

doi:10.11835/j.issn.1000.582X.2023.09.013

城市交通网格集群的 Bi-LSTM 的流量预测

贾现广^a, 冯超琴^a, 苏治文^a, 钱正富^a, 宋腾飞^a, 刘欢^a, 吕英英^b

(昆明理工大学 a. 交通工程学院; b. 信息工程与自动化学院, 昆明 650500)

摘要:为提升交通流预测精度,深层次地挖掘交通流数据的时空特征,提出一种基于双向长短时记忆网络(Bi-LSTM)的城市交通网格集群流量预测模型。将所获得的网约车轨迹数据集网格化;考虑人为确定集群个数对结果的影响,用贝叶斯信息准则进行参数估计确定集群数,利用高斯混合模型对交通状况相似的网格进行聚类得到不同交通网格集群;利用集群内部交通网络的输入时间序列的相互影响设计多对多模型,构建 Bi-LSTM 模型预测不相交集群上的流量;以经典多元线性回归模型(MLRA)作为对照组进行实验验证,采用平均绝对误差(MAE)、均方根误差(RMSE)、平均绝对百分比误差(MAPE)和动态时间规整(DTW)这四类评价指标对预测结果进行综合评价,验证基于 Bi-LSTM 模型的城市交通网格集群流量预测的可行性。实验结果表明:MLRA 模型和 Bi-LSTM 模型对城市交通网格集群流量的预测值小于真实值,早高峰时段尤为明显;各交通网格集群的交通状态态势相似,集群的簇内相关性较强,两类模型均可实现较好的流量预测效果,Bi-LSTM 表现更优;MLRA 和 Bi-LSTM 预测模型的精度 MAE、RMSE、MAPE 分别为 3.2011、4.4009、0.3187、3.0687、4.2943、0.3045,Bi-LSTM 与 MLRA 相比,模型精度分别提高了 4.14%、2.40%、4.46%,说明所构建的 Bi-LSTM 交通流网格集群流量预测精度高、误差低,要优于 MLRA 模型,表现出较好的泛化性能;MLRA 和 Bi-LSTM 的 DTW 结果分别为 52938.6356、54815.1055,构建的 Bi-LSTM 模型较 MLRA 模型各自工作日和节假日时间序列相似性 DTW 结果提高 3.42%,表现出更好的鲁棒性。利用城市交通流量的特点和交通轨迹数据网格化的优点,基于 Bi-LSTM 模型的城市交通网格集群流量预测与 MLRA 交通流量预测模型相比,具有精度高、误差低的特点。同时,DTW 指标方面,基于 Bi-LSTM 对城市交通网格集群流量模型与真实流量变化趋势一致,表现出较好的鲁棒性。

关键词:城市交通;交通流量预测;Bi-LSTM;交通网格集群;时空单元划分

中图分类号:U491.2

文献标志码:A

文章编号:1000-582X(2023)09-130-12

Forecasting for urban traffic grid clusters based on Bi-LSTM

JIA Xianguang^a, FENG Chaoqin^a, SU Zhiwen^a, QIAN Zhengfu^a, SONG Tengfei^a,
LIU Huan^a, LYU Yingying^b

(a. Faculty of Transportation Engineering; b. Faculty of Information Engineering and Automation,
Kunming University of Science and Technology, Kunming 650500, P. R. China)

Abstract: This study aims to improve the accuracy of traffic prediction and to explore the spatio-temporal characteristics of traffic data in urban areas by proposing a traffic flow prediction model based on Bi-LSTM

收稿日期:2022-11-09 网络出版日期:2023-07-06

基金项目:国家自然科学基金资助项目(71961012)。

Supported by National Natural Science Foundation of China(71961012).

作者简介:贾现广(1997—),男,副教授,主要从事智能交通与大数据方向研究,(E-mail)jxg@kust.edu.cn。

通信作者:吕英英,女,主要从事智能交通与大数据方向研究,(E-mail)20070102@kust.edu.cn。

(Bidirectional Long Short Term Memory) for urban traffic grid clusters. The trajectory dataset of ride-hailing vehicles was gridded, and the Bayesian information criterion was used for parameter estimation to determine the cluster number, with considering the influence of manually determining the number of clusters on the results. The Gaussian mixture model was then employed to cluster grids with similar traffic conditions, resulting in distinct traffic grid clusters. A Multi-to-Multi model was designed by considering the mutual influence of input time series of traffic grids within each cluster. The Bi-LSTM model was established to predict traffic flow in non-overlapping clusters. Experimental validation was conducted using the classical MLRA (multiple linear regression analysis) as a control group, and four performance metrics: MAE(mean absolute error), RMSE(mean squared root error), MAPE(mean absolute percentage error) and DTW (dynamic time warping) were used to comprehensively evaluate the prediction results, confirming the feasibility and superiority of the Bi-LSTM model for city traffic grid cluster flow prediction. The results showed that both MLRA and Bi-LSTM models predicted urban traffic grid cluster traffic values were generally smaller than the real value, with more pronounced discrepancies observed during morning peak hours. Increasing data volume improved the prediction performance of the models. Traffic state dynamics within each traffic grid cluster were similar, displaying strong intra-cluster correlation. Both models achieved better traffic prediction results, with Bi-LSTM outperforming MLRA. In terms of model accuracy, the Bi-LSTM model showed improved MAE, RMSE and MAPE(3.068 7, 4.294 3, 0.304 5, respectively) compared to MLRA(3.201 1, 4.400 9, 0.318 7, respectively), representing a 4.14%, 2.40% and 4.46% enhancement, respectively. The constructed Bi-LSTM model exhibited higher accuracy, lower error and better generalization performance. In terms of similarity result evaluation, the DTW results of MLRA and Bi-LSTM were 52 938.635 6 and 54 815.105 5 respectively. The Bi-LSTM model showed better robustness by 3.42% compared to the respective weekday and holiday time series similarity DTW results of the MLRA model. By considering the characteristics of urban traffic flow and leveraging the advantages of gridding traffic trajectory data, the Bi-LSTM-based model for urban traffic grid cluster traffic prediction exhibited high accuracy, low error and superior robustness compared to the MLRA traffic flow prediction model. Meanwhile, in terms of DTW metrics, the Bi-LSTM-based urban traffic grid cluster traffic model captured the real traffic variation trend and demonstrated excellent performance in traffic flow prediction for urban areas.

Keywords: urban transportation; traffic flow forecasting; Bi-LSTM; traffic grid clusters; spatiotemporal unit division

智能交通系统(intelligent transport system, ITS)是智慧城市的重要组成部分,可有效解决交通拥堵问题,城市交通流量预测是ITS的基石,精准高效的交通流量预测对缓解交通拥堵具有重要意义^[1]。城市交通流量预测是城市交通管理不可或缺的重要手段,准确的交通流量预测可为智能城市规划和交通管理提供可靠管理,为拥挤区域提供预警确保公共安全,为乘客提供建议方便日常通勤^[2]。

早期交通流量预测是基于统计学模型的方法。Vythoukas^[3]采用卡尔曼滤波模型利用线性系统状态方程对交通流量进行预测,Ahmed等^[4]将流量预测转换成时序预测提出ARIMA方法(Auto-Regressive integrated moving average)对高速路交通流量进行预测。但基于统计模型的预测方法大多为线性模型,随着交通数据的非线性和不确定性增强,模型预测精度逐渐变差,学者开始利用机器学习对交通流量预测。Feng等^[5]提出具有时空相关性的自适应多核SVM模型(support vector machine)用于短期交通流量预测, Lee等^[6]提出基于天气状况的多元线性回归分析模型(multiple linear regression analysis, MLRA)用于交通预测。随着交通数据的数据量剧增,交通流量数据也呈现出多维度的趋势,机器学习对于复杂庞大的连续数据深层次特征的提取日益困难,研究者开始运用深度学习进行交通流量预测。

Zhang等^[7]对交通数据进行时空分析,引入CNN(convolutional neural network)进行交通流量预测,单一模

型预测难免会使数据特征信息丢失,将神经网络拓展延伸形成组合预测方法,逐渐成为发展趋势。Fu等^[8]使用LSTM(long short term memory)-GRU(gate recurrent unit)组合模型对交通流量进行预测,Zheng等^[9]将注意力机制引入图神经网络用于交通流量预测。基于神经网络创新取得较好预测效果,但模型复杂度增加且效率变低。为挖掘深层次特征,Wang等^[10]构建动态感知系统收集数据,对交通道路数据进行AP(affinity propagation)聚类,利用LSTM进行交通流量预测^[11],Chiabaut等^[12]对高速道路数据进行高斯混合模型(gaussian mixture model,GMM)聚类得到拥堵图,运用共识学习技术进行交通预测。研究均是基于道路聚类,聚类前并未对道路数据进行其它处理以得到其隐含的时空特征。

针对交通流量数据的非线性和不确定性增加、复杂庞大的连续数据特征难以提取、交通轨迹数据的隐含特征提取不充分问题,笔者提出基于深度学习的交通流量预测方法,引入双向长短时记忆网络模型(bidirectional long short-term memory network, Bi-LSTM),贡献如下:

1) 网格化处理和交通网络集群建模:通过将交通轨迹数据进行网格化处理,将复杂庞大的连续数据转化为网格化的交通网络集群。利用贝叶斯信息准则(bayesian information criterion, BIC)和GMM模型,将网格化的交通流量数据序列分解为若干个加速度、速度比率、平均停车次数等相似交通状况的网络集群,有效提取交通网络集群的时空特征,深入挖掘路网的时空关联性。

2) 设计多对多模型进行序列预测:全面考虑交通网络之间的相互影响和时序关系,设计了多对多模型,将分解得到的交通网络集群作为Bi-LSTM的输入,实现对网络集群交通流量的序列预测。

3) 通过使用滴滴盖亚数据开放计划的数据集进行实验,实验结果表明,基于Bi-LSTM的城市交通网络集群流量预测具有良好的预测能力,且优于基准模型。

1 城市交通网络集群的Bi-LSTM流量预测模型

基于多对多模型设计思想,构建Bi-LSTM预测模型对城市交通流网络集群流量进行预测。考虑输入的交通网络集群序列 $C = \{c_1, c_2, \dots, c_k\}$ 遵循正态分布,为便于计算,对各个集群进行标准化处理;交通网络集群可用的时序数据可确定2个网格之间的时间和空间相关性,使所输入数据具有较强的时空关联,利用多对多模型设计将分解的若干个交通网络集群作为Bi-LSTM的输入;根据预测精度调整Bi-LSTM模型相关参数权重,以此预测 n 时刻该交通网络集群每个网格未来的交通流量。基于Bi-LSTM对城市交通网络集群流量模型预测框架如图1所示。

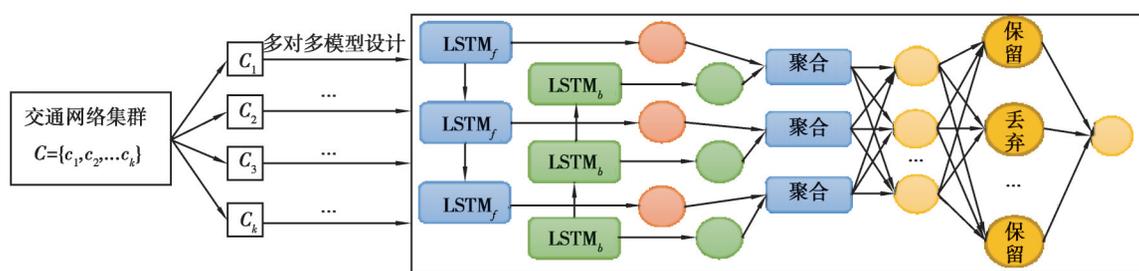


图1 基于Bi-LSTM对城市交通网络集群流量模型预测框架

Fig. 1 Prediction model framework for urban traffic rasters cluster flow based on Bi-LSTM

1.1 多对多模型设计

1.1.1 交通传播

交通传播模式是在一个道路网络中特定路段的流量水平所拥有的时空关系^[13],在空间上的周期性和时间上的相似性可从历史数据中获得有价值信息,进而挖掘深层次时空特征。交通流量在不同交通网络中的传播随时间和空间动态变化,图2是某集群某日早高峰时段各空间网络的交通流量图,横纵坐标表示轨迹数据网格化后的空间索引值,从图2可以看出交通网络(6,34)的交通流量大,周围其他交通网络的流量越靠近这个网络流量影响越大,说明这个交通网络影响着周围其它交通网络的流量大小。

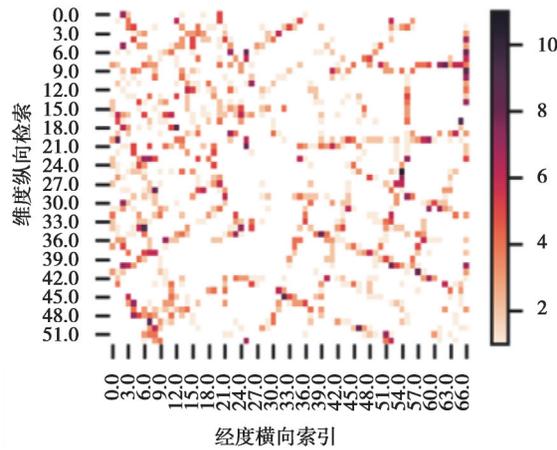


图 2 网格流量图

Fig. 2 Grid flow diagram

1.1.2 多对多模型设计

交通网络可用的时序数据可确定 2 个网格之间的时空相关性,得到 t 时刻当前网络的交通流量对其它网格影响。考虑到多对多模型对整个交通网络的交通流量预测效率更高^[11],为挖掘集群内各网格的时空特征,进行多对多模型设计。多对多模型设计思路图如图 3 所示,左图表示将网格化的交通数据进行聚类后所得到的某集群,为充分利用每个网格的时序数据将某集群 $t = 1, 2 \dots n-1$ 个时间序列数据作为 Bi-LSTM 模型的输入,预测该集群的每个网格在 $t=n$ 时刻的流量。多对多模型设计是在被网格化的交通网络中,读取网格的所有集群,将相似交通网络的聚类数据输入 Bi-LSTM,预测 n 时刻每个网格未来的交通流量。

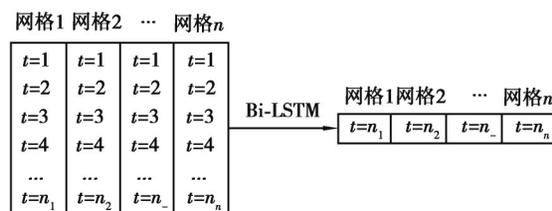


图 3 多对多模型设计

Fig. 3 Multi-to-Multi model design

1.2 Bi-LSTM 模型框架

交通流量当前状态不仅与之前的状态有关,还与未来的状态有关。Bi-LSTM 具有较强的序列建模能力,可以对交通数据的时间序列特征进行建模和预测,包括小时、天、周等不同时间尺度。Bi-LSTM 其结构由上下堆叠的 2 个单向 LSTM 组成,输入在 2 个方向流动^[14],输入包含预测时间前后的时间序列,可以同时考虑历史时刻和未来时刻的信息,通过结合前后时刻的信息可更好地捕获交通数据的动态变化,较好地预测流量。在交通流量预测中,其模型的结构如图 4 所示。该网络克服了 RNN(recurrent neural network, RNN)衰退问题,利用了 LSTM 可长时间储存记忆的优点。

求解时分别计算前项隐向量 \vec{h}_{t-1} 产生新的隐向量 \vec{h}_t , 如式(2),后项隐向量 \overleftarrow{h}_{t-1} 产生新的隐向量 \overleftarrow{h}_t , 如式(3),将正反 2 项输入序列的输出结果结合,得到 Y_t , 如式(4)。

$$h_t = o_t \cdot \tanh(C_t), \tag{1}$$

$$\vec{h}_t = \text{LSTM}(x_t, \vec{h}_{t-1}), \tag{2}$$

$$\overleftarrow{h}_t = \text{LSTM}(x_t, \overleftarrow{h}_{t-1}), \tag{3}$$

$$Y_t = \tanh(W_{hy} \vec{h}_t + W_{hy} \overleftarrow{h}_t + b_y), \tag{4}$$

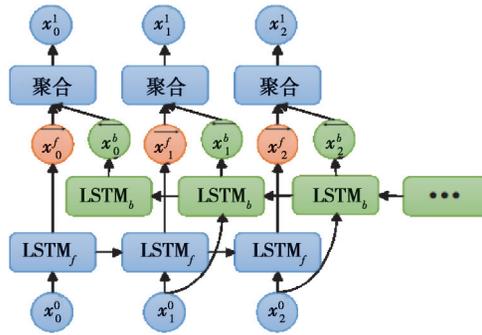


图4 Bi-LSTM结构图

Fig. 4 Bi-LSTM structure diagram

式中: \vec{h}_t 和 \overleftarrow{h}_t 分别为 t 时刻流量序列前向传播层和后向传播层中隐藏层的输出; W_{hy} 和 $W_{\overleftarrow{h}y}$ 表示每一层连接到前一个隐藏状态的权重矩阵, b_y 表示偏置项。

2 数据说明

研究的数据集来源于滴滴出行设立的盖亚数据开放计划平台,地理区域是四川省成都市二环内局部范围,原始数据大小约 180 GB,轨迹点采集间隔为 2-4 s,时间区间是某年 11 月 1 日-30 日,轨迹数据与实际路段相对应。司机及乘客订单的隐私信息进行了加密脱敏匿名化处理可保证用户安全。为了使轨迹数据更加集中,便于提取时空特征,对大规模杂乱轨迹数据进行清洗、错误筛查、压缩等处理,处理前后区域范围如图 5 所示,(a)表示处理前区域范围,(b)表示处理后区域范围。从图看出处理后的区域范围更加清晰,数据更加集中,便于进行时空网格划分。



图5 处理前后区域范围

Fig. 5 Range of areas before and after treatment

大规模轨迹数据时空相关性较强,其时间数据只有一维属性且具有严格的序列性,空间数据具有多维属性且空间相关性更复杂。网格划分可高效地挖掘数据中隐含的时空特征,对不同类别的时空数据进行融合^[15]。笔者利用网格划分方法将研究区域划分为棋盘状的网格以便于提取深层次的交通特征,时空网格划分如图 6 所示,左图中若干个黑色圆圈表示网约车轨迹数据点的分布情况,红色箭头表示将其归类到各个空间网格的过程,右图表示划分时间网格。取空间网格大小为 50,时间窗为 600 s,将空间数据点映射到各个网格中,在对定位点进行映射的同时也将矢量的轨迹网格化为网格序列,实现将轨迹数据打散成点坐标转化为张量,从而可直接或间接得到每个网格的索引值以及网格平均加速度、网格浮动车流量、网格平均停车次数等。



图 6 时空网格划分

Fig. 6 Spatiotemporal gridding

3 交通网格集群模型构建

3.1 确定最佳组分

基于Bi-LSTM的城市交通网格集群流量进行预测,确定交通网格的最佳集群数是进行GMM模型聚类的基础。为避免人为确定最佳集群数对实验结果产生影响,利用BIC进行参数估计确定交通网格最优集群数 k 。

很多参数估计问题采用似然函数作为目标函数,当训练数据集足够大时,该方法可不断提高模型的精度,但也会随着模型复杂度增加而出现过拟合问题^[16]。在模型选择上,在模型复杂度与模型对数据集描述能力之间寻求最佳平衡,Schwarz提出BIC^[17]信息准则,避免样本数量过多导致模型精度过高造成模型复杂度过高的问题。该准则如式(5)所示,通过在模型中加入与模型参数个数相关的惩罚项 $k \ln(n)$ 来避免模型的过拟合。

$$BIC = k \ln(n) - 2 \ln(L), \tag{5}$$

式中: k 为模型参数的个数; n 为样本数量; L 为似然函数; $k \ln(n)$ 惩罚项在维数过大且训练样本数据相对较少的情况下,可有效避免维度灾难现象。

BIC倾向于选择较小的值,而模型参数的个数 k 过大会导致模型出现过拟合^[17]。依据样本数据,GMM模型参数个数与BIC值关系如图7所示。由图7可知BIC的值在4出现了大转折,7以后虽然BIC的值一直在下降,但下降变缓。同时, k 值过大会使BIC值变化不明显且极易导致模型出现过拟合现象,为使BIC值和 k 值都相对小,取 k 值为6。

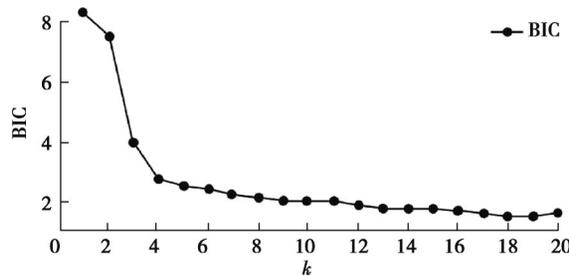


图 7 模型参数个数与 BIC 值关系图

Fig. 7 Number of GMM model parameters versus BIC value

3.2 构建 GMM 模型

某种程度上 GMM 模型是改进型的 K-means 模型。K-means 模型的作用是将数据收集到方差相等的组中,使组内总和最小^[18],与其他聚类相比收敛速度快,算法可解释度较强,但 K-means 存在集群形状缺乏灵活性和概率集群分配概念的问题,无法将 2 个聚类中心相同的类聚类。GMM 模型是多个高斯概率密度函数的线性组合,其曲线由若干个单高斯函数叠加而成,任何一个复杂曲线都可用若干个高斯曲线无限逼近。

GMM 聚类方法重视数据点分布而不仅是数据点之间的距离,它是通过选择成分最大化后验概率来完成聚类,提供的聚类有不同的大小^[19]。基于构成不同聚类的交通网格集群遵循正态分布,故此选择 GMM 算法。GMM 聚类模型如式(6)所示。

$$p_M(x) = \sum_{i=1}^k \alpha_i \cdot p(x | \mu_i, \Sigma_i), \quad (6)$$

式中: $p(x | \mu_i, \Sigma_i)$ 为概率密度函数; k 表示 k 个混合成分组成,每个混合成分对应一个高斯分布; μ_i 与 Σ_i 是第 i 个高斯混合成分参数; $\alpha_i > 0$ 为相应的“混合系数”, $\sum_{i=1}^k \alpha_i = 1$ 。

利用 K-means 迭代获取最优值后将聚类中心点作为 GMM 的初始值进行训练,以避免 GMM 收敛到局部最优解;在 GMM 聚类时,利用 EM 算法估算参数使每个分量都有自己的通用协方差矩阵。图 8 为交通网格聚类的示意图,网格 id 和时间 id 表示将交通数据网格化后所得到的时空索引值,网格 id 和时间 id 构成时空网格,clusters 表示集群的类别,其中(48,3),3, C_2 和(19,15),92, C_2 属于同一集群,构成了交通网格集群。

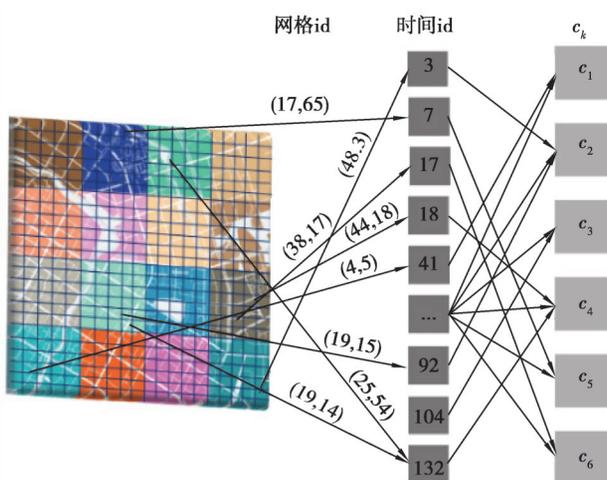


图 8 交通网格聚类示意图

Fig. 8 traffic condition grid clustering diagram

4 模型验证

4.1 模型训练

使用 1~24 日的数据作为训练数据,25~30 日的数据作为测试数据,取训练集的 15% 作为验证集,利用过去 1 h 的交通流量数据(6 个时间步长)来预测未来 10 min(1 个时间步长)的交通流量。采用 Python 语言进行编程,该程序使用 Keras 库,在 AMD Ryzen 7 5800H with Radeon Graphics 处理器、16gb 内存、Window10 操作系统和 Tensorflow2.0 深度学习框架上实现。经过 GMM 聚类后所生成的交通网格集群利用多对多模型输入 Bi-LSTM 模型进行训练。Bi-LSTM 设计有 6 个隐藏层,前两层是 bi-lstm 层,各有 64 个神经元;第三层有 32 个神经元;第四层 16 个神经元;第五层为 dropout 层,为防止模型过拟合,丢弃率设为 0.2;最后一层为全连接层。LSTM 模型学习率为 0.001,损失函数使用 MSE,反向传播算法使用 Adam 优化器,激活函数采用 ReLU。由于不同集群的交通状况不一致,故每批次所训练的样本数(batch sizes)、总迭代次数(epochs)和训练时间(times)如表 1 所示。为对模型预测效果进行评价,使用经典的多元线性回归分析(MLRA)模型作为对照组,在同一实验环境中进行对比实验验证。

表 1 不同集群的参数大小

Table 1 Parameter sizes of c_k

| c_k | c_1 | c_2 | c_3 | c_4 | c_5 | c_6 |
|------------|-------|-------|-------|-------|-------|-------|
| 批次数 | 64 | 64 | 256 | 256 | 64 | 16.0 |
| 迭代次数 | 98 | 36 | 179 | 71 | 68 | 39.0 |
| 训练时间 t/s | 522 | 291 | 1030 | 727 | 152 | 77.6 |

4.2 模型误差分析

利用过去 1 h 的交通数据预测未来 10 min 的交通流量,图 9 为某交通网格集群 MLRA 模型和 Bi-LSTM 各自相同工作日早高峰时段网格流量的预测结果图,其中 a 为真实值, b 为预测值,右下角为同一区域网格的放大图。

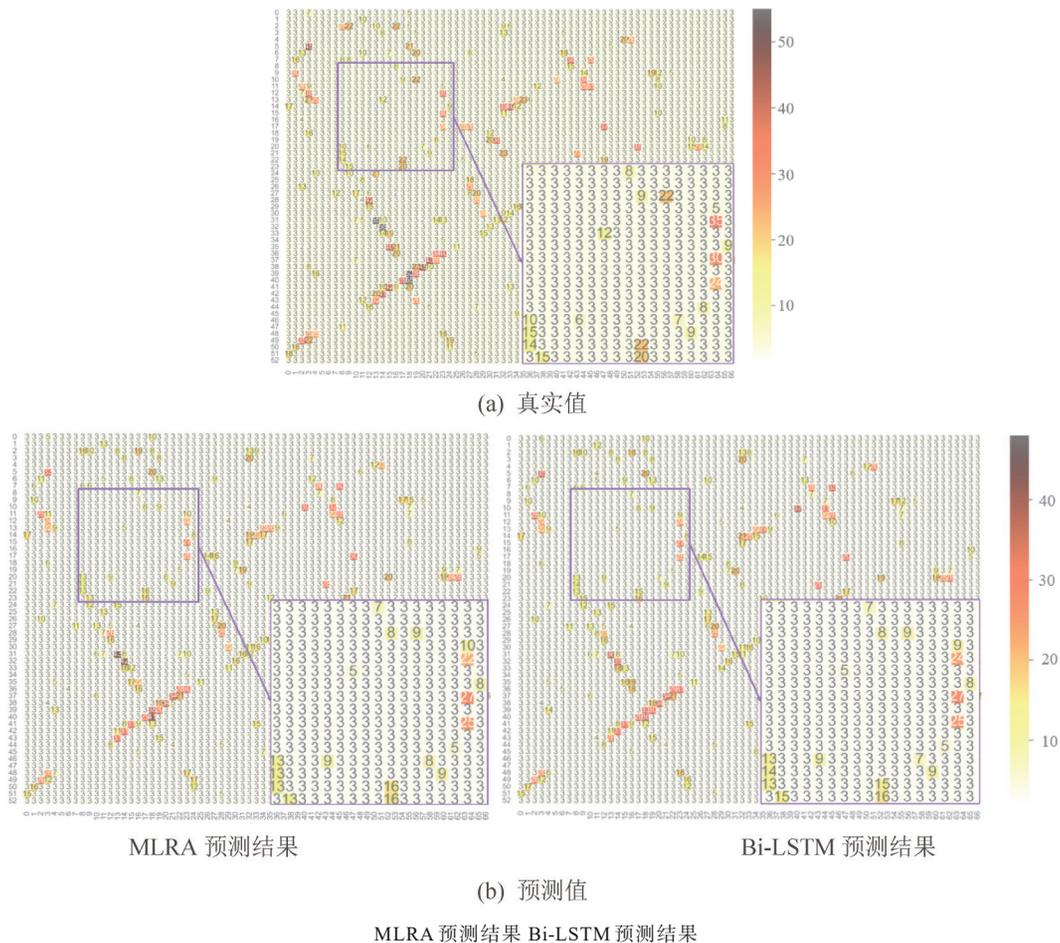


图 9 模型预测结果对比
Fig. 9 Comparison of prediction results of models

从图 9 可以看出:1)总体来看 MLRA 模型和 Bi-LSTM 模型预测结果均小于真实值,可能与早高峰时段的交通流量大、交通状况复杂等因素有关,使 2 个模型的预测值都偏小;2)从局部放大网格图直观上发现每个网格流量趋势大体相同,但 MLRA 模型的流量预测值较真实值波动大,不稳定, Bi-LSTM 模型的预测精度高于 MLRA 模型。

模型时间序列相似性结果:图 10、图 11 表示的是某交通网格集群 MLRA 模型和 Bi-LSTM 模型在工作日和节假日早高峰和平峰时段时间序列相似性的结果,其中 a 为早高峰(7:00-9:00)、 b 为平峰(9:00-11:00),纵坐标表示流量,横坐标表示时间步长,一个时间步长代表 10 min,时间窗取 600 s, 42 代表早上 7:00, 48 代表 8:00,以此类推。

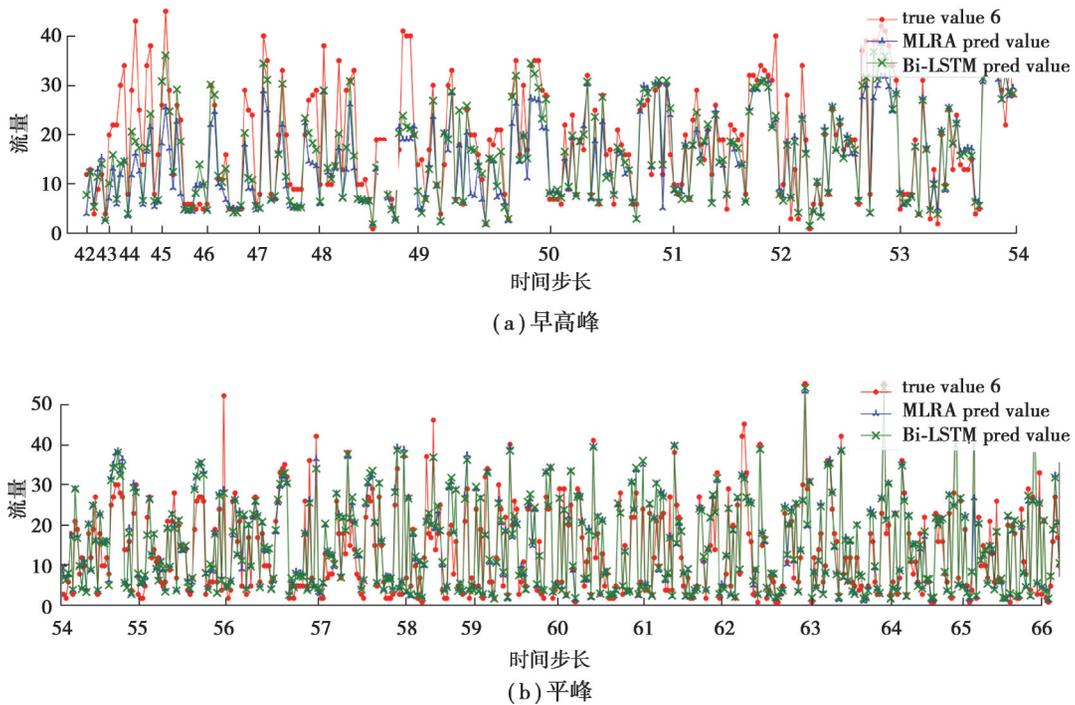


图 10 工作日时间序列相似性结果

Fig. 10 Time series similarity results(working day)

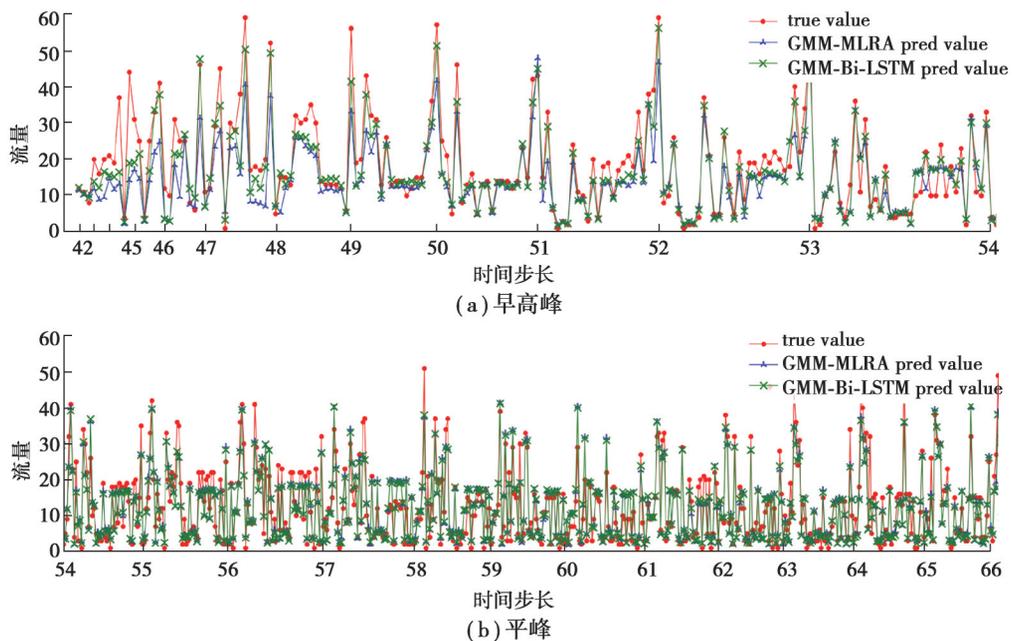


图 11 节假日时间序列相似性结果

Fig. 11 Time series similarity results(holiday)

从图 10、图 11 可以看出:1)整体来看,MLRA 模型和 Bi-LSTM 模型对于工作日和节假日的时间序列相似性结果,预测值与实际交通流量值的流动形态非常接近,说明 2 个模型能够进行较好预测。2 个模型的预测值大多时间小于真实值,早高峰时段更为明显,且 2 个模型对早高峰的预测没有平峰的预测好,可能与早高峰时段流量大、交通状况复杂等因素有关。2)在工作日早高峰时段,MLRA 模型和 Bi-LSTM 模型流量预测值大多小于真实值。在 7:00-7:30 时间段和 8:00-8:10 时间段 MLRA 模型和 Bi-LSTM 模型预测值严重偏离真

实值,但Bi-LSTM更加接近真实值,其它时间段2个模型的预测值直观上与MLRA模型更为接近,预测效果好;3)在节假日7:10-7:30时间段,MLRA模型时间序列相似性结果的预测值和实际值流动状态相差较大,随着数据增加预测值和真实值逐渐接近,表现出较好预测效果。在多数时段Bi-LSTM模型优于MLRA模型;4)对于平峰时段的预测,MLRA模型和Bi-LSTM模型流量预测值较为接近真实值,可能与平峰时段交通状况稳定有关。Bi-LSTM模型在工作日和节假日上午时段流量的预测值和真实值更加接近也更加稳定,表现出更好的鲁棒性。

为量化模型预测结果,采用经典的误差评价指标^[20]平均绝对误差(mean absolute error, MAE)、均方根误差(mean squared root error, RMSE)、平均绝对百分比误差(mean absolute percentage error, MAPE)和动态时间规整(dynamic time warping, DTW)对模型结果进行评价。

1) MAE:用于评估交通流量预测值和真实值的接近程度,其值越小拟合效果越好,

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (7)$$

2) RMSE:该指标计算的是交通流量预测结果和真实值对应样本点误差的平方根,其值越小说明预测速度值精度越高,

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}} \quad (8)$$

3)MAPE:用于测量测试集上平均流量测试值与实际测试值之间的相对误差,范围 $[0, +\infty)$,MAPE为0%表示完美模型,MAPE大于100%则表示劣质模型,

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - x_i}{y_i} \right| \quad (9)$$

式中: y_i 为预测值; x_i 为实际值; \bar{y} 为平均值。

4) DTW:用于测量交通流量预测值和真实值2个时间序列之间的相似性,所得到的距离越小,说明2个时间序列越相似。

$$\begin{aligned} \text{dist} &= \text{dtw}(x, y) \\ [\text{dist}, ix, iy] &= \text{dtw}(x, y) \\ [\quad] &= \text{dtw}(x, y, \text{maxsamp}), \\ [\quad] &= \text{dtw}(\quad, \text{metric}) \\ &\text{dtw}(\quad) \end{aligned} \quad (10)$$

$\text{dist} = \text{dtw}(x, y)$ 将2个流量预测值 x 和真实值 y 拉伸到一个共同的时刻集合上,使得对应点之间的欧几里得距离之和 dist 最小。 $[\text{dist}, ix, iy] = \text{dtw}(x, y)$ 返回共同翘曲路径,使得流量预测值 $x(ix)$ 和真实值 $y(iy)$ 之间距离尽可能小。 $[\quad] = \text{dtw}(x, y, \text{maxsamp})$ 限制扭曲路径在流量预测值 x 和真实值 y 之间拟合的最大样本范围内。 $[\quad] = \text{dtw}(\quad, \text{metric})$ 指定除前面语法中的任何输入参数之外要使用的距离度量。 $\text{dtw}(\quad)$ 交通流量预测值和真实值2个时间序列之间的相似性结果。

不同模型和集群的评价结果如表2所示,整体来看,在每个交通网格集群中,MLRA模型、Bi-LSTM模型对交通流量的预测都取得了不错效果,这与将网格化的交通流量数据通过GMM聚类所得到的一系列交通网格集群有关。因为每个交通网格集群内部相关性较强,交通状况极为相似,从而使2个模型的预测效果较好,但各自在模型精度以及时间序列相似性方面存在些许差异。为使MLRA模型、Bi-LSTM模型评价更直观,利用不同交通网格集群所占整个交通网络的权重方法,使每个交通网格集群乘以各自的权重再进行叠加,得到如表3所示总的模型评价结果。

模型精度:模型的平均评价结果如表3所示,通过表3发现,MLRA、Bi-LSTM的MAE、MSE、MAPE分别为3.201 1、4.400 9、0.318 7、3.068 7、4.294 3、0.304 5。

1)Bi-LSTM与MLRA相比,MAE降低0.1324,提高4.14%,Bi-LSTM的流量预测值较为接近真实值,拟合效果较好;

表2 不同模型和集群的评价结果

Table 2 Evaluation results of different models and c_k

| c_k | 评价指标 | MLRA | Bi-LSTM |
|-------|------|---------------|--------------|
| c_1 | MAE | 3.085 2 | 2.988 7 |
| | RMSE | 4.802 | 4.725 1 |
| | MAPE | 0.140 5 | 0.138 7 |
| | DTW | 20 407.815 1 | 19 811.135 6 |
| c_2 | MAE | 2.285 3 | 2.326 8 |
| | RMSE | 3.027 5 | 3.075 4 |
| | MAPE | 0.415 5 | 0.432 8 |
| | DTW | 30 479.499 3 | 30 473.800 1 |
| c_3 | MAE | 3.119 5 | 3.273 9 |
| | RMSE | 4.320 9 | 4.493 6 |
| | MAPE | 0.370 4 | 0.389 7 |
| | DTW | 75 837.802 4 | 72 737.685 4 |
| c_4 | MAE | 3.753 9 | 3.885 5 |
| | RMSE | 5.485 6 | 5.526 7 |
| | MAPE | 0.141 6 | 0.143 4 |
| | DTW | 103 322.079 8 | 99 625.153 7 |
| c_5 | MAE | 3.268 | 3.464 7 |
| | RMSE | 4.235 | 4.338 2 |
| | MAPE | 0.261 | 0.275 6 |
| | DTW | 5 367.451 5 | 5 343.821 6 |
| c_6 | MAE | 2.481 6 | 2.678 5 |
| | RMSE | 3.472 3 | 3.593 9 |
| | MAPE | 0.473 4 | 0.509 9 |
| | DTW | 1 522.672 2 | 1 517.951 2 |

2)Bi-LSTM与MLRA相比, RMSE降低0.1066, 提高2.42%, Bi-LSTM模型的交通流量预测结果和真实值对应样本点的误差平方根较小, 预测精度比经典MLRA模型流量预测高, 预测效果较好;

3)Bi-LSTM与MLRA相比, MAPE降低0.0142, 提高4.46%, Bi-LSTM模型的测试集上平均流量测试值与实际测试值之间的相对误差小, 说明Bi-LSTM模型优于MLRA模型。

DTW结果: 表3中MLRA模型、Bi-LSTM模型时间序列相似性结果DTW分别为52938.6356、54815.1055, Bi-LSTM模型较MLRA模型提高3.42%, 说明Bi-LSTM的流量预测值更加接近真实值, Bi-LSTM优于MLRA模型, 表现出较好的鲁棒性。

Bi-LSTM模型较经典MLRA模型在对交通网络集群进行流量预测时表现出较好的泛化性和鲁棒性, Bi-LSTM在交通流量集群预测方面比MLRA模型更有优势。

表3 模型的平均评价结果

Table 3 Average evaluation results of the model

| Models | MAE | RMSE | MAPE | DTW |
|---------|---------|---------|---------|--------------|
| MLRA | 3.201 1 | 4.400 9 | 0.318 7 | 52 938.635 6 |
| Bi-LSTM | 3.068 7 | 4.294 3 | 0.304 5 | 54 815.105 5 |
| 模型提升比例 | +4.14% | +2.42% | +4.46% | +3.42% |

5 结 语

在Tensorflow2.0深度学习的理论框架下, 基于Python语言, 利用贝叶斯信息准则和高斯混合模型, 将网络化的交通流量数据序列分解为若干个相似交通状况的网络集群, 设计了多对多模型, 根据交通网络可用的

时间序列数据,构建了基于Bi-LSTM交通流量时间序列预测框架,借助Keras完成逐层网络构建和精细化调参,实现基于Bi-LSTM对城市交通网格集群的短时交通流预测。Bi-LSTM模型较MLRA模型在MAE、MSE、MAPE方面分别提高了4.14%、2.40%、4.46%,Bi-LSTM模型表现出更好的泛化性能。2个时间序列之间相似性表现出的DTW结果为Bi-LSTM组合预测模型较MLRA模型提高了3.42%,Bi-LSTM要优于MLRA,表现出较好的鲁棒性。说明基于Bi-LSTM对城市交通网格集群流量预测是可行的,具有应用于城市交通网格流量预测的潜力。

参考文献

- [1] 刘宜成,李志鹏,吕淳朴,等. 基于动态时间调整的时空图卷积路网交通流量预测[J]. 交通运输系统工程与信息, 2022, 22(3): 147-157, 178.
Liu Y C, Li Z P, Lv C P, et al. Network-wide traffic flow prediction research based on DTW algorithm spatial-temporal graph convolution[J]. Journal of Transportation Systems Engineering and Information Technology, 2022, 22(3): 147-157, 178. (in Chinese)
- [2] Chen C L, Liu Y B, Chen L, et al. Bidirectional spatial-temporal adaptive transformer for urban traffic flow forecasting[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, PP(99): 1-13.
- [3] Vythoulkas P. Alternative approaches to short term traffic forecasting for use in driver information systems[J]. Transportation and Traffic Theory, 1993, 12: 485-506.
- [4] Ahmed M S. Analysis of freeway traffic time series data and their application to incident detection[M]. Oklahoma; The University of Oklahoma, 1979.
- [5] Feng X X, Ling X Y, Zheng H F, et al. Adaptive multi-kernel SVM with spatial - temporal correlation for short-term traffic flow prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(6): 2001-2013.
- [6] Lee J W, Hong B, Lee K, et al. A prediction model of traffic congestion using weather data[C]//2015 IEEE International Conference on Data Science and Data Intensive Systems. December 11-13, 2015. Sydney, NSW, Australia:IEEE, 2016: 81-88.
- [7] Zhang W B, Yu Y H, Qi Y, et al. Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning [J]. Transportmetrica a Transport Science, 2019, 15(2): 1688-1711.
- [8] Fu R, Zhang Z, Li L. Using LSTM and GRU neural network methods for traffic flow prediction[C]//2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC). November 11-13, 2016. Wuhan, China:IEEE, 2017: 324-328.
- [9] Zheng C P, Fan X L, Wang C, et al. GMAN: a graph multi-attention network for traffic prediction[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 1234-1241.
- [10] Wang Z Y, Thulasiraman P, Thulasiram R. A dynamic traffic awareness system for urban driving[C]//2019 International Conference on Internet of Things (iThings), IEEE Green Computing and Communications (GreenCom), IEEE Cyber, Physical, Social Computing (CPSCom) and IEEE Smart Data (SmartData), July 14-17, 2019. Atlanta, GA, USA: IEEE, 2019: 945-952.
- [11] Wang Z Y, Thulasiraman P. Foreseeing congestion using LSTM on urban traffic flow clusters[C]//2019 6th International Conference on Systems and Informatics (ICSAI). November 2-4, 2019, Shanghai, China. IEEE, 2020: 768-774.
- [12] Chiabaut N, Faitout R. Traffic congestion and travel time prediction based on historical congestion maps and identification of consensual days[J]. Transportation Research Part C: Emerging Technologies, 2021, 124: 102920.
- [13] Long J C, Gao Z Y, Ren H L, et al. Urban traffic congestion propagation and bottleneck identification[J]. Science in China Series F: Information Sciences, 2008, 51(7): 948-964.
- [14] Chen K H, Jia J P. Network evasion detection with Bi-LSTM model[J]. Journal of Physics: Conference Series, 2019, 1168: 052009.
- [15] 刘志远, 张文波. 交通大数据: 理论与方法[M]. 杭州: 浙江大学出版社, 2020.
Liu Z Y, Zhang W B. Traffic big data: theory and method[M]. Hangzhou: Zhejiang University Press, 2020. (in Chinese)
- [16] 赵杨璐, 段丹丹, 胡饶敏, 等. 基于EM算法的混合模型中子总体个数的研究[J]. 数理统计与管理, 2020, 39(1): 35-50.
Zhao Y L, Duan D D, Hu R M, et al. On the number of components in mixture model based on EM algorithm[J]. Journal of Applied Statistics and Management, 2020, 39(1): 35-50. (in Chinese)
- [17] Schwarz G. Estimating the dimension of a model[J]. The Annals of Statistics, 1978, 6(2): 461-464.
- [18] Likas A, Vlassis N, Verbeek J J. The global K-means clustering algorithm[J]. Pattern Recognition, 2003, 36(2): 451-461.
- [19] Reynolds D. Gaussian mixture models[M]. Boston, MA: Springer, 2009: 659-663.
- [20] 王博文, 王景升, 王统一, 等. 基于长短时记忆网络的Encoder-Decoder多步交通流预测模型[J]. 重庆大学学报, 2021, 44(11): 71-80.
Wang B W, Wang J S, Wang T Y, et al. An encoder-decoder multi-step traffic flow prediction model based on long short-time memory network[J]. Journal of Chongqing University, 2021, 44(11): 71-80. (in Chinese)

(编辑 侯 湘)