_{[t]} = (1 - z_{[t]}) \odot k_{[t]} + z_{[t]} \odot h_{[t-1]} \quad (4)$$

其中, σ 是 S 形激活函数, $\tanh$ 是双曲正切激活函数, $\odot$ 表示元素乘法。隐藏传递权重矩阵的 3 个输入是 W_{xr} 、 W_{xz} 和 W_{xk} , 而 b_{xr} 、 b_{xz} 和 b_{xk} 是相应的偏差。遗忘转移权重矩阵的 3 个输入和相应的偏差为 W_{hr} 、 W_{hz} 、 W_{hk} 、 b_{hr} 、 b_{hz} 、 b_{hk} 。

通过不断缩小 LSTM 输出和实际训练样本之间的差值, 迭代调整所有权重系数和偏差, 存储和维护当前时间步长的信息, 从而在可接受的学习时长内影响未来的 LSTM 模型输出。

1.4 电力负荷预测

通过建立模型结构, 并通过多次迭代对相应的权重矩阵和初始参数进行微调, 预测模型在其隐藏的存储单元中可以获得足够的信息, 从而能够预测具有可接受偏差的未知负荷。

2 模型的测试与验证

2.1 电力负荷预测

2.1.1 历史负荷数据

作为样本的历史电力负荷数据来自石家庄市供电公司 2021 年电力用户用电信息采集系统的原始文件, 这些数据对应于一个居住社区中的 314 个公寓用户, 用电数据每 15 min 采集一次, 总共有 10 000 余万条记录, 每条负荷记录有 3 个主要属性, 分别是时间戳、公寓/电表 ID 和有功功率 (单位为 kW)。

然而, 智能电表本地数据采集采用高速电力线载波通信方式, 在吞吐量和通信时延方面

的不足导致大量数据不完整或者丢失。历史采集数据的预处理规则是: 如果电力负荷值超出边界或者丢失, 则使用线性插值法, 调整负荷值为相邻读数值值的平均值。此外, 功耗数值以小时为分辨单位进行转换以适应每小时的天气数据, 这样处理可将负荷数据丢失率从 8.46% 降至 2.86%。

所用样本中 314 个公寓用户的个人日常生活和用电行为会有很大的不同, 因此将用户分割成不同的群体对于 STLF 至关重要。这里使用经典的 K 均值算法进行分类^[9]。给定一个参数 K , 该算法选取 K 个中心并按欧氏距离将每个点分布到最近的中心, 然后计算每个组的新中心, 并迭代该过程, 直到中心不再变化。最后, 通过迭代分类运算全年的负荷数据将消费者分为 5 个聚类, 分别为聚类 0、聚类 1、聚类 2、聚类 3 和聚类 4。5 个聚类的占比如图 3 所示, 其中聚类 1 的用户数量最多, 聚类 4 次之。

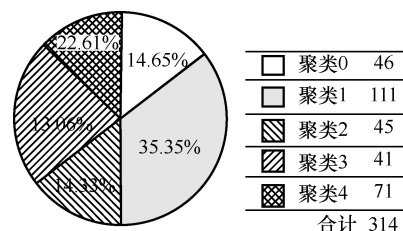


图 3 5 个聚类的占比

2.1.2 天气数据

石家庄市位于中国河北省, 该地区为暖温带半湿润大陆性季风气候, 四季分明, 年平均气温为 13.3℃; 冬夏季长, 春秋季短, 冬季盛行西北风, 天气晴朗, 气候寒冷, 月均气温可达 -2.9℃; 夏季炎热多雨, 月均气温可达 26.5℃。

获取的每小时天气数据可用于模拟分析, 如温度、体感温度气压、相对湿度、能见度、风速、风向、降水概率。为了区分相关性最大的气象因素, 使用单变量线性回归检验的 SelectKBest 和 scikit-learn Python 工具包中的 Lasso 正则化的

LassoCV 算法。天气数据与电力负荷的相关性见表 1, 最后选择了 3 个关键因素, 即温度、体感温度和气压。

表 1 天气数据与电力负荷的相关性

影响因素	SelectKBest	LassoCV
温度	2 213.146 5	0.831 2
体感温度	1 945.258 6	0.694 6
气压	314.621 4	0.152 3
相对湿度	129.316 7	0.028 7
降水概率	131.612 9	0.101 2
风速	2.644 2	-0.004 5
风向	1.571 7	0
能见度	0.482 4	0

2.1.3 日历数据

观察工作日(周一至周五)和周末(周六和周日)的用电情况差异, 梳理出工作日(2021 年 12 月 8 日)和周末(2021 年 12 月 12 日)5 个聚类的小时负荷总量, 工作日与周末的电力负荷对比如图 4 所示。与工作日相比, 周末用电负荷明显更高, 用电高峰持续时间长, 没有明显的中午和晚上两个相对用电高峰期。推测出现上述趋势的原因是大多数居民周末在家的时间更长, 周末家庭电器的工作时间也更长。

在每个公寓用户被分配到一个特定的聚类, 并

且聚类结果作为预测模型的输入, 最后将 6 类参数传递给预测模型。 X_1 表示用户集群, $X_1 \in \{1, 2, 3, 4, 5\}$; X_2 表示温度; X_3 表示体感温度; X_4 表示气压; X_5 表示日历日类型, $X_5 \in \{1, 2\}$; X_6 表示小时数。

2.2 构建 LSTM 预测模型

考虑计算性能, 本文的 LSTM 有 250 个神经元, 由于预测了 5 个聚类的负荷, 因此在输出层有 5 个神经元。在神经网络中, 激活函数是引导网络通过训练集拟合过程的关键。激活函数有两类: 线性函数和非线性函数, 后者被应用得更多。模型选用 LSTM 中传统的两种激活函数, 即 S 形函数和双曲线正切函数。此外, 机器学习中成本函数的作用是在训练阶段评估实际值和预测值之间的差距, 并使用反向传播方法修改神经元和神经链的权重。反向传播使用梯度下降法计算偏导数得出斜率, 并修改权重。最常见的代价函数是均方误差和交叉熵, 分别适用于回归问题和分类问题。本文研究的负荷预测本质上是一个回归问题, 因此采用均方误差作为模型的代价函数。

学习率参数对于机器学习过程有非常重要的意义, 学习率和梯度下降之间有很强的相关性。每次执行梯度下降时, 学习速率可以确定步长。大的学习率没有适合的、满意的解决方案, 而小的学习率会导致很长的训练时间。

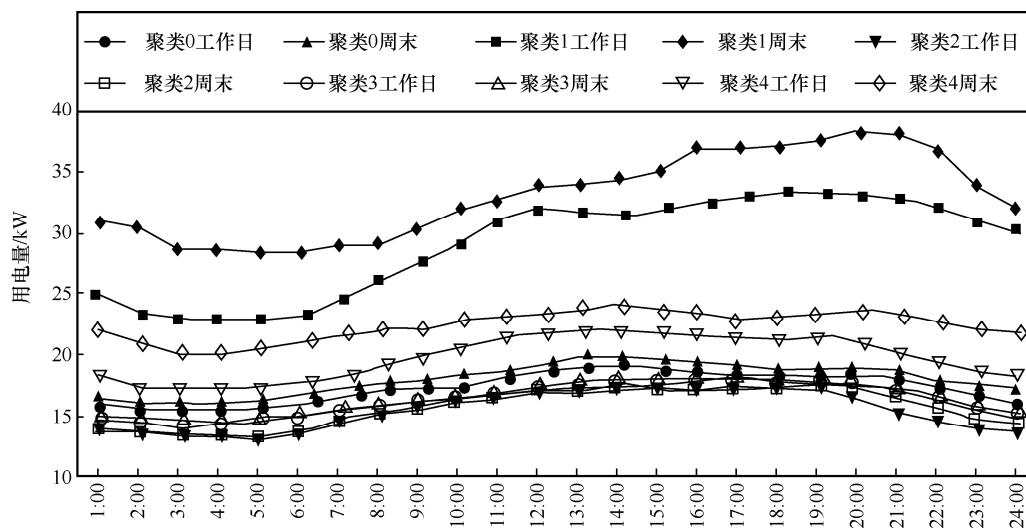


图 4 工作日与周末的电力负荷对比



在本文所提模型中,学习率最初设置为传统缺省值 0.1,并通过每轮减少 10%进行迭代以获得更好的错误值。通过实验,模型中的学习率选择设置为 0.001。另外,过拟合问题在机器学习过程中比较突出,随着训练集的误差越来越小,验证集的误差则越来越大。常用的训练方法有早期停止和中途放弃两种,本文模型选用直接的早期停止法,即在

训练集和验证集之间存在明显差异之前停止训练过程,以提高训练效率。

2.3 模型训练

使用本文模型、多元线性回归和 BP 神经网络 3 种模型预测目标社区的短期负荷。首先针对每一个聚类进行预测计算,得出每个聚类的小时负荷数值,将 24 个预测小时值相加得到一日的负荷用电

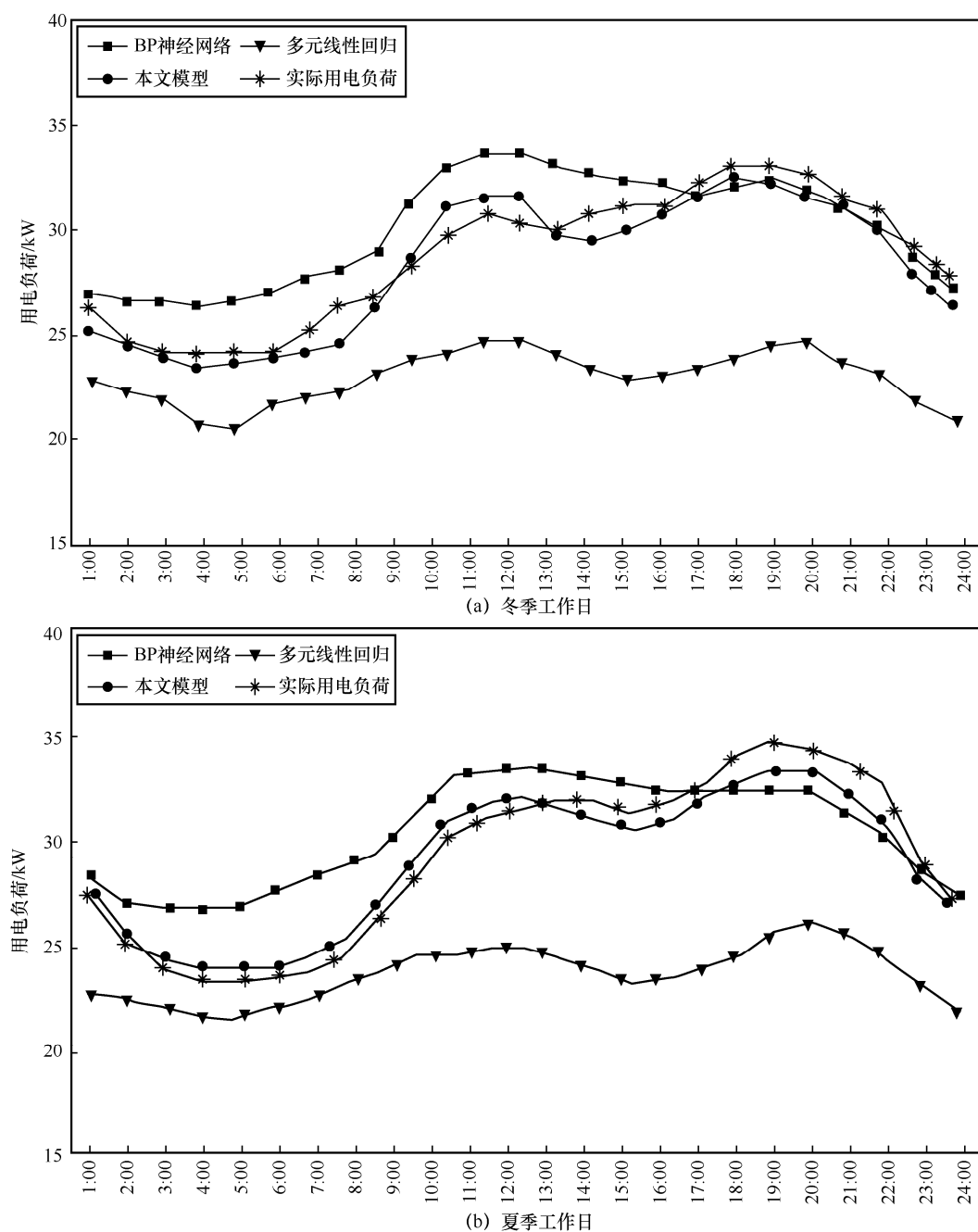


图5 聚类1的工作日负荷预测结果

量。同样,将5个聚类的预测值相加,得到整个社区的总体负荷用电量。

聚类1的工作日负荷预测结果如图5所示,描述了用户数量占比最大的聚类1在冬季和夏季工作日(2021年12月16日、2021年7月6日)的小时预测值和实际值。聚类1的周末负荷预测结果如图6所示,描述了聚类1在周末(2021年12月19日、2021年7月10日)的小时预测值和实际值。从图5

和图6可以看出,该社区居民在夏季和冬季的日用电量差异较小,预测结果与实际负荷值比较接近,而且本文模型的预测结果偏差略小于其他两种模型。

使用均方根误差(root mean square error, RMSE)和平均绝对百分比误差(mean absolute percentage error, MAPE)作为模型性能评估指标。RMSE和MAPE的定义分别如式(5)和式(6)所示。其中, $\bar{Y}$ 表示实际负荷数值, Y 表示预测

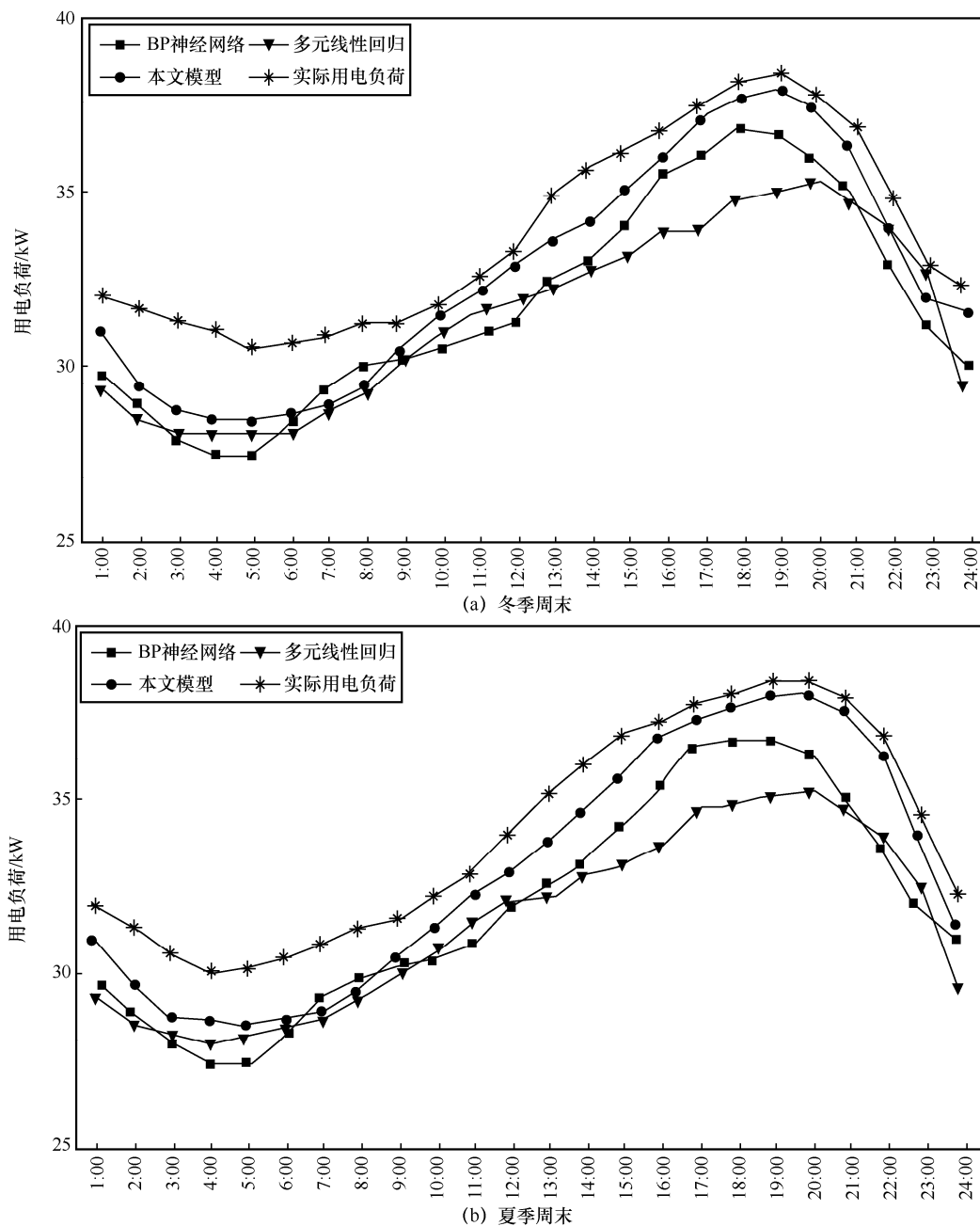


图6 聚类1的周末负荷预测结果



负荷值, Q 表示预测计算的聚类总数。

$$RMSE = \sqrt{\frac{1}{Q} \sum_{q=1}^Q (\bar{Y}_q - Y_q)^2} \quad (5)$$

$$MAPE = \frac{1}{Q} \times \sum_{q=1}^Q \frac{|\bar{Y}_q - Y_q|}{\bar{Y}_q} \times 100\% \quad (6)$$

采用 3 种模型进行对 2021 年 12 月 16 日的负荷进行预测, 3 种模型的 RMSE 对比见表 2。本文模型的 RMSE 最低, 领先 0.29~2.22 kW。此外, 表 2 表明, 聚类 1 的 RMSE 最低, 仅为 1.31 kW; 聚类 4 的 RMSE 指标次之, 为 1.42 kW, 这两个聚类的用户数量占样本用户数量近 60%。

表 2 3 种模型的 RMSE 对比/kW

模型	聚类 0	聚类 1	聚类 2	聚类 3	聚类 4
本文模型	2.31	1.31	2.09	2.31	1.42
多元线性回归	3.12	3.53	2.84	3.08	2.51
BP 神经网络	2.96	2.86	2.38	3.15	3.24

采用 3 种模型对 2021 年 12 月 16 日进行负荷预测, 3 种模型的 MAPE 对比见表 3。与多元线性回归和 BP 神经网络相比, 本文模型的 MAPE 指标明显领先。此外, 表 3 表明, 本文模型中 聚类 1 的 MAPE 指标最高, 仅为 7.28%, 聚类 2 次之, 为 7.32%, 尚可接受。

表 3 3 种模型的 MAPE 对比

模型	聚类 0	聚类 1	聚类 2	聚类 3	聚类 4
本文模型	8.14%	7.28%	7.32%	7.98%	8.06%
多元线性回归	23.83%	24.56%	23.76%	19.26%	20.44%
BP 神经网络	14.32%	13.38%	14.27%	13.48%	14.34%

为检验本文模型对不同年份负荷预测的适用性, 选用石家庄市同一个社区 2018 年的 5 500 多万条历史负荷记录进行大数据训练和分析, 用户群被分为 5 个簇, 就一个春季工作日(2018 年 5 月 21 日)的负荷进行预测和偏差比对, 3 种模型的 RMSE 和 MAPE 对比分别见表 4 和表 5。从表 4 和表 5 可以看出, 本文模型的 RMSE 和 MAPE 两个指标均优于其他两种模型。

表 4 3 种模型的 RMSE 对比/kW

模型	聚类 0	聚类 1	聚类 2	聚类 3	聚类 4
本文模型	2.24	1.28	2.18	2.42	1.46
多元线性回归	3.64	4.21	3.78	3.97	2.89
BP 神经网络	3.49	3.25	2.98	3.65	4.03

表 5 3 种模型的 MAPE 对比

模型	聚类 0	聚类 1	聚类 2	聚类 3	聚类 4
本文模型	8.34%	7.42%	7.61%	8.32%	8.12%
多元线性回归	35.13%	43.74%	36.46%	37.13%	31.93%
BP 神经网络	24.52%	35.83%	26.18%	27.12%	30.82%

3 结束语

本文研究了居民用户短期负荷预测问题, 提出了一种在大数据框架下使用 LSTM 算法的短期负荷预测模型, 主要通过处理分析以不同时间分辨率记录的大量用户电力负荷数据, 动态确定最重要的属性为预测因素, 并考虑它们对未来电力负荷的影响, 同时, 通过对非相关属性进行噪声过滤, 减少了 LSTM 算法的配置, 提高了性能。与传统直接预测某个社区总负荷的策略相比, 通过考虑消费模式的相似性将消费者分组到相关的集群中, STLF 的性能得到了提高。通过 RMSE 和 MAPE 两个指标, 将本文模型与传统线性回归和 BP 神经网络进行对比分析, 从中可以看出, 本文模型具有良好的预测性能。本文的工作可以为电力部门进行新能源接入背景下的配电控制和电价动态调整提供决策依据。下一步研究将在更广的不同区域维度上验证预测方案的性能和应用模式, 提高适用性, 推动该电力短期负荷预测方案的实用化。

参考文献:

- [1] 张东霞, 王继业, 刘科研, 等. 大数据技术在配用电系统的应用[J]. 供用电, 2015, 32(8): 6-11.
ZHANG D X, WANG J Y, LIU K Y, et al. Application of big data technologies in power distribution and utilization system[J]. Distribution & Utilization, 2015, 32(8): 6-11.
- [2] 赵明欣, 张春雷, 宋曼诗, 等. 基于居民小区负荷成长周期的负荷预测方法[J]. 供用电, 2016, 33(8): 64-68.

- ZHAO M X, ZHANG C L, SONG M S, et al. A load forecasting method based on load growth cycle of residential community[J]. Distribution & Utilization, 2016, 33(8): 64-68.
- [3] JIANG H, WANG K, WANG Y H, et al. Energy big data: a survey[J]. IEEE Access, 2016(4): 3844-3861.
- [4] DAKI H, EL HANNANI A, AQQAL A, et al. Big data management in smart grid: concepts, requirements and implementation[J]. Journal of Big Data, 2017(4): 13.
- [5] HONG T, GUI M, BARAN M E, et al. Modeling and forecasting hourly electric load by multiple linear regression with interactions[C]//Proceedings of IEEE PES General Meeting. Piscataway: IEEE Press, 2010.
- [6] 王克杰, 张瑞. 基于改进 BP 神经网络的短期电力负荷预测方法研究[J]. 电测与仪表, 2019, 56(24): 70-75.
- WANG K J, ZHANG R. Research on short-term power load forecasting method based on improved BP neural network[J]. Electrical Measurement & Instrumentation, 2019, 56(24): 70-75.
- [7] OPREA S V, BARA A. Machine learning algorithms for short-term load forecast in residential buildings using smart meters, sensors and big data solutions[J]. IEEE Access, 2019(7): 177874-177889.
- [8] 王德文, 周昉昉. 基于无监督极限学习机的用电负荷模式提取[J]. 电网技术, 2018, 42(10): 3393-3400.
- WANG D W, ZHOU F F. Extraction of electricity consumption load pattern based on unsupervised extreme learning machine[J]. Power System Technology, 2018, 42(10): 3393-3400.
- [9] ZHOU H X, ZHANG Y J, YANG L F, et al. Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism[J]. IEEE Access, 2019(7): 78063-78074.
- [10] MOR G, VILAPLANA J, DANOV S, et al. EMPOWERING, a smart big data framework for sustainable electricity suppliers[J]. IEEE Access, 2018(6): 71132-71142.

- [11] MOGHADDASS R, WANG J H. A hierarchical framework for smart grid anomaly detection using large-scale smart meter data[J]. IEEE Transactions on Smart Grid, 2018, 9(6): 5820-5830.
- [12] ZHAO T, ZHOU Z Q, ZHANG Y, et al. Spatio-temporal analysis and forecasting of distributed PV systems diffusion: a case study of Shanghai using a data-driven approach[J]. IEEE Access, 2017(5): 5135-5148.

[作者简介]



赵海波 (1978-), 男, 博士, 威胜信息技术股份有限公司高级工程师, 主要研究方向为电力大数据分析、移动通信、电力终端设计。



相志军 (1984-), 男, 中国电力技术装备有限公司高级工程师, 主要研究方向为 AMI 系统数据采集。



肖林松 (1980-), 男, 威胜信息技术股份有限公司高级工程师, 主要研究方向为电力终端设备开发、电力大数据分析。