

基于谱聚类和 LSTM 神经网络的 电动公交车充电负荷预测方法

王哲¹, 万宝², 凌天晗², 董晓红³, 穆云飞³, 邓友均³, 唐舒懿³

(1. 国网天津市电力公司, 天津市 300010; 2. 国网天津市电力公司滨海供电分公司, 天津市 300450;
3. 智能电网教育部重点实验室(天津大学), 天津市 300072)

摘要: 目前电动公交车的渗透率较大, 且充电频率和充电量较高, 故而其充电负荷对电网运行与调度产生着不可忽略的影响。因此, 电动公交车的充电负荷预测研究具有重要的理论与现实意义, 但由于公交车间歇性与随机性的充电行为在时间上给充电负荷预测增加了难度。为此, 提出基于谱聚类和长短期记忆(long short-term memory, LSTM)神经网络的电动公交车充电负荷预测方法。首先, 利用考虑距离与形态的谱聚类, 对充电负荷曲线进行聚类; 其次, 综合考虑影响充电负荷的关键因素, 如温度、日类型等, 利用不同簇的总充电负荷, 分别训练 LSTM 神经网络的模型参数, 并预测每簇的充电负荷; 接着, 对不同簇的预测结果求和即可得到预测日的总充电负荷; 最后, 通过利用某市实际数据, 验证本文所提方法。结果表明, 所提方法充电负荷预测结果的平均绝对百分误差(mean absolute percentage error, MAPE)在 11% 以下, 预测准确度有所提升。

关键词: 谱聚类; 长短期记忆网络(LSTM); 电动公交车; 负荷预测

Electric Bus Charging Load Forecasting Method Based on Spectral Clustering and LSTM Neural Network

WANG Zhe¹, WAN Bao², LING Tianhan², DONG Xiaohong³,
MU Yunfei³, DENG Youjun³, TANG Shuyi³

(1. State Grid Tianjin Electric Power Company, Tianjin 300010, China; 2. State Grid Tianjin Binhai Electric Power Company, Tianjin 300450, China; 3. Key Laboratory of Smart Grid of Ministry of Education (Tianjin University), Tianjin 300072, China)

ABSTRACT: At present, the penetration rate, charging frequency and charging capacity of electric buses are relatively high, so the charging load has a non-negligible impact on the operation and dispatch of the power grid. So, the charging load forecasting research has important theoretical and practical significance, but the intermittent and random charging behavior increase the spatial forecasting difficulty. Therefore, the charging load forecasting method of electric buses is proposed on the basis of spectral clustering and long short-term memory (LSTM) neural network. First of all, the charging load curve is clustered according to spectral clustering considering the distance and the shape. And then, considering the key factors that affect the charging load, such as historical load, temperature and day type, the model parameter of LSTM neural network is trained using each cluster charging load, and the charging load of each cluster is predicted. Then, the total charging load of the forecasting day is to sum the forecasting results of different clusters. Finally, on the basis of the historical real data in a certain city, the proposed method is verified. The result shows the mean absolute percentage error (MAPE) of charging load prediction result of the proposed method is below 11%, and the accuracy of load forecasting is improved.

This work is supported by Science and Technology Program of State Grid Tianjin Electric Power Company (No. KJ20-1-38).

KEYWORDS: spectral clustering; long short-term memory network (LSTM); electric bus; load forecasting

中图分类号: TM 715 文献标志码: A 文章编号: 1000-7229(2021)06-0058-09
DOI: 10.12204/j.issn.1000-7229.2021.06.006

0 引言

近年来, 随着电动汽车(electric vehicle, EV)的

发展, 作为智能电网和智能交通网络集成^[1]关键环节的 EV 及其充电设施的应用和研究正在如火如荼地进行中。其中公交车作为社会公共资源, 在该领域

基金项目: 国网天津市电力公司科技项目(KJ20-1-38)

的交通电力化方面有较好的推广和应用。目前电动公交车的渗透率较大,且充电频率和充电量较高,故其充电负荷对电网运行与调度产生着不可忽略的影响。因此,电动公交车充电负荷预测研究具有重要的理论和现实意义^[2]。而公交车白天发车班次较为固定,停留时间较短,因此公交车间歇性和随机性的充电行为在时间上增加了充电负荷预测的难度。

传统的电力负荷预测方法有回归分析法、相似日法等^[3]。随着分布式电源和电动汽车等新型负荷类型的大规模接入,对传统的负荷预测方法带来了极大挑战。相比于传统的电力负荷,考虑不同的充电方式、出行规律、充电效率、充电频率等特点,EV充电负荷时间上的分布与电力负荷的规律不同。此外,EV充电负荷受出行路况、天气、运营状态等多因素的影响,在时间上具有较大的随机性。

近年来,人工智能方法广泛应用于EV充电负荷预测领域,包括浅层网络和深度学习算法。但基于浅层网络的传统方法,如人工神经网络算法和支持向量机(support vector machine, SVM)通常会陷入学习速度慢、收敛速度慢等问题^[4]。而深度学习则具有出色的特征学习能力,并且可以用较少的参数表示复杂的功能。其中,文献[5]提出基于长短期记忆网络(long short-term memory, LSTM)的新型深度学习方法用于多时间尺度EV充电负荷预测;文献[6]提出了基于核主成分分析和非劣排序遗传算法优化卷积神经网络参数的EV充电站短期负荷预测方法;文献[7]将LSTM网络模型应用于电动汽车充电负荷的短期负荷预测中,实验证明使用LSTM预测具有精确性和有效性。此外,由于电动公交车发车的间隔时间较短,其充电负荷具有间歇性和时序特征,而利用LSTM神经网络模型作为时间循环神经网络可以有效解决EV充电的时间尺度问题,提高负荷预测的准确性。

另一方面,相比于其他类型电动汽车,如电动出租车和私家车等,电动公交车的工作时间和行驶路线较为稳定,但不同线路公交车受电动公交车工作模式、频率和发车时间等特征的影响,其行驶规律差异性较大,导致个体充电负荷差异性较大。而目前电动公交车负荷预测方法大多对电动公交车群体充电负荷进行预测^[8]。因此,可通过聚类算法将具有相同特性的EV用户划分为同一类簇,对不同类簇进行建模分析以考虑个体的差异性,提高总体负荷预测精度^[9]。在负荷聚类时,考虑负荷曲线趋势性和周期性的相似程度,正确衡量负荷随时间变化的形态和轮廓的相似性,才能把握用户的用电习惯和特性,从而

能够较好地实现对负荷的聚类^[10]。

近年来,针对用户负荷的聚类分析,已有大量研究。常用聚类算法主要包括K均值(K-means)聚类算法^[11]、基于密度含噪声的空间聚类(density-based spatial clustering of applications with noise, DBSCAN)算法^[12]、模糊C均值聚类算法(fuzzy C-means, FCM)^[13]和谱聚类算法^[14]等。文献[15]提出了加入凝聚层次聚类的一种改进K-means算法,实现对负荷曲线的聚类分析;文献[16]针对海量的负荷数据,利用信息熵分段聚合近似方法对数据集进行降维,再应用谱聚类算法对负荷曲线进行分类。其中谱聚类在数据降维、负荷分类有效性、稳定性和降低运算量等方面具有一定优势。但目前聚类方法多数仅将距离作为曲线相似度的判断依据。

为解决以上问题,本文采用基于距离与形态相似性度量的谱聚类算法,考虑电动公交车个体的差异性,对数据进行聚类处理。此外,EV充电负荷作为一种典型的时间序列数据,具有趋势性、周期性等时序特征。因此,本文提出一种基于谱聚类和LSTM神经网络的电动公交车充电负荷预测方法,以解决以上问题。

1 充电负荷预测方法框架

本文所提的充电负荷预测方法框架如图1所示。电动公交车充电负荷特征与其工作时间和行驶路线等特性密切相关。因此,在进行简单的数据预处理与清洗后,首先利用考虑距离与形态的谱聚类算法进行充电负荷聚类,将用电规律相似的电动公交车充电负荷聚为一簇;然后对各类电动公交车群体利用LSTM神经网络进行分组训练并预测充电负荷;最后,对不同LSTM神经网络得到的预测结果进行累加,得到最终的总充电负荷预测值。

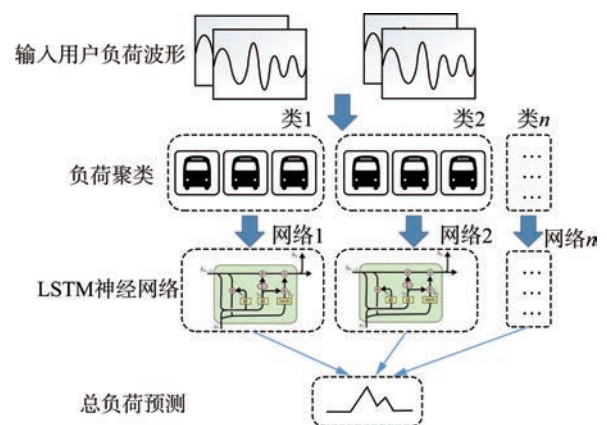


图1 充电负荷预测方法框架

Fig. 1 Framework of charging load forecasting method

公交车的工作时间和行驶路线有明确规定,其充电地点一般设在公交车始发站^[17],因此其充电负荷空间特性较为固定。但由于电动公交车发车时间间隔较短,不能进行长时间的持续充电,因此其充电负荷曲线具有间歇型特征。某个电动公交车日负荷曲线如图2所示。

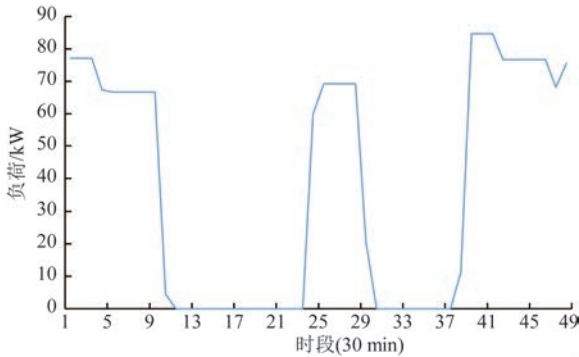


图2 电动公交车日负荷曲线

Fig. 2 Daily load curve of the electric bus

受电动公交车工作模式的影响,不同线路的充电负荷差异性较大。直接按照线路进行分组,就会忽略电动公交车个体每日的实际负荷情况,不能更好地在负荷量和负荷曲线趋势上区分日负荷,因此可利用聚类,考虑个体工作特性的差异性。在利用真实数据进行聚类之前,需要对原始数据进行如下的预处理和清洗。

1.1 数据预处理

负荷曲线主要依据其形状和距离特征,因此需要对数据进行标么化处理:

$$x'_i(t) = \frac{x_i(t) - x_{i \min}}{x_{i \max} - x_{i \min}} \quad (1)$$

式中: $x_i(t)$ 为充电负荷曲线 i 在时段 t 的实际负荷值; $x'_i(t)$ 为相应的标么化值; $x_{i \max}$ 、 $x_{i \min}$ 分别为充电负荷曲线 i 中的最大、最小负荷值。

负荷曲线经标么化,形成负荷矩阵 X ,如下式所示:

$$X = \begin{bmatrix} x'_{11} & \cdots & x'_{1T} \\ \vdots & & \vdots \\ x'_{n1} & \cdots & x'_{nT} \end{bmatrix}_{n \times T} \quad (2)$$

式中: n 为负荷曲线个数; T 为一日内的数据维数,即时段数; x'_{nT} 是第 n 条负荷曲线的第 T 时段数据。

1.2 数据清洗

在采集电动公交车的充电数据时,由于数据丢失、通信错误或者用户操作不当等原因,会造成数据异常^[18]。这些数据会影响聚类算法的效果,从而影响预测模型的预测精度,因此需要进行数据清洗。

2 考虑距离与形态的谱聚类算法

谱聚类算法是一种基于谱图理论的聚类方法。它构建基于相似度的无向加权图,从而将聚类问题转化为图分区问题。该算法是通过将以上进行过预处理和清洗后的每个数据,设置为图形的顶点,将顶点之间的相似性量化设置为相应连接的权重^[19]。其中,图论的最佳准则是最大化子图内部的相似度,并最小化子图间的相似度,从而利用图论的最佳准则来划分图的区域。本文考虑距离与形态的相似性度量,计算负荷曲线的相似度,并利用谱聚类算法实现负荷曲线的分类。

2.1 基于曲线距离的相似性度量

基于曲线距离的相似性度量用于衡量个体在空间上存在的距离,距离越远说明个体间的差异越大。本文选取欧氏距离作为曲线距离的相似性度量,欧氏距离衡量的是多维空间中各个点之间的绝对距离。负荷曲线 i 与 j 之间的距离计算公式如下:

$$d_{ij} = \sqrt{\sum_{t=1}^T |x'_i(t) - x'_j(t)|^2} \quad (3)$$

式中: d_{ij} 为日充电负荷曲线 i 与 j 之间的欧式距离。

利用欧氏距离,构建基于曲线距离的相似度矩阵 D ,如式(4)所示:

$$D = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & & \vdots \\ d_{n1} & \cdots & d_{nn} \end{bmatrix} \quad (4)$$

式中: d_{11} 、 d_{22} 、 \dots 、 d_{nn} 均为0。

2.2 基于形态特性的相似性度量

负荷曲线的形态特性更能反映出负荷时间序列的形态或者轮廓的相似性^[20]。其中灰色关联分析法是根据因素之间发展趋势的相似或相异程度,来衡量因素间的关联程度,它能反映出曲线间几何形状的差别程度,度量曲线形态特征。本文选取灰色关联分析法作为度量曲线形态相似性的方法。负荷曲线 i 与 j 间的关联系数和关联度如下:

$$\xi_{ij}(t) = \frac{\min_{ij} |x'_i(t) - x'_j(t)| + \rho \max_{ij} |x'_i(t) - x'_j(t)|}{|x'_i(t) - x'_j(t)| + \rho \max_{ij} |x'_i(t) - x'_j(t)|} \quad (5)$$

$$\gamma_{ij} = \frac{1}{T} \sum_{t=1}^T \xi_{ij}(t) \quad (6)$$

式中: $\xi_{ij}(t)$ 为负荷曲线 i 与 j 在时段 t 的关联系数; $\rho \in (0, 1)$ 为分辨系数; γ_{ij} 为负荷曲线 i 与 j 之间的关联度。本文选取关联度作为曲线形态特性的相似性度量。

利用关联度,构建基于形态特性的曲线相似度矩

阵 Y ,如下所示:

$$Y = \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ \vdots & & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} \end{bmatrix} \quad (7)$$

2.3 充电负荷曲线的相似度

本文基于曲线距离与形态特性的相似性度量,计算负荷曲线的相似度矩阵 W ,如下所示:

$$\begin{cases} W = \alpha D + \beta Y \\ \alpha + \beta = 1 \end{cases} \quad (8)$$

式中: α 、 β 分别为相似度矩阵 D 、 Y 的权重系数。

2.4 具体步骤

步骤 1: 输入聚类所需的 $n \times T$ 维矩阵 X , 权重系数 $\alpha = \alpha_0$, 权重系数 $\beta = \beta_0$ 和聚类数 $K = K_0$;

步骤 2: 计算基于曲线距离与形态特性的相似度矩阵 D 、 Y 。根据式(8)可确定相似度矩阵 W ,且矩阵 W 是 $n \times n$ 对称矩阵,如下所示:

$$W = \begin{bmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & & \vdots \\ w_{n1} & \cdots & w_{nn} \end{bmatrix}_{n \times n} \quad (9)$$

式中: $w_{11}, w_{22}, \dots, w_{nn}$ 均 0。

构造度矩阵 S 如下:

$$S = \begin{bmatrix} s_1 & 0 & \cdots & 0 & 0 \\ 0 & & & & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & 0 & s_n \end{bmatrix}_{n \times n} \quad (10)$$

度矩阵 S 中元素 s_i ,可由式(11)求得:

$$s_i = \sum_{j=1}^n w_{ij} \quad (11)$$

步骤 3: 构造如下拉普拉斯矩阵 L

$$L = S - W \quad (12)$$

步骤 4: 根据式(13),计算标准化后的拉普拉斯矩阵 L' :

$$L' = S^{-\frac{1}{2}} \times L \times S^{-\frac{1}{2}} \quad (13)$$

步骤 5: 取前 K 个最小特征值,计算特征向量;利用以上 K 个特征向量,组成新的矩阵 $V_{n \times K}$,构成特征向量空间;

步骤 6: 利用 K -means 聚类算法对特征向量空间 $V_{n \times K}$ 进行聚类,对应于原始数据的聚类,得到 K 簇的划分:

$$B_k = \{r_1, r_2, \dots, r_u\} \quad (14)$$

式中: B_k 为第 k 簇的负荷曲线编号集合; r_u 为第 u 条负荷曲线; u 为该簇中负荷曲线的个数。

步骤 7: 判断是否满足终止条件:若满足,则执行步骤 8;若不满足,执行 $\alpha = \alpha + \tau$, $\beta = 1 - \alpha$, τ 为固定步长。重复步骤 4~7,直到满足终止条件,并记录下

聚类数、相应的轮廓系数(silhouette coefficient, SC)指数、戴维森堡丁指数(Davies - Bouldin index, DBI)和每簇的负荷曲线。其中,SC 指数结合内聚度和分离度两种因素评价聚类效果。SC 指数越接近边界值表示聚类效果越好。DBI 表示类内距离平均值之和与类间距离之比的最大值,其值越小表示不同类相离越远,聚类效果越好。终止条件如下:

- 1) $\alpha = 1$;
- 2) 相邻两次 SC 或 DBI 变化小于阈值;

步骤 8: 执行 $K = K + 1$,判断聚类数 K 是否小于预设的最大值 K_{\max} ;若小于,则重复步骤 5~8,直到 $K = K_{\max}$ 。并从中选择 SC 或 DBI 指数最优对应的聚类数 K ,从而确定将负荷分成几簇以及每簇的负荷曲线。

3 LSTM 神经网络预测

EV 充电负荷的大小会随着用电时间有明显的变化,属于时序性很强的数据。其中循环神经网络(recurrent neural network, RNN)是处理时间序列强大而有用的工具^[21],但 RNN 对长期依赖信息的学习能力不足,会出现梯度消失等问题。而本文选用的 LSTM 神经网络可以有效解决 RNN 网络梯度消失和梯度爆炸等问题。电动公交车日负荷曲线经过上述的数据处理及谱聚类后,根据得到的不同聚类群体,将不同簇分别采用 LSTM 神经网络进行训练,得到适用于每个群体的网络结构,从而实现较为精准的充电负荷预测。

进行 LSTM 网络训练前,需要确定 LSTM 网络输入特征。输入特征主要为影响预测点充电负荷大小的因素,包括社会因素(如日期、星期、工作日或节假日等)及天气因素(如温度、湿度、降雨量等)^[22]。处于不同环境时,EV 充放电效果和续航能力都不同,负荷需求量差异较大,如在较低温度开启空调后,同一辆公交车的充电量将会增加。考虑电动公交车的负荷特性,本文选取以下影响因素作为 LSTM 神经网络的输入,其中历史负荷是将该簇内的实际充电负荷按照相应日期和相应时刻点累加,作为该簇的历史负荷值:

- 1) 负荷点所属日期 D_{ATE} ;
- 2) 负荷点所属时段 t ;
- 3) 共 T 个时段数;
- 4) 负荷点所属星期类别 e ;
- 5) 负荷点的温度 T_{emp} ;
- 6) 负荷点的降雨量 P_{rei} ;
- 7) 预测点前一时段 $t-1$ 的历史负荷。

上节所述的谱聚类算法主要是通过分析日负荷曲线特征进行分组,然后将具有相同相似性的曲线划分为一类,就是为每一条日负荷曲线打好标签(如 1, 2, …, 13)。得到每类标签后,将相同类别的日负荷曲线中每日的每时刻数据累加,就能得到该簇的总负荷数据,并可通过标签找到该日对应的日期和星期,从而找到对应的天气数据。然后,按照一定比例 p_r 将该簇的历史负荷划分为训练集与测试集。训练集用于训练 LSTM 神经网络,调整选取最优网络参数,测试集用于测试最终的神经网络性能。将该簇的总日负荷数据、日期、时间、星期类别与相应的天气数据作为神经网络的输入。通过训练 LSTM 网络结构,每个簇对应一种 LSTM 结构,其输出为预测的该簇电动公交车的充电负荷。最后将每簇的预测结果直接相加就可得到预测的电动公交车总群体的充电负荷。其中, LSTM 神经网络单元结构如图 3 所示。图中 x_t 和 h_t 分别是输入值与中间状态; x_{t-1} 、 x_{t+1} 、 h_{t-1} 和 h_{t+1} 分别是 $t-1$ 时刻和 $t+1$ 时刻的输入值与 $t-1$ 时刻与 $t+1$ 时刻的中间状态; σ 为 sigmoid 函数, \tanh 为双曲正切函数, A_1 、 A_2 分别表示上一、下一时刻的门。

LSTM 通过利用“门”来控制丢弃或者增加信息,从而实现遗忘或记忆的功能,以避免 RNN 长期依赖的问题^[23]。一个 LSTM 单元有 3 个这样的门,分别是遗忘门、输入门、输出门,计算公式如(15)~(20)所示:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (15)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (16)$$

$$\hat{S}_t = \varphi(W_s[h_{t-1}, x_t] + b_s) \quad (17)$$

$$S_t = i_t \odot \hat{S}_t + S_{t-1} \odot f_t \quad (18)$$

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (19)$$

$$h_t = O_t \odot \varphi(S_t) \quad (20)$$

式中: f_t 、 W_f 、 b_f 分别为遗忘门的输出、权重和偏置; σ

为 sigmoid 函数; i_t 为需要记忆的信息; W_i 为输入门权重; W_s 为记忆单元权重; b_i 为输入门偏置; b_s 为记忆单元偏置; \hat{S}_t 为候选记忆单元,用于更新记忆单元; φ 为 \tanh ; S_t 为记忆单元; S_{t-1} 为上一时刻记忆单元; \odot 为哈达玛积; O_t 、 W_o 、 b_o 分别为输出门的输出、权重和偏置。

每个门由一个 σ 神经网络层和一个乘法运算组成。 σ 函数的输出在 0 和 1 之间,指示输入数据信息的保持程度。0 表示完全屏蔽,1 表示完全通过。因此,通过将每个矢量相加并相乘,最终实现了充电负荷的输出。

本文在评价 LSTM 预测模型时,选择平均绝对百分误差(mean absolute percentage error, MAPE)、平均绝对误差(mean absolute error, MAE)以及均方根误差(root mean square error, RMSE)来描述预测效果,计算公式分别如下:

$$\xi_{\text{MAPE}}(k) = \frac{1}{T} \sum_{t=1}^T \left| \frac{f_{kt} - \hat{f}_{kt}}{f_{kt}} \right| \times 100\% \quad (21)$$

$$\xi_{\text{MAE}}(k) = \frac{1}{T} \sum_{t=1}^T |f_{kt} - \hat{f}_{kt}| \times 100\% \quad (22)$$

$$\xi_{\text{RMSE}}(k) = \sqrt{\frac{1}{T} \sum_{t=1}^T (f_{kt} - \hat{f}_{kt})^2} \times 100\% \quad (23)$$

式中: f_{kt} 为第 k 簇负荷曲线在预测日 t 时刻的实际值, \hat{f}_{kt} 为对应预测值。

4 算例分析

本文采用某市 2019 年 12 月份全天所有电动公交车充电数据作为测试数据,具体包含 103 个电动公交车编号,及其每次充电的交易电量,充电开始时间和充电结束时间。2019 年 12 月份全天的天气数据可由中国气象数据网获得^[24],并作为电动公交车充电负荷的影响因素加以考虑。按 30 min 间隔统计充电负荷,并假设一日内气象数据不变。

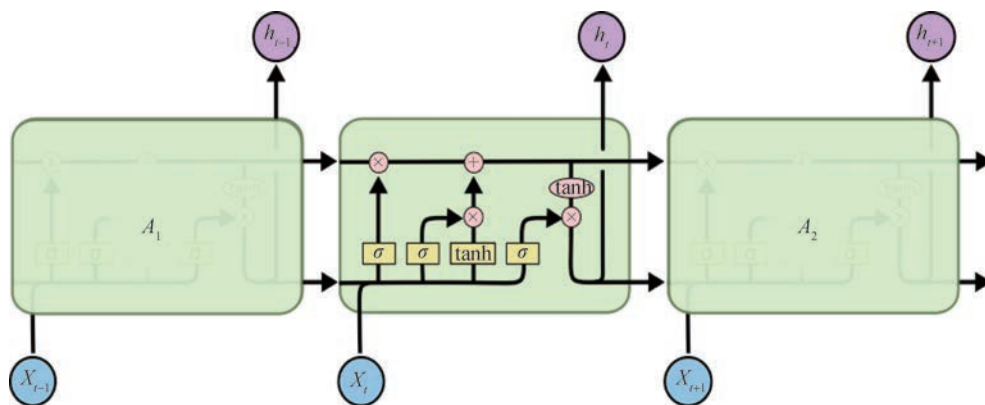


图 3 LSTM 网络单元结构
Fig. 3 Structure of LSTM network unit

本文利用的是 31 天的充电负荷数据,即 $D_{ATE} \in \{1, 2, \dots, 31\}$;一天内的数据个数,即 $t \in \{1, 2, \dots, 48\}$;一周内有 7 种星期类型,即 $e \in \{1, 2, \dots, 7\}$ 。

4.1 谱聚类结果分析

本文选取 103 辆电动公交车 31 天日充电负荷数据,利用第 2 节所提的谱聚类对 3 193 组数据进行聚类分析,取 $\rho = 0.5$,初始权重系数 α 和 β 分别为 0.05 和 0.95,固定步长 τ 为 0.05。SC 指数和 DBI 指数随聚类数 K 变化趋势如图 4 所示。由图 4 可以看出,当 $K = 13$ 时 SC 指数和 DBI 指数都达到最优取值。此时,权重系数 α 和 β 分别为 0.95 和 0.05。因此,最后聚类数选定为 13,每类的中心线如图 5 所示。

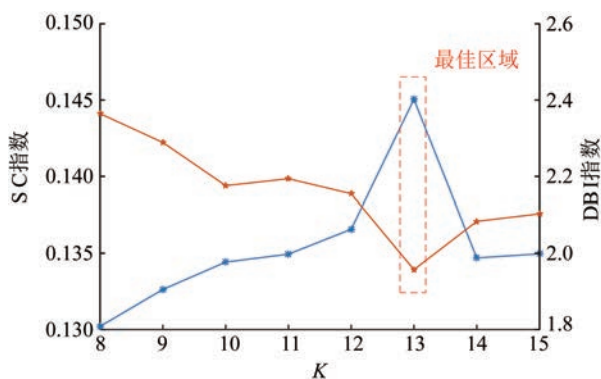


图 4 SC 指数和 DBI 指数随聚类数 K 的变化趋势
Fig. 4 Change of SC and DBI indices with the amount of the amount K of the clusters

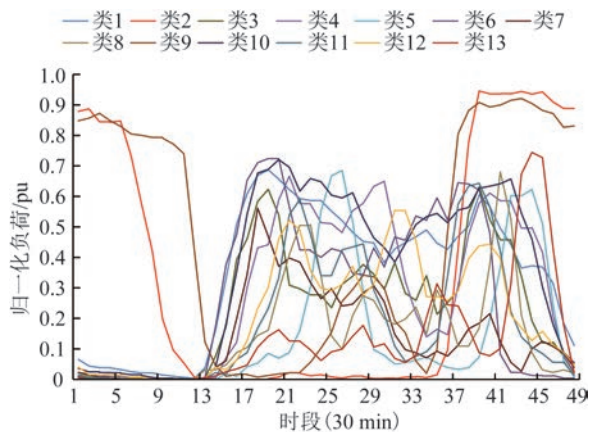


图 5 聚类中心线
Fig. 5 Cluster center line

由图 5 可以看出,不同线路的电动公交车日充电负荷曲线存在明显的差异。主要是由于公交车的工作性质,白天需要按点发车,没有大量时间集中充电,因此只能利用早中午、晚上休息时间以及到达始发站的停车时间进行充电。而不同线路的公交车发车班次、行驶里程都不一样,从而充电负荷曲线具有不同的特征。其中类 2 与类 9 的充电大量集中在早晚

时段,同时类 9 在中午也进行充电;类 5 与类 12 存在明显的用电尖峰,但充电时间与频率有所不同;类 13 集中在晚上充电,白天充电负荷较少;其余各类在白天到达始发站及中晚休息时间均有充电,但充电负荷大小和班次停靠时间的不同,负荷也不相同,波动性较强。

4.2 负荷预测结果分析

通过第 2 部分的聚类分析,得到 13 簇充电负荷曲线分类,每簇中既包含不同线路车辆也包含不同日期的充电负荷曲线。将每簇中充电负荷真实值按照日期和时刻进行累加,以作为每个 LSTM 网络的输入,进行负荷训练与预测。对每类数据按照 8:2 划分训练集与测试集。

实验在 Windows 系统下利用 TensorFlow 开发完成。LSTM 神经网络使用 Keras Python 包对时间序列预测。在建立 LSTM 网络结构后,通过观测训练时的损失值来衡量模型训练的效果。当所设参数能使每一类训练损失值达到收敛且趋于稳定,则为该类最佳参数,得到各类别 LSTM 网络预测性能如表 1 所示。由表 1 可知,由于各类群体自身特征不同,预测性能也各不相同。将各类预测日结果累加,可得到最终总日负荷预测结果。13 类数据均达到最佳预测效果,每类损失值均能收敛且达到稳定。

表 1 每类的 LSTM 结构参数与预测性能
Table 1 Structural parameters and forecasting performance of LSTM networks for each cluster

| 类别 | LSTM 隐含层 神经元/个 | MAPE/% | MAE/% | RMSE/% |
|----|-------------------|--------|-------|--------|
| 1 | 79 | 2.29 | 2.003 | 2.967 |
| 2 | 88 | 6.06 | 1.643 | 2.803 |
| 3 | 70 | 4.92 | 1.617 | 2.801 |
| 4 | 45 | 6.83 | 2.082 | 2.719 |
| 5 | 70 | 4.78 | 1.949 | 2.806 |
| 6 | 80 | 5.65 | 2.671 | 3.661 |
| 7 | 77 | 7.46 | 2.218 | 3.108 |
| 8 | 89 | 7.52 | 2.741 | 4.240 |
| 9 | 44 | 3.35 | 1.320 | 2.208 |
| 10 | 56 | 3.38 | 1.929 | 2.614 |
| 11 | 64 | 5.81 | 2.289 | 3.424 |
| 12 | 42 | 4.46 | 1.894 | 2.740 |
| 13 | 79 | 5.57 | 1.776 | 2.772 |

训练时,类 1 和类 2 的训练和预测数据分别如图 6 和图 7 所示,可以看出不同类的 EV 充电负荷具有明显差异,但类间的日负荷分布规律较为一致,且预测精度较高。图 8 为类 1 和类 2 训练时随着训练次数

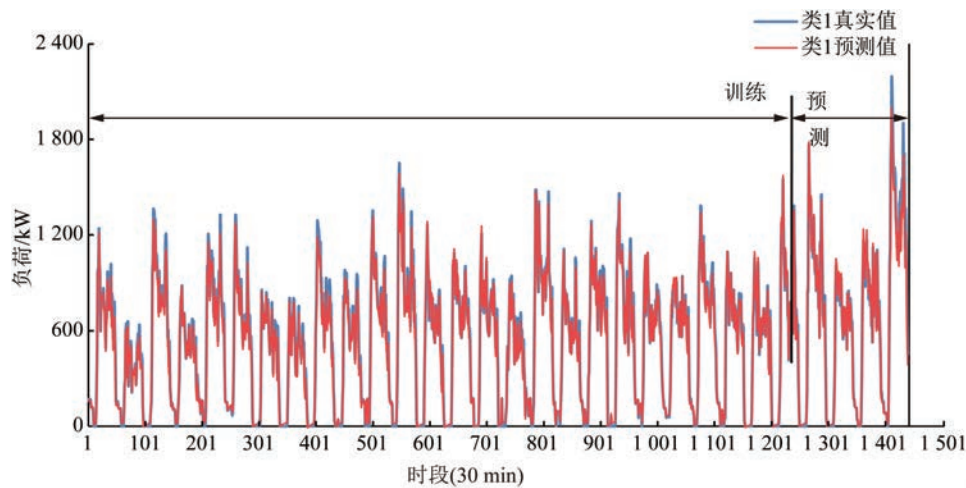


图 6 类 1 训练和预测值

Fig. 6 Training and predictive values of Class 1

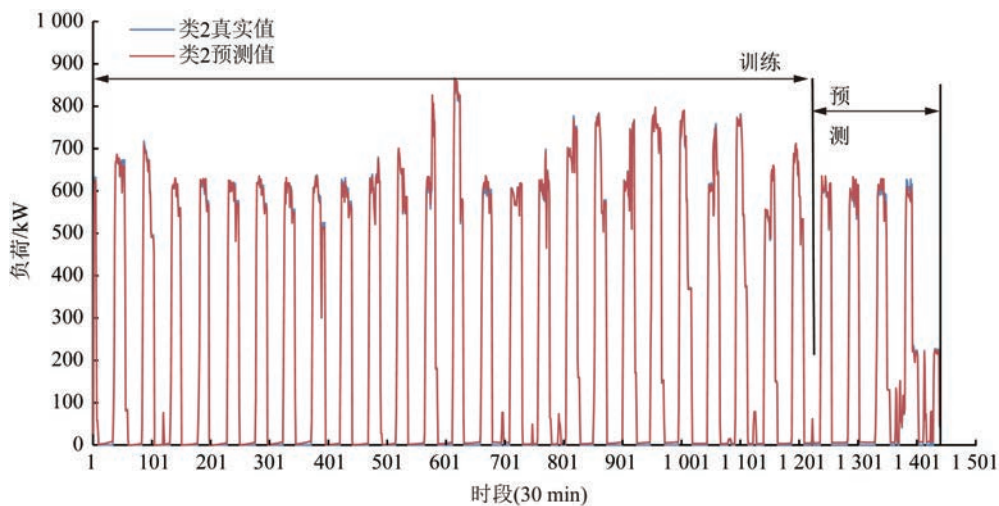


图 7 类 2 训练和预测值

Fig. 7 Training and predictive values of Class 2

增加的损失值变化,由图可知 模型训练能稳定收敛, 有较好的预测效果。

为验证所提方法的有效性,分别将本文方法与 Kmeans-LSTM 网络的预测方法和单一 LSTM 网络的预测方法做对比,结果如图 9 所示(12 月 26 日至 28 日)。其中,单一 LSTM 网络的预测方法是把所有电动公交车按相同日期累加得到日总负荷,再进行 LSTM 神经网络训练。

由图 9 可以看出,在负荷值预测上,虽然单一 LSTM 网络的预测方法结果比其他方法的预测结果在数值上比较准确,但预测值在时间上明显滞后,而本文所提方法的预测值在时间上能更好地跟踪真实值。Kmeans-LSTM 网络预测结果则在负荷值与增减趋势上都存在较多误差,但预测结果准确度 MAPE 比单 LSTM 网络预测方法有所提升,如表 2 所示。从中可以看出,相比于其他预测方法,本文提出的方法准确度较高,预测结果的 MAPE 在 11% 以下。

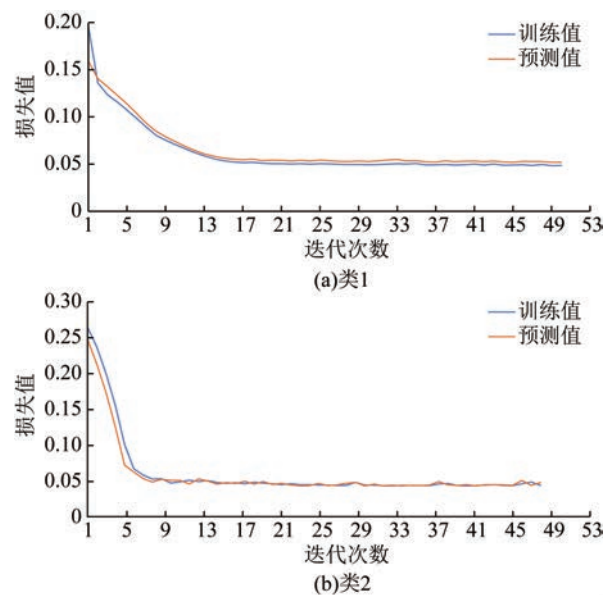


图 8 类 1 和类 2 训练时的损失值变化

Fig. 8 Changes in loss value during the training for Class 1 and Class 2

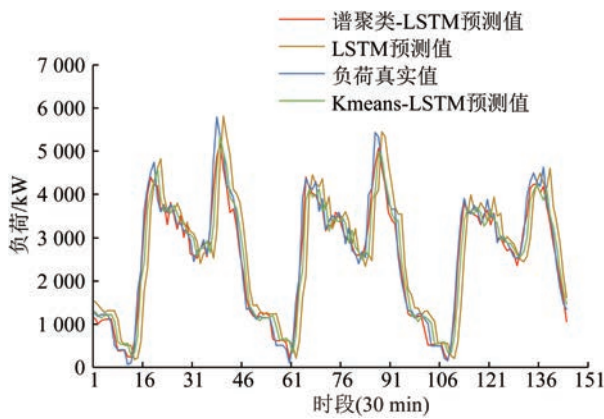


图 9 负荷预测模型对比
Fig. 9 Comparison of load forecasting model

表 2 预测模型 MAPE 误差对比

Table 2 MAPE error comparison of the prediction model

| 预测日 | MAPE/% | | |
|--------|----------|-------------|-------|
| | 谱聚类-LSTM | Kmeans-LSTM | LSTM |
| 12月26日 | 10.29 | 34.04 | 36.38 |
| 12月27日 | 7.98 | 25.94 | 33.41 |
| 12月28日 | 7.67 | 19.25 | 33.05 |

5 结 论

本文通过分析电动公交车的充电负荷特性,结合 LSTM 神经网络在时间序列上有较好的处理能力,提出了基于谱聚类和 LSTM 神经网络的电动公交车充电负荷预测方法。首先通过对数据进行简单的预处理和清洗工作,分析电动公交车日负荷曲线特征,利用谱聚类算法对电动公交车日负荷曲线进行聚类;然后针对每簇群体的充电负荷数据,利用其相应数据训练 LSTM 网络结构参数,并预测该类负荷;通过将所有 LSTM 预测结果求和得到总群体的充电负荷预测值;最后利用某市电动公交车数据,通过与其他方法进行比较,验证了所提方法的有效性。

6 参 考 文 献

[1] 罗卓伟,胡泽春,宋永华,等. 电动汽车充电负荷计算方法[J]. 电力系统自动化, 2011, 35(14): 36-42.
LUO Zhuowei, HU Zechun, SONG Yonghua, et al. Study on plug-in electric vehicles charging load calculating [J]. Automation of Electric Power Systems, 2011, 35(14): 36-42.

[2] WANG H L, ZHANG Y J, MAO H P. Load forecasting method of EVs based on time charging probability [C]//2018 International Conference on Power System Technology (POWERCON). November 6-8, 2018, Guangzhou, China. IEEE, 2018: 1731-1735.

[3] 袁正平,周伟,王文斌. 电动汽车充电负荷预测方法研究[J]. 华东电力, 2013, 41(12): 2567-2572.
YUAN Zhengping, ZHOU Wei, WANG Wenbin. Charging load

forecasting method for electric vehicles [J]. East China Electric Power, 2013, 41(12): 2567-2572.

[4] WANG Y, XIA Q, KANG C Q. Secondary forecasting based on deviation analysis for short-term load forecasting [J]. IEEE Transactions on Power Systems, 2011, 26(2): 500-507.

[5] ZHU J C, YANG Z L, CHANG Y, et al. A novel LSTM based deep learning approach for multi-time scale electric vehicles charging load prediction [C]//2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia). May 21-24, 2019, Chengdu, China. IEEE, 2019: 3531-3536.

[6] 牛东晓,马天男,王海潮,等. 基于 KPCA 和 NSGA II 优化 CNN 参数的电动汽车充电站短期负荷预测[J]. 电力建设, 2017, 38(3): 85-92.
NIU Dongxiao, MA Tiannan, WANG Haichao, et al. Short-term load forecasting of electric vehicle charging station based on KPCA and CNN parameters optimized by NSGA II [J]. Electric Power Construction, 2017, 38(3): 85-92.

[7] 王琨,高敬更,张勇红,等. 基于 LSTM 神经网络的复合变量电动汽车充电负荷预测方法技术研究[J]. 工业仪表与自动化装置, 2019(1): 27-31.
WANG Kun, GAO Jinggeng, ZHANG Yonghong, et al. Study on forecasting method of charging load of hybrid variable electric vehicle based on LSTM neural network [J]. Industrial Instrumentation & Automation, 2019(1): 27-31.

[8] 王潇笛,刘俊勇,刘友波,等. 采用自适应分段聚合近似的典型负荷曲线形态聚类算法[J]. 电力系统自动化, 2019, 43(1): 110-118.
WANG Xiaodi, LIU Junyong, LIU Youbo, et al. Shape clustering algorithm of typical load curves based on adaptive piecewise aggregate approximation [J]. Automation of Electric Power Systems, 2019, 43(1): 110-118.

[9] DONG X S, QIAN L J, HUANG L. Short-term load forecasting in smart grid: A combined CNN and K-means clustering approach [C]//2017 IEEE International Conference on Big Data and Smart Computing (BigComp). February 13-16, 2017, Jeju, Korea (South). IEEE, 2017: 119-125.

[10] 庞传军,余建明,冯长有,等. 基于 LSTM 自动编码器的电力负荷聚类建模及特性分析[J]. 电力系统自动化, 2020, 44(23): 57-63.
PANG Chuanjun, YU Jianming, FENG Changyou, et al. Clustering modeling and characteristic analysis of power load based on long-short-term-[J]. Automation of Electric Power Systems, 2020, 44(23): 57-63.

[11] SUN R Q, XIAO X Q, ZHOU F, et al. Research of power user load classification method based on K-means and FSVM [C]//2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). March 15-17, 2019, Chengdu, China. IEEE, 2019: 2138-2142.

[12] 金之榆,王毛毛,史会磊. 基于 DBSCAN 和改进 K-means 聚类算法的电力负荷聚类研究[J]. 东北电力技术, 2019, 40(6): 10-14.
JIN Zhiyu, WANG Maomao, SHI Huilei. Research on power load clustering based on DBSCAN and improved K-means clustering algorithm [J]. Northeast Electric Power Technology, 2019, 40(6):

- 10-14.
- [13] YU T M, YANG J H, LU W. Dynamic background subtraction using histograms based on fuzzy C-means clustering and fuzzy nearness degree [J]. IEEE Access, 2019, 7: 14671-14679.
- [14] DINESH C, MAKONIN S, BAJC I V. Residential power forecasting based on affinity aggregation spectral clustering [J]. IEEE Access, 2020, 8: 99431-99444.
- [15] 艾欣, 杨子豪, 胡寰宇, 等. 基于改进 k-means 算法的 VPP 负荷曲线聚类方法及应用 [J]. 电力建设, 2020, 41(5): 28-36.
AI Xin, YANG Zihao, HU Huanyu, et al. A load curve clustering method based on improved K-means algorithm for virtual power plant and its application [J]. Electric Power Construction, 2020, 41(5): 28-36.
- [16] 林顺富, 田二伟, 符杨, 等. 基于信息熵分段聚合近似和谱聚类的负荷分类方法 [J]. 中国电机工程学报, 2017, 37(8): 2242-2253.
LIN Shunfu, TIAN Erwei, FU Yang, et al. Power load classification method based on information entropy piecewise aggregate approximation and spectral clustering [J]. Proceedings of the CSEE, 2017, 37(8): 2242-2253.
- [17] 端祝超. 基于数据简化拟合的电动公交车充电负荷预测 [J]. 电工电气, 2020(3): 23-27.
DUAN Zhuchao. Electric bus charging load forecast based on data simplification fitting [J]. Electrotechnics Electric, 2020(3): 23-27.
- [18] 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 [J]. 电力系统自动化, 2019, 43(8): 131-137.
LU Jixiang, ZHANG Qipei, YANG Zhihong, et al. Short-term load forecasting method based on CNN-LSTM hybrid neural network model [J]. Automation of Electric Power Systems, 2019, 43(8): 131-137.
- [19] SUN Y P, GAO Y J, YANG W H, et al. Research on user optimal aggregation based on demand response potential spectrum clustering analysis [J]. The Journal of Engineering, 2017, 2017(13): 2152-2157.
- [20] 彭勃, 张逸, 熊军, 等. 结合负荷形态指标的电力负荷曲线两步聚类算法 [J]. 电力建设, 2016, 37(6): 96-102.
PENG Bo, ZHANG Yi, XIONG Jun, et al. A two-step clustering algorithm combined with load shape index for power load curve [J]. Electric Power Construction, 2016, 37(6): 96-102.
- [21] 杨甲甲, 刘国龙, 赵俊华, 等. 采用长短期记忆深度学习模型的工业负荷短期预测方法 [J]. 电力建设, 2018, 39(10): 20-27.
YANG Jiajia, LIU Guolong, ZHAO Junhua, et al. A long short term memory based deep learning method for industrial load forecasting [J]. Electric Power Construction, 2018, 39(10): 20-27.
- [22] 陈吕鹏, 殷林飞, 余涛, 等. 基于深度森林算法的电力系统短期负荷预测 [J]. 电力建设, 2018, 39(11): 42-50.
CHEN Lypeng, YIN Linfei, YU Tao, et al. Short-term power load forecasting based on deep forest algorithm [J]. Electric Power Construction, 2018, 39(11): 42-50.
- [23] MUZAFFAR S, AFSHARI A. Short-term load forecasts using LSTM networks [J]. Energy Procedia, 2019, 158: 2922-2927.
- [24] 中国气象数据网. 中国地面国际交换站气候资料日值数据集 (V3.0). 2019 [2020-12-30]. <https://data.cma.cn>.
-
- 收稿日期: 2020-10-20
- 作者简介:
王哲(1984) 男, 硕士, 高级工程师, 主要研究方向为配电网规划与电网智能化;
万宝(1988) 男, 学士, 工程师, 主要研究方向为配电网规划;
凌天晗(1989) 男, 学士, 工程师, 主要研究方向为配电网规划;
董晓红(1989) 女, 博士研究生, 主要研究方向为电动汽车与电网互动控制, 配电网规划;
穆云飞(1984) 男, 博士, 教授, 主要研究方向为电力系统安全与稳定、综合能源集成与应用;
邓友均(1990) 男, 博士研究生, 主要研究方向为电动汽车充放电管理、配电网规划;
唐舒懿(1996) 女, 博士研究生, 主要研究方向为电动汽车负荷预测及电动汽车有序充电控制、配电网规划。
- (编辑 白婉欣)