

基于注意力机制的 CNN-GRU 短期电力负荷预测方法

赵兵^{1,2}, 王增平¹, 纪维佳³, 高欣³, 李晓兵³

(1. 华北电力大学 电气与电子工程学院, 北京市 昌平区 102206;

2. 中国电力科学研究院有限公司, 北京市 海淀区 100192;

3. 北京邮电大学 自动化学院, 北京市 海淀区 100876)

A Short-term Power Load Forecasting Method Based on Attention Mechanism of CNN-GRU

ZHAO Bing^{1,2}, WANG Zengping¹, JI Weijia³, GAO Xin³, LI Xiaobing³

(1. School of Electrical and Electronic Engineering, North China Electric Power University, Changping District, Beijing 102206, China;

2. China Electric Power Research Institute, Haidian District, Beijing 100192, China;

3. School of Automation, Beijing University of Posts and Telecommunications, Haidian District, Beijing 100876, China)

ABSTRACT: Efficient and accurate short-term load forecasting can help power utilities to rationally formulate production scheduling plans and reduce resource waste. RNN (recurrent neural network)-based predictive model in deep learning is a typical method of short-term load forecasting, but it is difficult to effectively extract potential high-dimensional features in historical sequences and important information is easily lost when the time series is too long, resulting in decrease of prediction accuracy. This paper proposes a short-term load forecasting method of CNN (convolutional neural network)-GRU (gated recurrent unit) based on attention mechanism. This model takes historical load data as input and builds a CNN architecture composed of one-dimensional convolution layers and pooling layers to extract complex dynamic changes of load. A high-dimensional feature is constructed by constructing the proposed feature vector into time series as GRU input. It models the internal dynamic change of the feature, and the attention mechanism is introduced to the GRU implicit state through mapping weight and learning parameter matrix, reducing the loss of historical information and enhancing the impact of important information, thus completing short-term load forecasting. Taking the public data set provided by a public utility department in the United States and the load data of a certain region in China as practical examples, the prediction accuracy of the method reaches 97.15% and 97.44% respectively. Comparing with multi-layer perceptron (MLP), radial basis function (RBF) neural network, support vector regression (SVR), GRU, CNN, AE (autoencoder)-GRU and CNN-GRU, the proposed method has higher prediction accuracy

showed in experimental results.

KEY WORDS: short-term load forecasting; convolutional neural network; gated recurrent unit; attention mechanism

摘要: 高效准确的短期电力负荷预测能帮助电力部门合理制定生产调度计划, 减少资源浪费。深度学习中以循环神经网络(recurrent neural network, RNN)为主体构建的预测模型是短期负荷预测方法中的典型代表, 但存在难以有效提取历史序列中潜在高维特征且当时序过长时重要信息易丢失的问题。提出了一种基于 Attention 机制的卷积神经网络(convolutional neural network, CNN)-GRU (gated recurrent unit)短期电力负荷预测方法, 该方法将历史负荷数据作为输入, 搭建由一维卷积层和池化层等组成的 CNN 架构提取反映负荷复杂动态变化的高维特征; 将所提特征向量构造为时间序列形式作为 GRU 网络的输入, 建模学习特征内部动态变化规律, 并引入 Attention 机制通过映射加权和学习参数矩阵赋予 GRU 隐含状态不同的权重, 减少历史信息的丢失并加强重要信息的影响, 最后完成短期负荷预测。以美国某公共事业部门提供的公开数据集和中国西北某地区的负荷数据作为实际算例, 该方法预测精度分别达到了 97.15% 和 97.44%, 并与多层感知机(multi-layer perceptron, MLP)、径向基神经网络(radial basis function neural network, RBF)、支持向量回归(support vector regression, SVR)、GRU、CNN、自编码器(autoencoder, AE)-GRU 和未引入 Attention 机制的 CNN-GRU 进行对比, 实验结果表明所提方法具有更高的预测精度。

关键词: 短期负荷预测; 卷积神经网络; 门控循环单元; 注意力机制

DOI: 10.13335/j.1000-3673.pst.2019.1524

0 引言

电力负荷预测在电力系统的调度运行中起着重要作用, 根据预测时间可分为长期、中期和短期。

基金项目: 国家重点研发计划项目((2016YFF0201201)。

Project Supported by National Key R&D Program of China(2016YFF0201201)。

准确的短期负荷预测不仅能帮助系统安全可靠地运行,还能减少资源浪费,提高经济效益^[1]。

目前用于负荷预测的方法可分为统计方法和机器学习方法。统计方法主要包括自回归积分移动平均值^[2]、指数平滑法^[3]和卡尔曼滤波^[4]等。这些方法简单易行,但对原始数据的处理和时间序列的稳定性要求很高,难以反映非线性因素的影响,适用于影响因素较少的情况。机器学习方法包括专家系统(expert systems, ES)^[5]、支持向量机^[6]和人工神经网络(artificial neural network, ANN)^[7]等。这些方法虽都能较好地处理非线性问题,但专家系统不具备自学习的能力,支持向量机处理大规模数据困难。人工神经网络具有自学习、自适应能力,能充分解决大规模负荷数据中存在的非线性问题,因此在预测领域应用广泛。其典型方法包括多层感知机(MLP)^[8]、小波神经网络(wavelet neural network, WNN)^[9]和广义神经网络(generalized neural network, GNN)^[10]等。但研究者发现 ANN 难以学习序列数据的相关性,通常利用历史负荷数据人为提取时间特征,建立输入与输出变量之间的关系。电力负荷序列具有非线性、不平稳和动态变化等特点,即模型的输出不仅与当前输入有关,与过去的输入也有关^[11],人为提取特征会影响负荷数据的连续性,单一建立输入输出模型关系也会降低预测精度。

近年来,深度学习的高速发展和广泛应用使其成为负荷预测领域的热点。典型方法有深度神经网络(deep neural network, DNN)、深度置信网络(deep belief network, DBN)、自编码器(AE)、卷积神经网络(convolutional neural network, CNN)和循环神经网络(recurrent neural network, RNN)等。文献[12]和文献[13]分别将 DBM 和 DNN 应用到短期负荷预测问题中,较浅层人工神经网络在精度上虽有所提升;但仍存在人为提取时序特征的问题,忽略了负荷数据时序性的特点。AE 属于无监督学习,主要用于数据降维或特征提取,通常需与其他模型结合后应用到负荷预测领域^[14]。CNN 是包含卷积计算且具有深度结构的神经网络,使用卷积池化对数据特征进行提取,降低了人为提取特征带来的误差,在图像、语音等领域应用广泛^[15]。文献[16]和文献[17]通过利用 CNN 模型提取输入数据的特征并捕获季节性周期来进行负荷预测,发现在学习负荷数据这类高度非线性序列上准确率更高;但当负荷数据波动性和不稳定性较大时,单一的 CNN 模型难以较好地学习负荷的动态变化。RNN 在网络中引入了循环结构,使其相较于其他神经网络能更

好地对动态时间序列数据进行建模^[18]。GRU (gated recurrent unit)是一种特殊的 RNN,与长短期记忆网络(long short-term memory, LSTM)^[19-21]都为解决 RNN 存在的梯度消失问题提出的,它能更完整地考虑时间序列中长短期依赖性关系;相较于 LSTM 具有更快的收敛速度,并在精度上差距不大,因此也逐渐应用到短期负荷预测领域中^[22]。然而,当输入的时间序列较长时,LSTM、GRU 等 RNN 系列网络易出现丢失序列信息、难以建模数据间结构信息的问题,影响模型的准确率^[23]。Attention 机制是一种资源分配机制,可以通过对输入特征赋予不同的权重,使包含重要信息的特征不会随着步长的增加而消失,突出更加重要信息的影响,使模型更易学习序列中长距离相互依赖的关系^[24]。

为提高短期负荷预测的精度,本文提出了一种基于 Attention 机制的 CNN-GRU 的短期负荷预测模型。该方法利用 CNN 从历史负荷序列中提取有效的特征向量,采用 GRU 网络对 CNN 所提特征的动态变化进行建模学习,并引入 Attention 机制赋予 GRU 隐含状态不同的概率权重,加强重要信息对负荷需求的影响。此模型旨在利用不同结构更有效的分析处理电力负荷数据,以更准确的预测负荷值的需求变化,最后通过实验证明,与典型方法对比本文所提方法具有更高的预测精度。

1 深度学习模型原理

1.1 CNN 原理结构

CNN 模型采用局部连接和权重共享的方式,对原始数据进行更高层次和更抽象的处理,能够有效自动提取数据中的内部特征^[25]。其内部神经网络层主要由卷积层、池化层和全连接层等组成,此结构减少了权重的数量,降低了网络模型的复杂度。模型通过使用卷积层和池化层获取有效信息,自动提取数据中的特征向量,有效降低了特征提取和数据重构的复杂度,提高了数据特征质量。结构见图 1。

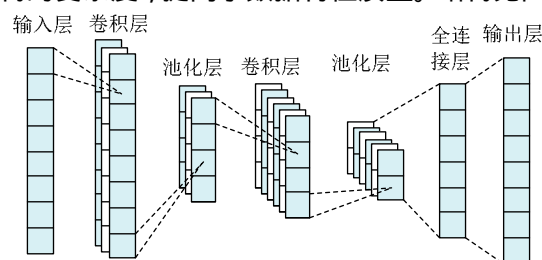


图 1 CNN 结构
Fig. 1 Structure of CNN

1.2 GRU 原理结构

LSTM 能够捕获长期依赖关系,适用于分析时间序列数据,但内部结构复杂导致训练时间较长,

GRU 对 LSTM 进行优化改进,减少了训练参数的同时保证了预测精度^[26]。GRU 包含更新门和重置门,相比 LSTM 三个门结构参数更少,收敛速度更快。更新门控制前一时刻的状态信息保留到当前状态中的程度,值越大表示前一时刻的状态信息保留越多。重置门控制当前状态与先前的信息结合的程度,值越小说明忽略的信息越多^[27]。GRU 结构图见图 2。在下图中,箭头所指方向为数据流动方向,其中, \times 为矩阵的数乘, σ 为激活函数 Sigmoid 函数, \tanh 为激活函数 \tanh ,1-表示该链路向前传播的数据为 $1-z_t$ 。

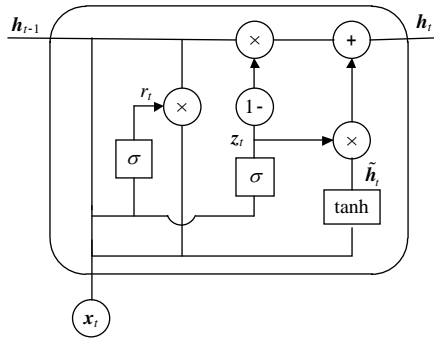


图 2 GRU 结构

Fig. 2 Structure of GRU

如前所述,更新门和重置门分别为 z_t 和 r_t , x_t 为输入, h_t 为隐藏层的输出。基于 GRU 的单元通过以下公式计算 h_t

$$z_t = \sigma(W^{(z)}x_t + U^{(z)}h_{t-1}) \quad (1)$$

$$r_t = \sigma(W^{(r)}x_t + U^{(r)}h_{t-1}) \quad (2)$$

$$\tilde{h}_t = \tanh(r_t \circ U h_{t-1} + W x_t) \quad (3)$$

$$h_t = (1 - z_t) \circ \tilde{h}_t + z_t \circ h_{t-1} \quad (4)$$

式中: \tilde{h}_t 是输入 x_t 和过去隐层状态 h_{t-1} 的汇总; $U^{(z)}$ 、 $W^{(z)}$ 、 $U^{(r)}$ 、 $W^{(r)}$ 、 U 和 W 为可训练参数矩阵。

1.3 Attention 机制原理

Attention 机制是一种模拟了人脑注意力的资源分配机制,人脑在某个特定时刻会将注意力集中在需要重点关注的区域,减少甚至忽略对其他区域的关注,以获取更多所需要关注的细节信息,抑制其他无用信息,其核心思想是巧妙合理的改变对信息的注意力,忽略无关信息并放大所需信息^[24]。Attention 通过概率分配的方式,对关键信息分配足够的关注,突出重要信息的影响,从而提高模型的准确率。Attention 结构如图 3 所示。其中 $x_t(t \in [1, n])$ 表示 GRU 网络的输入 $h_t(t \in [1, n])$ 对应于每一个输入通过 GRU 得到的隐藏层输出, $\alpha_t(t \in [1, n])$ 为 attention 机制对 GRU 隐藏层输出的注意力概率分布值, y 为引入 attention 机制的

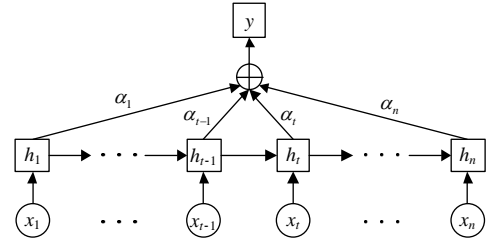


图 3 Attention 机制结构

Fig. 3 Structure of attention mechanism

GRU 输出值。

2 基于 Attention 机制的 CNN-GRU 模型

2.1 预测模型叙述

在短期电力负荷预测中,历史序列是负荷需求变化的表现,它包含了负荷预测规律的重要信息。传统机器学习方法、DBM 和 DNN 等在处理历史负荷数据时,通常是人为从历史负荷数据中选取时间特征的,如选取前一个小时的电量值、前一天同一时刻的电量值等作为特征,但这种特征提取方法会破坏历史负荷序列的时序性和蕴含的潜在规律。CNN 凭借其特殊结构能充分挖掘数据之间的相互关系,提取出较为重要的特征,捕获历史负荷数据中的周期性,为提高预测精度打下了基础。GRU 网络可以在负荷数据高波动性和不确定性情况下,有效的对动态时间序列数据建模,将 CNN 提取出的特征输入到 GRU 中能更好的学习负荷数据中的周期性变化和负荷需求规律。但是当短期负荷预测中输入时间序列过长时,GRU 网络易出现信息丢失、难以建模数据间结构信息的问题。Attention 机制通过对模型输入特征赋予不同的权重,加强了重要信息的影响以避免序列过长信息丢失问题,使模型更容易捕获序列中长距离相互依赖的特征。因此,本文提出基于 Attention 机制的 CNN-GRU 短期电力负荷预测模型,通过结合多种结构能有效学习输入数据中丰富的特性和规律,提高了短期负荷预测的精度。

2.2 预测模型结构

本文提出的模型结构如图 4 所示,主要分为输入层、CNN 层、GRU 层、Attention 层和输出层。历史负荷数据作为输入通过 CNN 层提取特征,设计卷积操作使其深度增加,为压缩参数数量,再经过池化处理进行特征降维,最后由全连接层将特征转换为一维结构,完成特征向量的提取。GRU 层与 Attention 层从所提特征中学习负荷内部的变化规律以实现预测功能,最后通过输出层得到预测结果。

模型中每层描述如下:

1) 输入层。输入层将历史电力负荷数据作为

预测模型的输入。长度为 n 的负荷数据经过预处理后输入到预测模型中,可用 $X=[x_1 \cdots x_{t-1}, x_t \cdots x_n]^T$ 表示。

2) CNN 层。CNN 层主要对输入的历史序列进行特征提取。构建由 2 个卷积层、2 个池化层和全连接层组成的 CNN 框架。根据负荷数据的特点,将卷积层 1 和卷积层 2 设计为一维卷积,并选取 ReLU 激活函数进行激活。为保留更多数据波动信息,池化层 1 和池化层 2 的池化方法选取最大池化。经过卷积层和池化层的处理后原始数据被映射到隐层特征空间,搭建全连接层结构将其转换输出,提取得到特征向量,全连接层选用激活函数 Sigmoid。CNN 层的输出特征向量 H_c ,可表示为

$$C_1 = f(X \otimes W_1 + b_1) = \text{ReLU}(X \otimes W_1 + b_1) \quad (5)$$

$$P_1 = \max(C_1) + b_2 \quad (6)$$

$$C_2 = f(P_1 \otimes W_2 + b_3) = \text{ReLU}(P_1 \otimes W_2 + b_3) \quad (7)$$

$$P_2 = \max(C_2) + b_4 \quad (8)$$

$$H_c = f(P_2 \times W_3 + b_5) = \text{Sigmoid}(P_2 \times W_3 + b_5) \quad (9)$$

式中: C_1 和 C_2 分别为卷积层 1 和卷积层 2 的输出;

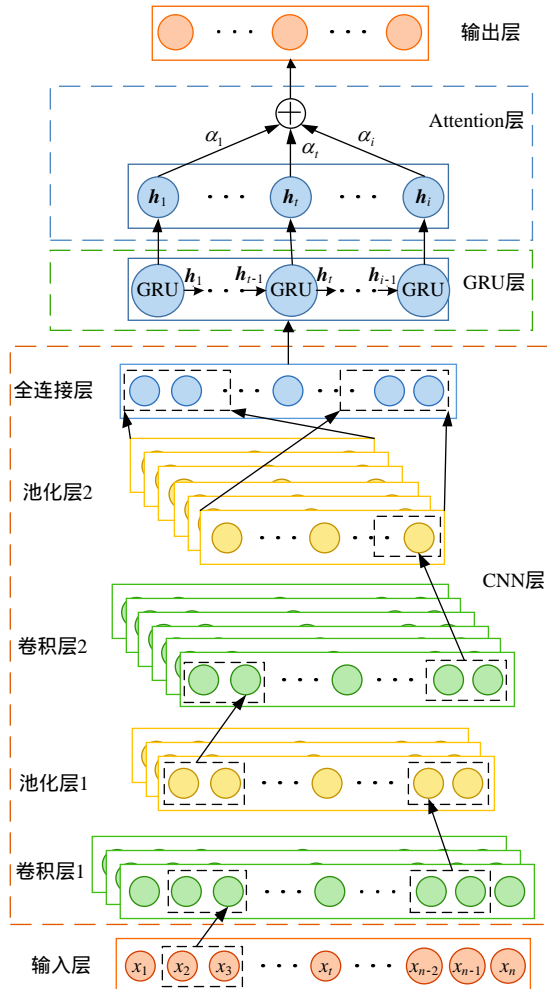


图4 基于 Attention 机制的 CNN-GRU 模型结构
Fig.4 Structure of CNN-GRU model based on attention mechanism

P_1 和 P_2 分别池化层 1 和池化层 2 的输出; W_1 、 W_2 和 W_3 为权重矩阵; b_1 、 b_2 、 b_3 、 b_4 和 b_5 为偏差; \otimes 和 $\max()$ 为卷积运算和最大值函数; CNN 层的输出长度为 i , 记为 $H_c=[h_{c1} \cdots h_{ct-1} \cdots h_{ct} \cdots h_{ci}]^T$ 。

3) GRU 层。GRU 层对 CNN 层提取的特征向量进行学习。搭建单层 GRU 结构,对所提特征进行充分学习,以捕获其内部的变化规律。GRU 层的输出记为 H , 在第 t 步的输出表示为

$$h_t = \text{GRU}(H_{c,t-1}, H_{c,t}), t \in [1, i] \quad (10)$$

4) Attention 层。Attention 机制层的输入为经过 GRU 网络层激活处理的输出向量 H , 根据权重分配原则计算不同特征向量对应的概率,不断更新迭代出较优的权重参数矩阵。Attention 机制层的权重系数计算公式可表示为

$$e_t = \text{utanh}(wh_t + b) \quad (11)$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^t e_j} \quad (12)$$

$$s_t = \sum_{i=1}^i \alpha_i h_i \quad (13)$$

式中: e_t 表示第 t 时刻由 GRU 网络层输出向量 h_t 所决定的注意力概率分布值; u 和 w 为权重系数; b 为偏置系数; Attention 层在第 t 时刻的输出由 s_t 表示。

5) 输出层。输出层的输入为 Attention 机制层的输出。输出层通过全连接层进行计算出预测步长为 m 的输出 $Y=[y_1, y_2, \cdots, y_m]^T$, 预测公式可表示为

$$y_t = \text{Sigmoid}(w_o s_t + b_o) \quad (14)$$

式中: y_t 表示第 t 时刻的预测输出值; w_o 为权重矩阵; b_o 为偏差向量, 本文 Dense 层的激活函数选取 Sigmoid 函数。

2.3 损失函数

本文在预测模型的训练过程中,选取 Adam (adaptive moment estimation)^[28] 优化算法对模型参数进行优化。Adam 是一种可以替代传统随机梯度下降过程的一阶优化算法,该算法能够基于训练数据迭代更新神经网络的权重,使损失函数输出值达到最优。模型的损失函数使用均方误差函数,即

$$L_{\text{oss}} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y}_i)^2 \quad (15)$$

式中: n 为样本个数; y_i 实际值; \bar{y}_i 为模型输出值。结合本文研究内容以日负荷 48 点预测为例, n 为负荷预测输出时刻的点数即 48, y_i 为 i 时刻的负荷实际值, \bar{y}_i 为模型运算后得到的 i 时刻负荷预测值。

3 算例分析

3.1 数据预处理及误差指标

本文采用 2 个实际负荷数据集对模型进行实验，分别为美国某公共事业部门提供的公开数据集和中国西北某地区的实际电力负荷数据集。美国公开数据集选取了 2006 年 1 月 1 日—2007 年 12 月 31 日共 2 年电力负荷数据，一天采集 24 点，时间间隔为 1h。中国某地区数据集提供了 2016 年 1 月 1 日—2017 年 8 月 31 日共 20 个月的电力负荷数据，一天采集 48 点，时间间隔为 30min。本文将 AE 和 GRU 结合的模型(记为 AE-GRU)、MLP、RBF、SVR、GRU、CNN 和未引入 Attention 机制的 CNN-GRU 模型作为对比方法。根据负荷数据特点，所有方法的输入都为相应的历史负荷序列数据，目的是公平有效的验证所提方法的科学性和可靠性。

为了便于模型网络的训练，采用 min-max 归一化法将原始数据归一化在(-1, 1)之间，计算公式为

$$x_n = \frac{x - x_{\text{mean}}}{x_{\text{max}} - x_{\text{min}}} \quad (16)$$

式中： x 为原始负荷数据； x_n 是归一化处理后的数据； x_{mean} 、 x_{max} 和 x_{min} 分别是样本数据的平均值、最大值和最小值。

为评估模型预测的精度，本文选取平均绝对百分误差 M_{APE} 和均方根误差 R_{MSE} 作为评价标准，其表达式分别为

$$M_{\text{APE}} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \bar{y}_i|}{y_i} \times 100\% \quad (17)$$

$$R_{\text{MSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (18)$$

式中： n 表示为预测结果的总个数； y_i 和 \bar{y}_i 分别为预测第 i 个采样点的实际负荷值和预测负荷值。MAPE 可以衡量模型预测结果的好坏，RMSE 可以评价预测的精度，对结果中极大或极小的误差反映敏感。电力负荷预测中，MAPE 和 RMSE 的值越小，负荷预测结果越准确。

3.2 美国公开数据集预测结果对比

本文对美国公开数据集提供的某一地区电力负荷进行训练与预测，使用此地区前 18 个月数据作为训练样本集，对后 6 个月的日负荷进行预测。

为验证模型的科学性和稳定性，对测试集中每月连续一周进行日负荷预测，各模型定量评价结果如表 1 所示。从预测结果平均值角度分析，本文方法预测精度相较于其他 7 种方法是最高的，平均预测精度达到了 97.15%，其中 MAPE 平均值相比其他 7 种方法分别降低了 3.18%、2.79%、3.02%、2.30%、2.19%、1.8%、0.42%；RMSE 相比于其他 7 种方法分别降低了 58.47%、55.4%、64.4%、49.01%、49.51%、42.78%、19.53%。综合分析，本文方法在 MAPE 和 RMSE 指标上都有明显的下降，表明在预测过程中整体预测精度和模型性能都有较大的提升。

图 5 为 2006 年 11 月 7—9 日连续三天的真实值和不同方法的短期负荷预测曲线对比图，由图 5 可知本文所提预测模型能更好贴近真实值，具有更高的预测精度。相较于其他模型，本文方法除了能预测出负荷的变化规律，在负荷变化较大的峰谷时段表现更为突出，能较为准确地分析峰谷值时段的负荷变化规律，提高了预测精度。

表 1 不同月份预测精度比较
Tab. 1 Comparison of forecasting accuracy in different months

日期	评价标准	结果							
		MLP	RBF	SVR	GRU	CNN	AE-GRU	CNN-GRU	本文方法
7 月 11 日—7 月 17 日	$M_{\text{APE}}/\%$	5.91	5.62	5.87	5.88	5.63	4.92	3.83	3.16
	R_{MSE}/kW	2.90	2.68	2.48	2.59	2.77	2.20	1.77	1.26
8 月 8 日—8 月 14 日	$M_{\text{APE}}/\%$	6.70	5.26	6.59	5.54	4.19	4.15	3.49	3.07
	R_{MSE}/kW	2.92	2.61	3.07	2.31	1.71	1.84	1.46	1.16
9 月 4 日—9 月 10 日	$M_{\text{APE}}/\%$	5.86	5.10	5.23	4.52	5.52	4.36	3.30	2.87
	R_{MSE}/kW	2.46	2.13	2.29	1.94	2.18	1.69	1.30	1.01
10 月 6 日—10 月 12 日	$M_{\text{APE}}/\%$	6.42	5.94	5.63	5.00	4.56	4.46	3.39	2.94
	R_{MSE}/kW	2.22	1.94	1.98	1.50	1.42	1.37	1.06	0.89
11 月 7 日—11 月 13 日	$M_{\text{APE}}/\%$	6.04	5.68	5.84	5.23	5.15	4.96	2.82	2.69
	R_{MSE}/kW	2.22	1.54	1.70	1.75	1.91	1.63	0.94	0.86
12 月 9 日—12 月 15 日	$M_{\text{APE}}/\%$	5.26	6.21	4.90	4.74	5.16	5.06	2.77	2.38
	R_{MSE}/kW	2.17	2.94	2.48	2.05	2.25	2.09	1.15	0.99
平均值	$M_{\text{APE}}/\%$	6.03	5.64	5.87	5.15	5.04	4.65	3.27	2.85
	R_{MSE}/kW	2.48	2.31	2.48	2.02	2.04	1.80	1.28	1.03

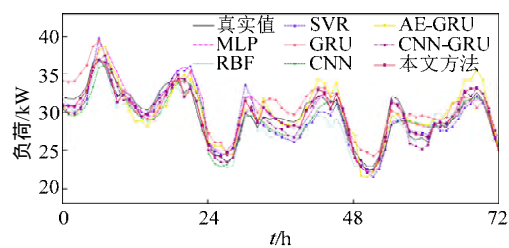


图 5 2006 年 11 月 7—9 日预测结果对比
Fig. 5 Comparison of forecasting results from 7 to 9 November 2006

3.3 中国西北某地区数据集预测结果对比

本文将中国西北某地区的电力负荷数据划分为 2 个数据集，训练集为前 16 个月，后 4 个月作

为测试集，对测试集的日负荷进行预测。

表 2 为不同模型在 2017 年 5 月连续一周的日负荷预测结果，可知本文提出模型的日负荷预测精度更高，其平均值达到了 97.44%，MAPE 平均值相比其他 7 种方法分别降低了 1.42%、1.31%、1.14%、1.57%、1.42%、1.39%、0.60%；RMSE 分别降低了 35.79%、37.5%、36.9%、39.46%、38.99%、39.19%、20.48%。本文方法相对于其他 7 种方法在 2 项精度指标上都有明显提升，预测精度更高。从日负荷精度角度分析，可看出本文模型预测结果较为稳定，都有较高的预测精度，表明本文方法具有更优的预测性能。

表 2 负荷预测精度结果
Tab. 2 Comparisons of load forecasting accuracy

日期	评价标准	结果							
		MLP	RBF	SVR	GRU	CNN	AE-GRU	CNN-GRU	本文方法
5 月 5 日	$M_{APE}/\%$	4.19	4.37	4.16	3.66	3.91	3.90	2.84	2.41
	R_{MSE}/kW	23.90	28.49	26.26	22.14	24.56	25.13	17.81	13.54
5 月 6 日	$M_{APE}/\%$	3.26	3.73	4.01	3.44	3.36	3.41	3.11	2.72
	R_{MSE}/kW	19.03	21.50	23.29	20.76	20.85	20.69	17.40	15.71
5 月 7 日	$M_{APE}/\%$	3.47	3.06	3.09	3.52	3.70	3.61	3.27	2.10
	R_{MSE}/kW	20.18	18.47	19.09	21.30	23.01	23.47	18.51	13.50
5 月 8 日	$M_{APE}/\%$	4.06	3.39	3.01	4.01	4.01	4.15	3.19	2.56
	R_{MSE}/kW	24.00	24.25	23.16	24.46	23.78	24.94	17.52	15.52
5 月 9 日	$M_{APE}/\%$	4.48	4.63	3.79	4.88	4.50	4.57	3.35	2.75
	R_{MSE}/kW	27.12	26.83	24.18	28.79	27.43	28.00	20.34	15.43
5 月 10 日	$M_{APE}/\%$	3.67	4.25	4.40	4.62	3.54	3.52	2.73	2.60
	R_{MSE}/kW	21.99	25.88	28.37	28.22	23.73	23.87	17.90	14.67
5 月 11 日	$M_{APE}/\%$	4.73	3.68	3.32	4.77	4.86	4.52	3.61	2.80
	R_{MSE}/kW	26.31	21.66	21.00	26.73	27.71	25.55	21.76	16.01
平均值	$M_{APE}/\%$	3.98	3.87	3.68	4.13	3.98	3.95	3.16	2.56
	R_{MSE}/kW	23.22	23.87	23.62	24.63	24.44	24.52	18.75	14.91

为了更加清晰直观的展现负荷预测结果，图 6 为 2016 年 7 月 30 日的真实值和不同方法的短期负荷预测曲线对比图。由图可看出本文预测模型拟合效果更好，且预测精度更高。此数据集的日负荷变化较大，早上负荷值变化有明显的上升，相较于其他模型，本文方法能在此时段更加准确平滑的预测负荷值，且在较为平缓的时段也能较好的捕捉到负

荷变化规律。

4 结论

本文提出一种基于 Attention 机制的 CNN-GRU 模型的短期电力负荷预测方法，该方法通过 CNN 对历史负荷序列提取特征，利用与 Attention 机制结合的 GRU 网络对所提特征进行动态时序建模，完成短期电力负荷预测。本文和 MLP、RBF、SVR、GRU、CNN、AE-GRU 和 CNN-GRU 进行对比，提出的预测模型能够获得更高的负荷预测精度，也具有更广泛的应用潜力。未来可以研究其他负荷预测方法相结合，进一步提高预测模型的普适性和处理数据的能力。

参考文献

[1] Khwaja A S , Zhang X , Anpalagan A , et al . Boosted neural networks for improved short-term electric load forecasting[J] . Electric Power

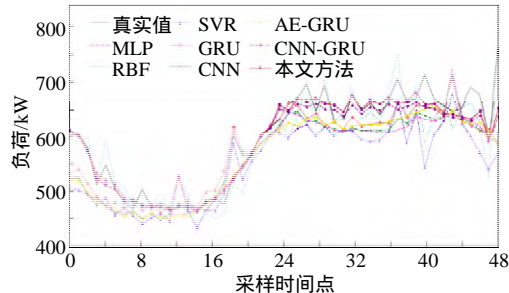


图 6 2016 年 7 月 30 日预测结果对比
Fig. 6 Comparison of forecasting results on July 30, 2016

- Systems Research, 2017(143): 431-437.
- [2] Nury A H, Hasan K, Alam M J B. Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh[J]. Journal of King Saud University-Science, 2017, 29(1): 47-61.
 - [3] Qiu X, Ren Y, Suganthan P N, et al. Empirical mode decomposition based ensemble deep learning for load demand time series forecasting[J]. Applied Soft Computing, 2017(54): 246-255.
 - [4] 赵峰, 孙波, 张承慧. 基于多变量相空间重构和卡尔曼滤波的冷热电联供系统负荷预测方法[J]. 中国电机工程学报, 2016, 36(2): 399-406.
Zhao Feng, Sun Bo, Zhang Chenghui. Cooling, heating and electrical load forecasting method for CCHP system based on multivariate phase space reconstruction and kalman filter[J]. Proceedings of the CSEE, 2016, 36(2): 399-406(in Chinese).
 - [5] Eroshenko S A, Poroshin V I, Senyuk M D, et al. Expert models for electric load forecasting of power system[C]// Young Researchers in Electrical & Electronic Engineering. St. Petersburg: IEEE, 2017: 1507-1513.
 - [6] 蒋敏, 顾东健, 孔军, 等. 基于在线序列极限支持向量回归的短期负荷预测模型[J]. 电网技术, 2018, 42(7): 2240-2247.
Jiang Min, Gu Dongjian, Kong Jun, et al. Short-term load forecasting model based on online sequential extreme support vector regression[J]. Power System Technology, 2018, 42(7): 2240-2247(in Chinese).
 - [7] Houimli R, Zmami M, Salha O B. Short-term electric load forecasting in Tunisia using artificial neural networks [J/OL]. Energy Systems, 2019: 1-19[2019-07-20]. <https://link.springer.com/article/10.1007/s12667-019-00324-4#citeas>.
 - [8] 冯凯. 基于大数据的居民用电行为分析与负荷预测[D]. 北京: 华北电力大学, 2017.
 - [9] 苏小林, 刘孝杰, 阎晓霞, 等. 计及需求响应的主动配电网短期负荷预测[J]. 电力系统自动化, 2018, 42(10): 60-66.
Su Xiaolin, Liu Xiaojie, Yan Xiaoxia, et al. Short-term load forecasting of active distribution network based on demand response[J]. Automation of Electric Power Systems, 2018, 42(10): 60-66(in Chinese).
 - [10] 李冬辉, 尹海燕, 郑博文. 基于 MFOA-GRNN 模型的年电力负荷预测[J]. 电网技术, 2018, 42(2): 585-590.
Li Donghui, Yin Haiyan, Zheng Bowen. An annual load forecasting model based on generalized regression neural network with multi-swarm fruit fly optimization algorithm[J]. Power System Technology, 2018, 42(2): 585-590(in Chinese).
 - [11] Wang S X, Wang X, Wang S M, et al. Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting[J]. Electrical Power and Energy Systems, 2019(109): 470-479.
 - [12] 孔祥玉, 郑锋, 鄂志君, 等. 基于深度信念网络的短期负荷预测方法[J]. 电力系统自动化, 2018, 42(5): 133-139.
Kong Xiangyu, Zheng Feng, E Zhijun, et al. Short-term load forecasting based on deep belief network[J]. Automation of Electric Power Systems, 2018, 42(5): 133-139(in Chinese).
 - [13] Mohammad F, Lee K B, Kim Y C. Short term load forecasting using deep neural networks[EB/OL]. (2018-10-01). <https://arxiv.org/pdf/1811.03242.pdf>.
 - [14] Tong C, Li J, Lang C, et al. An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders[J]. Journal of Parallel and Distributed Computing, 2018(117): 267-273.
 - [15] Hao Z, Liu G, Zhang H. Correlation filter-based visual tracking via adaptive weighted CNN features fusion[J]. IET Image Processing, 2018, 12(8): 1423-1431.
 - [16] Dong X, Qian L, Huang L. Short-term load forecasting in smart grid: a combined CNN and K-Means clustering approach[C]// IEEE International Conference on Big Data and Smart Computing. Jeju: IEEE, 2017: 119-125.
 - [17] Wang Y, Chen Q X, Gan D H, et al. Deep learning-based socio-demographic information identification from smart meter data[J]. IEEE Transactions on Smart Grid, 2019, 10(3): 2593-2602.
 - [18] Graves A. Generating sequences with recurrent neural networks [EB/OL]. (2014-06-05). <https://arxiv.org/pdf/1308.0850.pdf>.
 - [19] Rahman A, Srikumar V, Smith A D. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks[J]. Applied Energy, 2018(212): 372-385.
 - [20] 张宇帆, 艾芊, 林琳, 等. 基于深度长短期记忆网络的区域级超短期负荷预测方法[J]. 电网技术, 2019, 43(6): 1884-1892.
Zhang Yufan, Ai Qian, Lin Lin, et al. A very short-term load forecasting method based on deep LSTM RNN at zone level[J]. Power System Technology, 2019, 43(6): 1884-1892(in Chinese).
 - [21] 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, 12(12): 4045-4052.
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[J]. Power System Technology, 2018, 12(12): 4045-4052(in Chinese).
 - [22] Ke K, Hongbin S, Chengkang Z, et al. Short-term electrical load forecasting method based on stacked auto-encoding and GRU neural network[J]. Evolutionary Intelligence, 2019, 12(3): 385-394.
 - [23] Cinar Y G, Mirisae H, Goswami P, et al. Period-aware content attention RNNs for time series forecasting with missing values[J]. Neurocomputing, 2018(312): 177-186.
 - [24] Bahdanau D, Cho K H, Bengio Y. Neural machine translation by jointly learning to align and translate[EB/OL]. (2014-09-01). <https://arxiv.org/pdf/1409.0473v2.pdf>.
 - [25] Ping-Huan K, Chiou-Jye H. A high precision artificial neural networks model for short-term energy load forecasting[J]. Energies, 2018, 11(1): 213-226.
 - [26] Wang Y, Liao W, Chang Y. Gated recurrent unit network-based short-term photovoltaic forecasting[J]. Energies, 2018, 11(8): 1-14.
 - [27] Cho K, Merriënboer B V, Bahdanau D, et al. On the properties of neural machine translation: encoder-decoder approaches[EB/OL]. (2014-09-03). <https://arxiv.org/pdf/1409.1259.pdf>.
 - [28] Kingma D P, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2014-12-22). <https://arxiv.org/pdf/1412.6980.pdf>.



赵兵

收稿日期: 2019-07-22.

作者简介:

赵兵(1971), 男, 博士研究生, 从事智能用电及电力数据挖掘方面的研究, E-mail: zhaob@epri.sgcc.com.cn;

高欣(1974), 男, 通信作者, 副教授, 主要研究方向为电力系统自动化, E-mail: xlhhh74@bupt.edu.cn.

(责任编辑 徐梅)