[电动汽车]

基于Bi-LSTM的EV充电负荷预测

文/上海市城市科技学校智能工程系　闫普虹　葛根波　姜　萍　汪　迎
上海市松江区新桥职业学校　杜振东

针对电动汽车（Electric Vehicle，EV）充电负荷序列较强的随机性和间歇性而导致预测准确度不高的问题，利用双向长短期记忆网络（Bi-Long Short-Term Memory，Bi-LSTM）善于处理时间序列数据的特点，提出一种基于Bi-LSTM的EV充电负荷预测方法。验证了基于Bi-LSTM的EV预测模型比LSTM和ARMA预测模型具有更高的准确性，可为EV优化调度提供辅助决策。

近年来，由于EV的零碳排放具有很好的环保效益，因而EV得到了大规模的推广与使用。然而，EV充电负荷作为一种新兴的主动性负荷，具有非线性、间歇性和随机性等特殊性，经典的人工神经网络预测方法应用于EV充电负荷预测时，往往因自身的局限性而导致模型预测准确度偏低。常德政采用BP-NN和RBF-NN预测模型对EV短期负荷进行预测，仿真结果表明RBF-NN的预测准确度远高于BP-NN，但因容易陷入局部最小点，依然不能满足EV充电负荷预测的要求。罗浩成等采用蒙特卡洛法模拟EV的充电负荷曲线，这个方法适用于缺少原始数据的场合，它避免了真实数据中含有的不良数据造成的结果偏差，但是蒙特卡洛模拟真实数据这一方法也同样导致预测结果与实际结果存在的偏差较大。于惠鸣采用深度神经网络（Deep Neural Networks，DNN）的非线性结构将输入节点的信息传到隐含层节点，并利用粒子群优化算法（PSO）逐层优化，克服了传统网络易陷入局部最优的弊端，明显提升了预测准确度，但由于电力负荷序列是动态非线性时间序列，因DNN仅具有前馈结构，没有反馈机制，无法有效刻画负荷序列的动态特征，致使DNN仍很难达到更高的准确度。为此，本文构建基于Bi-LSTM的EV负荷预测模型，该模型不仅可以依据本身输入信息是否有用而判定保留或遗忘，而且可以从正反两方向对具备较强随机性特征的EV负荷数据进行预测，从而提高了预测模型的准确度。

LSTM网络模型

LSTM是循环神经网络（Recurrent Neural Network，RNN）的一种。最初，RNN的隐藏层是单一状态，它擅长处理短期的输入信息，LSTM在此基础上增加了可以处理保存长期信息的状态，从而解决了RNN长期依赖信息的学习问题。同RNN一样，LSTM能处理各种长短不一的时序信号， RNN不能进行特征提取，只适合于序列建模，而LSTM则能自动学习时间序列的相关性。EV充电负荷数据是一种兼具非线性和动态特征的负荷序列，LSTM能较好地适应EV充电负荷序列数据的这一特点，可以对其进行预测。LSTM“神经元”运行原理如图1所示。

项目基金：国家自然科学基金项目（61673528）

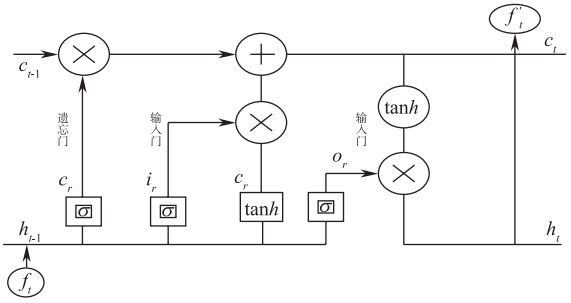

图1　LSTM“神经元”运行原理

图1中，圆形框代表“逐元素运算”，分别是乘运算和加运算两种，也就是说，两个相同维度的向量经过矩形框之后，对应元素进行相乘或相加。矩形框表示“激活运算”，激活函数分别是σ（Sigmoid）函数和tan$h$函数；与RNN相比，LSTM增加一条线；两条线按箭头方向融合，可以很好地表达输入中的长期依赖的信息。具体原理为

$$e_t=\sigma(W_f\times[h_{t-1},f_t]+b_f) \tag{1}$$

式中，e_t表示遗忘门限；h_{t-1}是上一个输出的信息；f_t为输入数据信息；W_f、b_f分别是遗忘门的权值矩阵和遗忘门的偏置。

$$i_t=\sigma(W_i\times[h_{t-1},f_i]+b_i) \tag{2}$$

$$\tilde{C}_t=\tan h(W_c\times[h_{t-1},f_t]+b_c) \tag{3}$$

$$C_t=e_t\times C_{t-1}+i_t\times\tilde{C}_t \tag{4}$$

式中，i_t表示输入门限；$\tilde{C}_t$表示前一刻的cell状态；W_i、W_c分别为Sigmoid函数和tan$h$函数训练的两个权值。Sigmoid函数决定需要更新的信息，tan$h$函数决定忘记的早期信息；$e_t\times C_{t-1}$表示忘记的信息，表示更新的信息，它们的累加就是当前的cell状态，即表示信息筛选更新的程度。

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

$$f'_t=o_t\times\tan h(C_t) \tag{6}$$

式（5）中表示输出门限，通过函数将提取出的信息输出，把细胞状态通过tan$h$函数（输出－1～1之间的值）进行处理，然后再乘上σ输出，最后输出的部分为式（6）中的f'_t。

构建基于Bi-LSTM的EV充电负荷预测模型

1. EV充电负荷特性分析

EV充电负荷特性主要分为内在负荷特性和外在负荷特性。前者是指EV充电负荷随着负荷端电压或系统频率变化而变化，研究内在负荷特性主要用于EV充电负荷的稳定性分析。后者是指EV充电负荷随着时间变化的规律而表现出来的外部特征，本文所研究的就是外在EV充电负荷特性。EV充电负荷的变化具有一定的随机性和周期性。

随机性：虽然每天的充电负荷按相似的规律变化，但是节假日与工作日的充电负荷需求相差较大，而且在工作日充电负荷曲线中，每天的充电负荷也不完全相同，这是由于每天的充电负荷会受到随机性因素的影响，如天气、温度等因素，节假日每天的充电负荷也会有不小的波动。

周期性：EV充电负荷作为电力负荷的重要组成部分，具有电力负荷变化的特点，一周的充电负荷当中，每天的充电负荷基本按相同的规律波动，呈现周期性的特点。把广东省某市的EV充电站每日24 h负荷数据按60 min为一个采样周期，共24点，7天共168点，绘制出一周的负荷曲线图如图2所示。

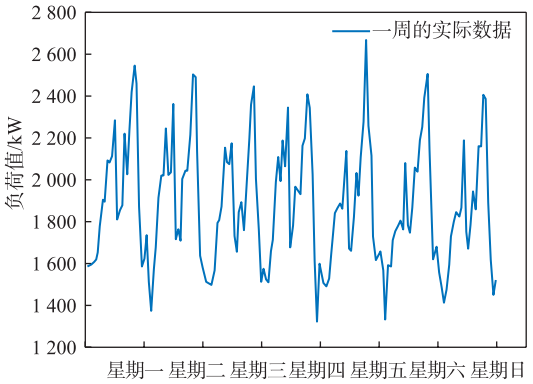

图2　一周的EV充电负荷变化曲线

2. Bi-LSTM应用于EV充电负荷预测的基本原理

Bi-LSTM是由RNN发展而来的学习网络，由前向LSTM与后向LSTM组合而成。Bi-LSTM可以理解为是对LSTM的改进，而LSTM是对RNN的改进。使用LSTM模型可以更好地捕捉到较长距离的依赖关系，虽然LSTM的训练过程可以记忆和遗忘信息；但是标准的LSTM按时间顺序处理序列，仅考虑历史信息并忽略未来信息，在预测未来负荷时，为了获得更准确的预测结果并充分利用历史和未来信息，LSTM显然不能满足这一需求，针对EV充电负荷存在的周期性和随机性特点，构建基于Bi-LSTM的EV充电负荷预测模型之前，应先对EV充电负荷数据进行预处理。

3. EV充电负荷数据预处理

本文的仿真试验的数据来源于广东省某市某年的部分EV充电负荷数据，Bi-LSTM的输入维度特征数据X包含历史数据L、时刻特征S、日期类型R和天气因素W共4个特征，即$X=\{L, S, R, W\}$。因EV充电负荷数据的相关性较强，预测方法是根据N个样本数据，生成第$N+1$、$N+2$、…个数据，每个数据含有24个负荷点，即根据N个样本数据$N\times 24$个负荷点得到第$N+1$天、$N+2$天的负荷值，然后再以此类推，不断滚动预测。输入数据维度汇总见表1。

表1　输入数据维度汇总

影响因素	不同维度	描述
历史数据	L（d-n，t）	前n天t时刻负荷值
时刻特征	S：不同时刻	高峰或低谷
日期类型	R：不同日期	工作日/节假日
天气因素	W：温度，湿度，风雨可能性	未来一天t时刻的预测温度、湿度与起风、降雨概率

由于EV充电负荷历史数据采集过程中多种不确定性因素影响导致的数据失真，从而严重影响模型预测的准确性，因此有必要在模型训练的前一步进行数据的预处理。因EV充电负荷历史数据的周期性，假定短时间内数据横向相似，也就是不会有突变，采用横向比较法进行不良数据的识别处理，利用设定阈值判断是否有不良数据，具体步骤为

1）首先计算历史数据的均值和方差，即

$$\overline{X}_{n,i}=\frac{1}{N}\sum_{n=1}^{N}X_{ni} \tag{7}$$

$$\sigma^2=\frac{1}{N}\sum_{n=1}^{N}(X_{n,i}-\overline{X}_{n,i})^2 \tag{8}$$

2）基于3σ原理，采用式（9）进行数据判断，其中ε为阈值，通常取1～1.5σ。所谓3σ原理又称为拉依达准则，先假设一组检测数据只含有随机误差，对其进行计算处理得到标准误差，按一定概率确定一个区间，认为凡超过这个区间的误差，就不属于随机误差而是粗大误差，含有该误差的数据应予以剔除。

$$\left|X_{n,i}-\overline{X}_{n,i}\right|>3\sigma_i\varepsilon \tag{9}$$

3）若数据满足式（9），则为异常数据，采用式（10）进行权重修正。

$$X_{n,i}^{*}=\frac{\alpha_1}{2}\sum_{n=1}^{N}X_{n+1,i}+\frac{\beta_1}{2}\sum_{n=1}^{N}X_{n,i}^{1,2}+\gamma_1\overline{X}_{n,i} \tag{10}$$

式中，$\alpha_1+\beta_1+\gamma_1=1$；$X_{n,i}^{*}$为第$n$天第$i$点修正数据；$X_{n+1,i}$为$X_{n,i}$附近两个横向负荷点；$X_{n,i}^{1,2}$为$X_{n,i}$距离最近的两个相似日的负荷点。

4）经过修正后，EV充电负荷数据集可以表示为式（11），采用式（12）、式（13）对其进行归一化处理。

$$\boldsymbol{X}=\begin{pmatrix}x_{11}\cdots x_{1i}\\ \vdots\quad\vdots\\ x_{n1}\cdots x_{ni}\end{pmatrix} \tag{11}$$

$$\boldsymbol{X}^{R}=\begin{pmatrix}x_{11}^{R}\cdots x_{1I}^{R}\\ \vdots\quad\vdots\\ x_{n1}^{R}\cdots x_{ni}^{R}\end{pmatrix} \tag{12}$$

$$x_{jk}^{R}=\frac{x_{jk}-\min x_{jk}}{\max x_{jk}-\min x_{jk}} \tag{13}$$

式中，$\min x_{jk}$为变量的最小值，$\max x_{jk}-\min x_{jk}$为变量的极值，经过式（13）变换后，所得充电负荷数据均值在0～1之间，由此完成归一化。

4. EV充电负荷预测的实现及算法流程

对原始EV充电负荷数据处理后，构建Bi-LSTM的EV充电负荷预测模型，工作原理如图3所示。双向LSTM由两个普通的LSTM组成，正向的LSTM利用过去的信息，逆向的LSTM利用未来的信息，这样在时刻t，既能够使用$t-1$时刻的信息，又能够利用$t+1$时刻的信息。其中x_1, x_2, …, x_i和y_1, y_2, …, y_i分别是EV充电负荷数据的输入和输出，s_i、s_o和M_i、M_o分别是Bi-LSTM正反向隐含层激活函数。A_0, A_1, …, A_n为正向的LSTM，参与正向计算，t时刻的输入为t时刻的序列数据x_t和$t-1$时刻的输出A_{t-1}。a_0, a_1, …, a_n为逆向的LSTM，参与反向计算，t时刻的输入为t时刻的序列数据x_t和$t+1$时刻的输出a_{t+1}。t时刻的最终输出值取决于A_{t-1}和a_{t+1}。

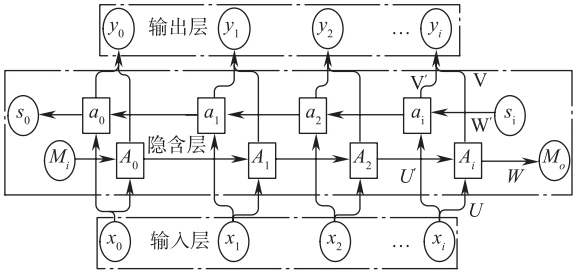

图3　基于Bi-LSTM的EV充电负荷模型工作原理

Bi-LSTM的隐藏层要保存两个值，即A_i和a_i，A_i是正向训练时刻值，a_i是反向训练时刻值。最终的EV充电负荷预测值y_i取决于A_i和a_i。基于Bi-LSTM的EV充电负荷预测模型为

$$y_i = g(VA_i + V'a_i) \tag{14}$$

$$a_i = f(W'a_i + U'x_i) \tag{15}$$

$$A_i = f(WA_i + Ux_i) \tag{16}$$

式中，g和f一般为非线性激活函数；U是正向计算从输入层连接到隐含层的权值；W是正向计算从隐含层到自己的连接权值；V是正向计算从隐含层到输出层的权值；而U'、V'、W'则是反向计算对应的权值。

LSTM循环神经网络的主要用途是处理和预测序列数据，在网络结构上，能记忆之前的信息，并利用之前的信息影响后面节点的输出，而图3的Bi-LSTM网络结构构成了一个无环图，输出是同时考虑了前后因素得到的，最终的结果为正向LSTM的结果与反向LSTM的结果的简单堆叠，因此更具有鲁棒性，提升了EV充电负荷预测模型的预测准确度。图4为基于Bi-LSTM的EV充电负荷预测流程图，从整体的角度给出了预测模型的运行过程与实现原理。

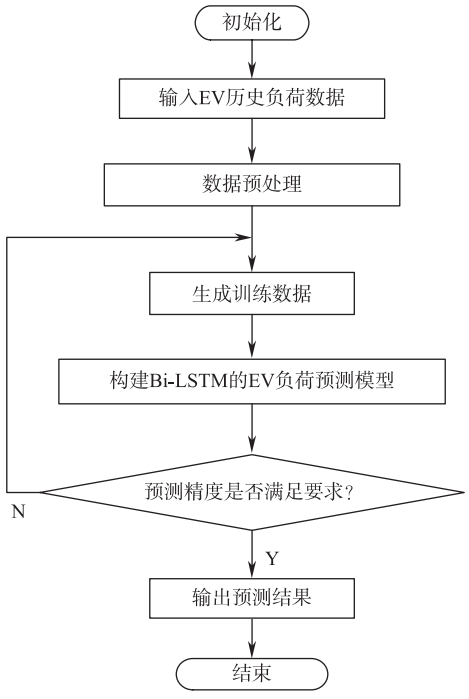

图4　基于Bi-LSTM的EV充电负荷预测流程

5. 评价指标

基于Bi-LSTM的EV充电负荷预测模型的评价指标为相关系数和均方根误差，定义如下：

1）相关系数是一种研究变量之间线性相关程度的量，取值范围为（－1，1），其大小决定了相关密切程度，当相关系数为1时，为完全正相关；当相关系数为－1时，为完全负相关；相关系数的绝对值越大，相关性越强；相关系数越接近于0，相关度越弱。具体计算公式为

$$R\sqrt{1-\frac{\sum_{i=1}^{n}(p_i-\bar{y})}{\sum_{i=1}^{n}(y_i-\bar{y})}} \tag{17}$$

式中，y_i为真实值；$\bar{y}$为真实值集的平均值；p_i为预测值，i=1, 2, …, n。

2）均方根误差（Root Mean Square Error，RMSE），也称为标准误差，是预测值和真实值偏差的二次方与观测次数n比值的平方根，能够很好地反映测量的准确度。具体计算公式为

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-p_i)^2} \tag{18}$$

试验与分析

为了验证Bi-LSTM模型的有效性，利用三组10天的负荷数据预测下周的负荷数据，分别使用ARMA、LSTM 和Bi-LSTM对处理后的EV负荷序列数据（共720+168组采样点）进行对比仿真试验，其中，前30天共720组数据为训练集，后7天共168组为验证集。其中LSTM 和Bi-LSTM配置的参数为：优化方法使用Adam（自适应矩估计迭代次数设置为100），选择参数均采用单层网络，隐藏层单元为100，训练轮数epoch1500，初始学习率为0.001。根据前文提到的两种评价指标分别对ARMA、LSTM 和Bi-LSTM模型得到的预测值与真实值进行计算，结果见表2。ARMA、LSTM 和Bi-LSTM模型得到的EV充电负荷预测分别如图5～图7所示。

表2　不同预测模型的结果对比

预测模型	RMSE	值
ARMA	130.344	0.886 53
LSTM	121.612 2	0.901 04
Bi-LSTM	14.095 5	0.999 88

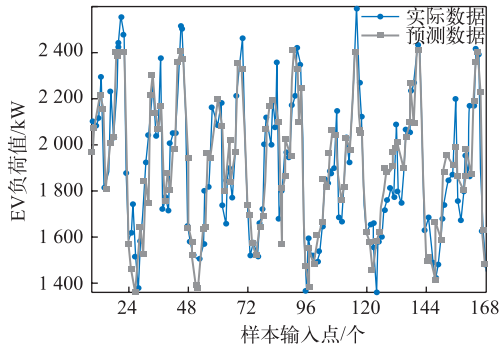

图5　基于ARMA的EV充电负荷预测

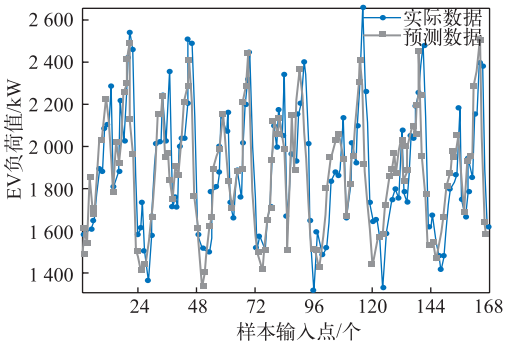

图6　基于LSTM的EV充电负荷预测

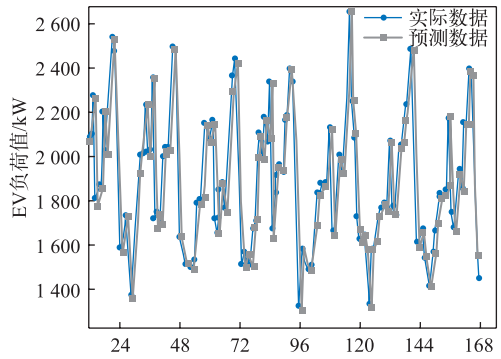

图7　基于Bi-LSTM的EV充电负荷预测

图5～图7中细实线是EV充电负荷的实际曲线，粗实线分别对应ARMA、LSTM和Bi-LSTM的预测曲线，可以得出EV充电负荷曲线呈现一定的周期性和随机性，三种预测方法跟踪EV实际充电负荷预测曲线的能力依次增强。

1）ARMA模型本质上只能描述平稳负荷序列的随机过程，而EV充电负荷序列是随时间变化的非平稳随机过程，具有一定的周期性和生长趋势，因此，ARMA模型无法表示出EV充电负荷序列数据这一复杂非线性序列的变化规律，在RSME和*R*上表现最差。

2）LSTM模型不仅能很好地捕捉EV充电负荷序列数据的时序性特点，而且可以在改进RNN的基础上，增强长期依赖信息的能力，因此可以兼顾EV充电负荷数据的非线性，比ARMA具有更好的预测效果。但是不能捕捉从后到前的关系（Bi-LSTM比LSTM在RSME和*R*指标上分别提升了88.41%、9.89%），因此，预测效果不及Bi-LSTM。

3）Bi-LSTM模型能很好地捕捉EV充电负荷序列数据的时序性特点，并且可以更好地捕捉双向的依赖信息，因此，预测效果在三种方法中最为突出。与ARMA模型相比，Bi-LSTM在RSME和*R*指标上分别提升了89.19%和11.33%。

结束语

本文针对存在时序性、受多个变量影响的非平稳的EV充电负荷序列数据，构建基于Bi-LSTM的EV充电负荷预测模型，并且对比了ARMA和LSTM预测模型。通过试验表明，Bi-LSTM模型能够更好地挖掘出EV充电负荷序列从后到前的内在变化规律，从而提升了预测准确度，验证了基于Bi-LSTM的EV充电负荷预测模型的有效性，可应用于EV短期充电负荷预测，为EV优化调度提供参考。

虽然Bi-LSTM能从两个方向上训练输入信息，提升了原有模型的预测准确度，但是Bi-LSTM的训练网络结构与RNN相比更复杂，GRU（Gated Recurrent Unit）是LSTM网络的一种效果较好的变体，较LSTM网络的结构更加简单，容易实现，面对EV充电负荷数据的不确定性，采用Bi-GRU模型理论上能取得更好的预测结果，因此，基于的Bi-GRU的EV充电负荷预测模型的建立有待于进一步研究和验证。**EA**

（收稿日期：2021-11-20）

参考文献

[1] 邢恩恺．基于改进VMD预处理与双向LSTM的风功率预测研究[D]．吉林：东北电力大学，2019．

[2] 常德政．智能电网中电动汽车充电站短期负荷预测模型研究[D]，青岛：青岛大学．2015．

[3] 杨廷方，刘沛，李浙，等．应用新型多方法组合预测模型估计变压器油中溶解气体浓度[J]．中国电机工程学报，2018，28(31)：108-113．