

基于全局时空图卷积神经网络的城市交通流量预测

王佳昊, 黎文斌, 郭仕尧, 向平

引用本文

王佳昊, 黎文斌, 郭仕尧, 向平. [基于全局时空图卷积神经网络的城市交通流量预测](#)[J]. 计算机科学, 2024, 51(11A): 240200045-9.

WANG Jiahao, LI Wenbin, GUO Shiyao, XIANG Ping. [Urban Traffic Flow Prediction Based on Global Spatiotemporal Graph Convolutional NeuralNetwork](#) [J]. Computer Science, 2024, 51(11A): 240200045-9.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于迁移学习的动态环境室内定位方法研究](#)

Indoor Location Algorithm in Dynamic Environment Based on Transfer Learning
计算机科学, 2024, 51(5): 277-283. <https://doi.org/10.11896/jsjcx.230300137>

[ST-WaveMLP:面向交通流量预测的时空全局感知网络模型](#)

ST-WaveMLP:Spatio-Temporal Global-aware Network for Traffic Flow Prediction
计算机科学, 2024, 51(5): 27-34. <https://doi.org/10.11896/jsjcx.230100086>

[基于可穿戴设备的开放集动作识别技术研究](#)

Study on Open Set Activity Recognition Technology Based on Wearable Devices
计算机科学, 2024, 51(4): 291-298. <https://doi.org/10.11896/jsjcx.230300158>

[基于图对比学习的多模态交通流量协同预测方法](#)

Co-Forecasting for Multi-modal Traffic Flow Based on Graph Contrastive Learning
计算机科学, 2023, 50(10): 135-145. <https://doi.org/10.11896/jsjcx.230700127>

[基于元学习和注意力机制的用户行为识别研究](#)

Human Activity Recognition with Meta-learning and Attention
计算机科学, 2023, 50(8): 193-201. <https://doi.org/10.11896/jsjcx.220900124>

基于全局时空图卷积神经网络的城市交通流量预测

王佳昊 黎文斌 郭仕尧 向平

电子科技大学信息与软件工程学院 成都 610051

(wangjh@uestc.edu.cn)

摘要 交通流量预测在智能交通系统(ITS)中发挥着重要作用,将城市中复杂的时空相关性高效且全面地提取出来是交通流量预测中面临的关键挑战。交通速度不仅在时间维度上具有短期和长期周期性依赖关系,而且在空间维度上具有局部和全局依赖性,现有方法对捕获交通数据的时空依赖关系有一定的局限。为此,文中提出了一种基于全局时空图卷积神经网络(Global Spatial-Temporal Graph Convolutional Network,GSTGCN)的深度学习模型,用于解决在城市交通速度预测的局限性。该模型中存在3种时空分量,可相应地对交通数据中的近期、天周期、周周期这3种不同的时空相关性进行建模。每个时空分量都由时间模块和空间模块组成,时间模块为了更好地获取交通数据的时间维度信息,引入了Informer机制以自适应地分配特征权重。空间模型为了更好地获取交通数据的空间关系,引入了图卷积神经网络来提取交通数据的局部和全局空间信息。在两个不同的真实数据集上进行了测试,结果表明所提出的GSTGCN优于最先进的基线模型。

关键词: 交通流量预测;全局时空图卷积网络;时空依赖性

中图分类号 TP181

Urban Traffic Flow Prediction Based on Global Spatiotemporal Graph Convolutional Neural Network

WANG Jiahao, LI Wenbin, GUO Shiyao and XIANG Ping

School of Information and Software Engineering, University of Electronic Science and Technology, Chengdu 610051, China

Abstract Traffic flow prediction plays an important role in intelligent transportation systems(ITS). The key challenge in traffic flow prediction is to efficiently and comprehensively extract the complex spatiotemporal correlations in cities. Traffic speed has not only short-term and long-term periodic dependencies in the temporal dimension, but also local and global dependencies in the spatial dimension. Existing methods have certain limitations in capturing the spatiotemporal dependencies of traffic data. To this end, this paper proposes a deep learning model based on the global spatiotemporal graph convolutional network(GSTGCN) to address the limitations of urban traffic speed prediction. There are three spatiotemporal components in the model, which can model the three different spatiotemporal correlations in traffic data, namely, recent, daily, and weekly cycles. Each spatiotemporal component consists of a time module and a spatial module. In order to better obtain the temporal dimension information of traffic data, the time module introduces the Informer mechanism to adaptively assign feature weights. In order to better obtain the spatial relationship of traffic data, the spatial model introduces a graph convolutional neural network to extract local and global spatial information of traffic data. In the experiments, the proposed model is tested on two different real-world datasets. The results show that the proposed GSTGCN outperforms the most advanced baseline models.

Keywords Traffic flow prediction, Global spatiotemporal graph convolutional neural network, spatiotemporal dependence

1 引言

随着城市化加速发展,城市道路愈加复杂,交通拥堵问题变得更加严重。统计数据表明,在我国各大城市中的交通拥堵现象造成了巨大的经济损失,每年造成高达1600亿元的经济损失且该损失还在逐年上升。为了缓解城市日益严重的交通拥堵现象同时降低拥堵造成的经济损失,城市中道路交

通状况和通行能力亟需得到改善。因此如果能够通过监控城市中道路的交通流量,并且预测未来的交通流量、道路中车流速度和其他相关数据,将极大地帮助交通部门采取有效的分流措施,缓解交通压力并提高交通利用率,以有利于交通建设和降低经济损失^[1]。

交通流量预测指根据历史交通流量数据,采用某种预测模型或算法对其进行训练或分析,从而预测未来一段时间的

基金项目:电子科技大学-智小金-智能家居联合研究中心项目(H04W210180);内江市科技孵化和成果转化专项资金(2021KJFH004);四川省科技支撑计划项目(2022YFG0212,2021YFG0024);泸州市科技计划项目(2022-XDY-192)

This work was supported by the UESTC-ZHIXIAOJING Joint Research Center of Smart Home(H04W210180),Nejiang Technology Incubation and Transformation Funds(2021KJFH004),Science and Technology Support Plan of Sichuan Province of China(2022YFG0212,2021YFG0024) and Luzhou Science and Technology Plan Project(2022-XDY-192).

通信作者:黎文斌(12916227265@163.com)

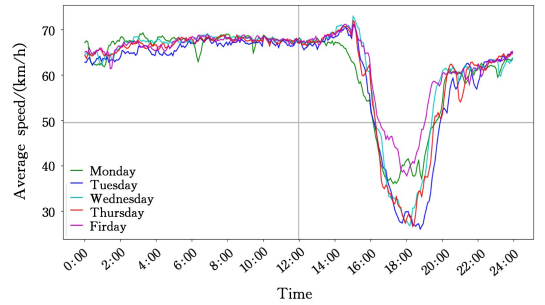
交通流量信息^[2]。交通流量预测的主要目标是利用历史交通数据开发一个模型,以预测未来时间段内道路网络的交通速度。随着预测时间的延长,任务的复杂性也在增加。与短期交通速度预测(如 15 min 内)相比,长期预测(超过 15 min)具有更大的价值。因此,为了提高研究的实际相关性,本文将专注于 1 h 的长时交通速度预测研究。

交通流量在时间维度和空间维度都具有依赖关系。在时间维度上有短期和长期两种周期性依赖关系,在空间维度上具有局部和全局两种依赖关系。

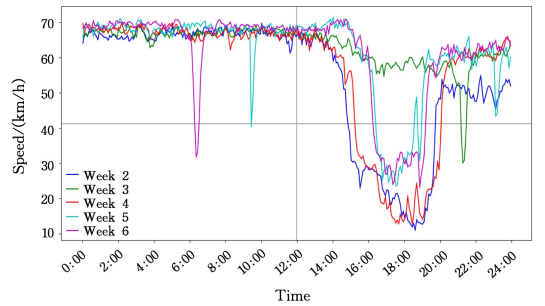
1) 交通流量的多重时间依赖性。图 2 给出了路网的拓扑结构和传感器之间复杂的时空相关性,其中传感器 1 在 $t+1$ 时刻下与 $t-l+1$ 时刻的相关性比在 t 时刻下的相关性更高。并且,图 1(a) 显示出在不同工作日下午,交通速度随时间的变化趋势具有高度一致性。此外,图 1(b) 也表明在同一工作日的不同星期中,交通速度的变化趋势也具有高度一致性。这表明交通速度同时呈现短期邻近和长期周期两种依赖关系。因此,我们同时考虑近期、每天和每周的周期模式来进行交通预测。

2) 交通流量的全局空间依赖性。如图 2 所示,交通网络上不同节点的空间相关性不仅是局部相关的,也是全局相关的。以传感器 1 为例子,其相邻传感器 2 和 3 的交通状况会影响 1 的交通状况,这是局部空间相关性。同时传感器 4 虽远离传感器 1 但是可以间接影响传感器 1 的交通状态。因此交通网络上不仅是与传感器 1 相邻的传感器会产生影响,网

路上的其他传感器也会对传感器 1 产生影响,这是全局空间相关性的表现。



(a) Average speed by time of day for each weekday



(b) Speed by time of day for each monday of weeks 2-6

图 1 PeMSD7 交通速度的多个时间依赖性

Fig. 1 Multiple temporal dependencies of traffic speed for PeMSD7

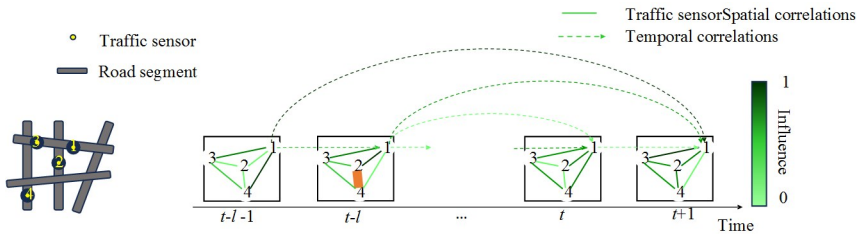


图 2 路网的拓扑结构和传感器之间复杂的时空相关性

Fig. 2 Topological structure of road network and complex spatial-temporal correlations between sensors

早期对于交通流量预测多使用统计方法和传统的机器学习方法,但不能有效地对交通数据的非线性时间相关性建模,同时也很少考虑空间的依赖关系。近年来多数研究者在交通流量预测领域使用基于深度学习的方法。将卷积神经网络(CNN)和循环神经网络(RNN)结合使用来进行交通流量预测,利用 CNN 捕捉空间依赖性,同时利用 RNN 对交通数据的时间相关性进行提取。

上述提到的传统深度学习算法有一定局限性,传统方法只能有效处理规则的网格结构,在面临拓扑不规则结构的数据时效果不甚理想。为此,本文采用一种新的深度学习模型——全局时空图卷积网络(GSTGCN),该模型使用 3 种不同的时空组件,能够有效捕捉交通路网中的不同周期性时空变化。首先,在每个时空组件中,在时间维度中使用 Informer 来捕捉交通速度的时间依赖性,在空间维度中使用图卷积神经网络(GCN)并结合全局相关机制,捕捉局部与全局的空间相关性。此外,考虑到其他外部因素也会对交通速度产生一些影响,因此本文对检测器的物理位置数据、POI 数据以及道路结构数据构建了一个静态邻接矩阵,用于表示整个交通道路网络,以降低外部因素的影响。在实验测试部分,本文模型

在两个真实交通数据集上进行了测试,实验验证了本文提出的模型在真实数据集上的有效性,尤其是在捕捉交通速度的时空特性方面表现良好。

2 相关工作

在过去的几十年里,交通流量预测得到了广泛的研究。早期的交通流量预测统计方法是简单的时间序列模型,包括自回归积分移动平均(ARIMA)及其变体^[3-4]、向量自回归(VAR)^[5]等,这些方法依赖于数据平稳假设,因此对复杂交通数据进行建模的能力有限。后来,基于传统机器学习方法的模型,如 k 近邻(KNN)^[6]、支持向量回归(SVR)^[7]和卡尔曼滤波(KF)^[8]被应用于流量预测,建立模型适用于更复杂的数据。然而,这些方法对捕捉交通数据中的非线性具有局限,并且几乎没有利用空间相关性^[9],此外它们都需要更详细的特征工程。

深度学习的方法已在许多领域得到应用并取得成功,这启发了交通预测研究使用基于深度学习的方法对交通数据复杂的时空依赖性进行建模。Lv 等^[10]利用堆叠自动编码器(SAE)来预测不同节点的流量状态。Luo 等^[11]集成 KNN 和

LSTM^[12]来预测交通流。Yu等^[13]将LSTM网络与SAE相结合来预测极端条件下的交通状态。Cui等^[14]提出了一种基于LSTM的网络,由双向网络和单向网络组成,用于流量预测。此外,Zhang等^[15]将道路网络转化为规则的二维网格,并使用卷积神经网络来预测全市人群流量。Liu等^[16]使用全连接神经网络和改进的残差网络来预测公交车交通流量。文献[17-19]将卷积神经网络(CNN)与循环神经网络(RNN)及其变体结合起来进行流量预测。然而,以上模型使用传统的卷积运算,传统卷积运算在处理不规则拓扑网格结构数据时无法有效捕捉信息。因此,以上模型不适用于捕获交通网络的空间相关性。

在具有复杂拓扑的交通数据中提取空间相关性,将神经网络应用于处理图结构数据是目前广泛讨论的热点^[20]。一系列研究将传统的卷积扩展到对光谱^[21-23]或空间^[24-26]域上的任意图进行建模。基于谱的方法是通过图谱滤波器对节点的输入信号进行平滑处理。基于空间的方法则通过收集邻居节点的特征信息来提取节点的高级表示。其他研究侧重于图嵌入,用于学习保留图结构信息的顶点的低维表示^[27-28]。为了克服传统卷积的局限性并捕获更复杂的时空依赖性,Li等^[29]提出了将扩散卷积与循环神经网络相结合的框架来预测交通状况。Fang等^[30]提出了用于交通流预测的全球时空网络(GSTNet)。GSTNet采用张量随意卷积和全局相关机制来提取动态时间依赖性和全局空间相关性。Yu等^[31]提出了时空GCN(ST-GCN),它采用图卷积与一维卷积相结合的全卷积结构。在ST-GCN中,图卷积用于获得空间相关性,

一维卷积用于提取时间依赖性。STGCN比上述使用RNN的模型计算效率更高。随后,ST-MetaNet^[32]利用序列到序列的结构,并将图注意网络(GAT)与循环神经网络(RNN)相结合来捕获时空相关性。Wu等^[33]将Wavenet^[34]集成到GCN中,用于在交通数据中提取动态时间依赖性,同时使用自适应邻接矩阵来捕捉道路网络中的隐含空间依赖性。该自适应邻接矩阵是基于道路网络中不同节点嵌入的相似性构建的。然而,节点之间学习到的空间依赖性缺乏领域知识的指导,这在训练阶段容易导致过拟合^[35]。此外,大多数现有的交通速度预测方法忽略了路网中不同节点之间的全局空间相关性,并且几乎没有利用多重时间相关性和外部因素。

3 全局时空图卷积网络(GSTGCN)模型

如图3所示,全局时空图卷积神经网络模型由3个独立但结构相同的时空组件构成,它们各自专注于捕获历史速度中的近期、天周期、周周期的时空依赖性。

3个独立的时空组件的输入数据分别为近期数据 X_h 、每日周期数据 X_d 和每周周期数据 X_w ,3种输入数据都提取自交通速度数据。在每个时空组件中都包含两个模块,分别为基于Informer的时间卷积模块和全局相关图卷积模块。最后将3个时空组件的输出 Y_h, Y_d, Y_w 进行参数矩阵融合和映射操作,为不同的输出赋予不同的权重,生成最终的预测结果 y_{res} 。

这一整合模型结构充分考虑了交通数据的时空特性,为准确预测交通速度提供了有力的工具。

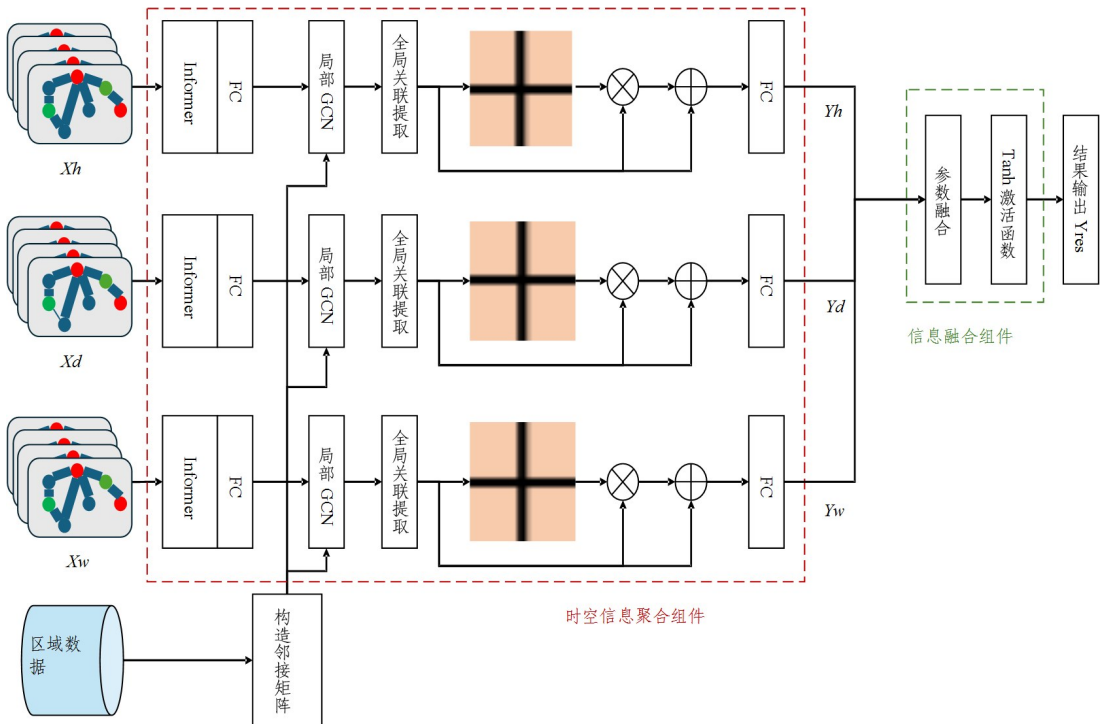


图3 GSTGCN结构

Fig.3 GSTGCN structure

3.1 静态邻接矩阵的构建

在以前的研究中,邻接矩阵的构建通常是利用两个或多个检测器间的物理位置距离作为参数构建。然而,即使两个检测器间的物理距离足够远,但如果两个检测器都处于相同或相似功能的区域(如都在学校区域或者商场区域内),它们

则会在某些特定的时段表现出相似的变化趋势^[36]。因此,本文在构建邻接矩阵时做了如下改进:1)检测器监测点之间的距离;2)检测器点所在区域的功能相似性。这些数据用于构建邻接矩阵 A 。本文首先使用Dijkstra算法计算了在交通网络中检测器中两两的最短路径距离 $dist(i, j)$ 。随后对于

具体的检测器 i , 结合采集到的 POI 数据和检测器 i 所处位置的特征数据, 将这两种特征数据进行拼接形成特征向量 \mathbf{Er} 。这一整合考虑了检测器之间的位置距离、功能相似性和道路的网络结构, 能为模型提供更综合的静态信息, 在建模时提高了对复杂空间相关性的捕捉能力。 \mathbf{Er} 的具体定义为:

$$\mathbf{Er} = (POI_1, \dots, POI_{10}, \text{type}, \text{length}, \text{lanes}, \text{is_bridge}, \text{speedlimit}, \text{oneway}) \quad (1)$$

其中, POI_i 表示第 i 种感兴趣区的数量。

随后, 对于检测器之间所在区域功能的相似度计算, 本文采用了余弦相似度公式。如检测器 i 和检测器 j 之间的相似度的具体计算式为:

$$\text{sim}(i, j) = \frac{\sum_{m=1}^C E_{i,m} \times E_{j,m}}{\|E_i^c\| \times \|E_j^c\|} \quad (2)$$

其中, C 为特征向量的长度, \mathbf{Er} 为检测器 i 的特征向量。最后使用基于阈值的高斯核来计算静态邻接矩阵, 其计算式如下:

$$\mathbf{A}_{i,j} = \begin{cases} S_{i,j} = w_1 \exp\left(-\frac{\text{dist}(i,j)^2}{2\sigma_1^2}\right) + w_2 \exp\left(-\frac{1-\text{sim}(i,j)^2}{2\sigma_2^2}\right), & \text{if } S_{i,j} \geq \kappa \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

其中, κ 是阈值, σ_1 和 σ_2 分别是距离的标准差和相似度的标准差, $w_1 + w_2 = 1$, 本文使用的静态邻接矩阵 \mathbf{A} 是上述过程构建的。

3.2 时间序列段截取

由于时空交通数据是由多种时间周期性变化模式共同作用的结果, 因此下文对模型使用到的时空数据的处理

$$\mathbf{X}_d = \begin{pmatrix} X_{t_0 - (\frac{T_d}{T_p}) * q + 1}, X_{t_0 - (\frac{T_d}{T_p}) * q + 2}, \dots, X_{t_0 - (\frac{T_d}{T_p}) * q + T_p}, X_{t_0 - (\frac{T_d}{T_p} - 1) * q + 1} \\ X_{t_0 - (\frac{T_d}{T_p} - 1) * q + 2}, \dots, X_{t_0 - (\frac{T_d}{T_p} - 1) * q + T_p}, \dots, X_{t_0 - q + 1}, \dots, X_{t_0 - q + T_p} \end{pmatrix} \in \mathbb{R}^{N \times F \times T_d} \quad (5)$$

图 4 中, 绿色部分的时间段是历史数据中若干天与模型需要预测的时间段相同的部分。其数据与预测时间段的数据有

$$\mathbf{X}_w = \begin{pmatrix} X_{t_0 - (\frac{T_w}{T_p}) * 7 * q + 1}, X_{t_0 - (\frac{T_w}{T_p}) * 7 * q + 2}, \dots, X_{t_0 - (\frac{T_w}{T_p}) * 7 * q + T_p}, X_{t_0 - (\frac{T_w}{T_p} - 1) * 7 * q + 1} \\ X_{t_0 - (\frac{T_w}{T_p} - 1) * 7 * q + 2}, \dots, X_{t_0 - (\frac{T_w}{T_p} - 1) * 7 * q + T_p}, \dots, X_{t_0 - 7 * q + 1}, \dots, X_{t_0 - 7 * q + T_p} \end{pmatrix} \in \mathbb{R}^{N \times F \times T_w} \quad (6)$$

图 4 中, 蓝色部分是周周期数据, 该部分的数据是由历史数据中若干周与模型预测时间段相同的一部分数据构成, 该部分的数据与预测时间段的数据拥有相同的周属性。周周期组件的作用是获取时空数据的周周期规律。

3.3 全局时空图卷积网络模型

3.3.1 Informer 时间块结构

Informer 模型是一种基于注意力机制的深度学习模型, 具备更为强大的能力来准确捕捉长序列中输入和输出的依赖关系。其具体结构如图 5 所示。

Informer 是一种用于时间序列预测的神经网络结构, 特别适用于长序列的预测任务, 具有高效、灵活和可扩展的特点。Informer 结构综合了自注意力机制 (Self-Attention)、卷积操作 (Convolution)、自回归机制 (Autoregressive)、门控机制 (Gating Mechanism) 等多种技术, 能够有效地捕获序列数据中的长期依赖关系。

输入序列经过嵌入层进行特征编码, 将原始输入序列转换为固定维度的特征向量表示。接着, Informer 采用了一个多层的编码器结构, 每一层包含了多头自注意力机制、卷积层

规则进行阐述。

表 1 特征向量字段含义

Table 1 Meaning of the field of feature vector

字段名称	具体含义
POI_i	第 i 类感兴趣区的数量
oneway	路段是否为单行道
speedLimit	路段的限速值
type	路段的类型
length	路段的长度
isBridge	路段是否为桥梁
lanes	路段车道数量

假设 q 为每天时空数据采样的次数, t_0 为当前时间, T_p 为预测的长度。如图 4 所示, T_h , T_d , T_w 为按照特定规则截取的 3 个时间序列片段, 其详细截取规则如下。

近期模式的数据获取:

$$\mathbf{X}_h = (X_{t_0 - T_h + 1}, X_{t_0 - T_h + 2}, \dots, X_{t_0}) \in \mathbb{R}^{N \times F \times T_h} \quad (4)$$

图 4 中, 红色部分的时间段和模型未来预测的时间段是紧挨着的, 在时序预测问题中, 与模型未来预测时间段直接相邻的时间段的数据, 会对预测结果产生较大的影响。

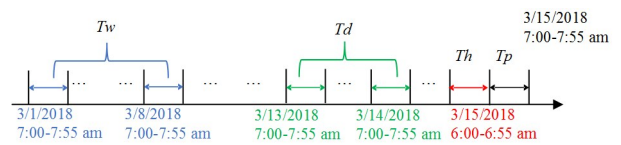


图 4 3 种数据的截取示例图

Fig. 4 Three examples of data extraction diagrams

天周期的数据获取:

相同的天属性。天周期组件用于捕捉数据的天周期依赖关系。

周期数据获取:

和门控循环单元 (GRU) 组成。自注意力机制用于学习序列中不同时间步之间的关联关系, 允许模型在编码过程中同时考虑到序列中的所有位置信息。卷积层用于捕获序列中的局部模式和时间特征。门控循环单元用于建模序列中的时间依赖关系。

编码器输入向量表达式为:

$$x_{en} = x^t \quad (7)$$

式 (7) 中原始输入序列转换为 x_{en} 。

解码器部分根据时序自回归模块的输出以及编码器的输入特征, 来预测未来的时间步的输出。解码器采用了多层的解码器结构, 每一层包含了多头自注意力机制和门控循环单元。自注意力机制用于聚合编码器和时序自回归模块的信息, 以便生成最终的输出序列。解码器的表达式如下:

$$x_{ed} = x_{\text{end-half}} + x_{t_0} \quad (8)$$

Informer 模型在编码层和解码层中引入了时序自回归模块 (Temporal Attention), 用于处理时序数据中的时间依赖关系。该模块将编码器的输出作为输入, 通过多层的时序自注意力机制来建模序列中不同时间步之间的依赖关系。时序自

注意力机制允许模型对序列中的每个时间步都赋予不同的权重,以便更好地捕捉序列中的长期依赖关系。

Informer 模型在编码层和解码层中引入了时序自回归模块(Temporal Attention),用于处理时序数据中的时间依赖关系。该模块将编码器的输出作为输入,通过多层的时序自注意力机制来建模序列中不同时间步之间的依赖关系。时序自注意力机制允许模型对序列中的每个时间步都赋予不同的权重,以便更好地捕捉序列中的长期依赖关系。概率计算式如下:

$$A(q_i, k, v) = \sum_j \frac{k(q_i, k_j)}{\sum_l k(q_i, k_l)} = E_p(k_j | q_i)[v_j] \quad (9)$$

$$p(k_j, q_i) = \frac{k(q_i, k_j)}{\sum_l k(q_i, k_l)} \quad (10)$$

$$q(k_j, q_i) = \frac{1}{L_k} \quad (11)$$

其中,令 q_i, k_i, v_i 分别表示 Q, K, V 中的第 i 行, $p(k_j, q_i)$ 表示第 i 个查询 query 对所有 keys 的注意力概率分布, $q(k_j, q_i)$ 表示查询 query 的均匀分布。

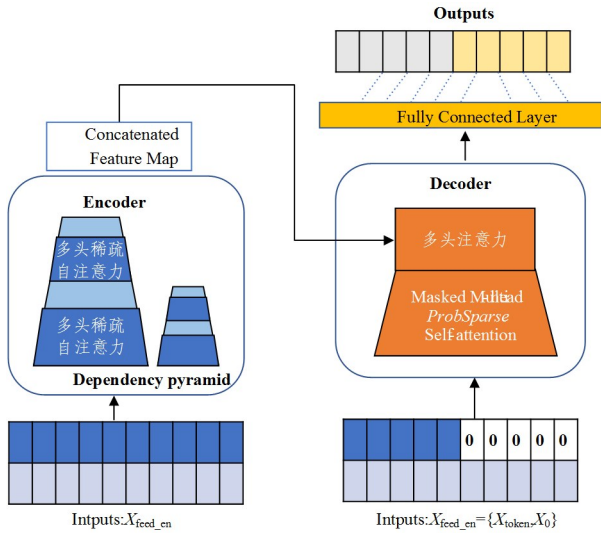


图5 Informer 结构
Fig. 5 Informer structure

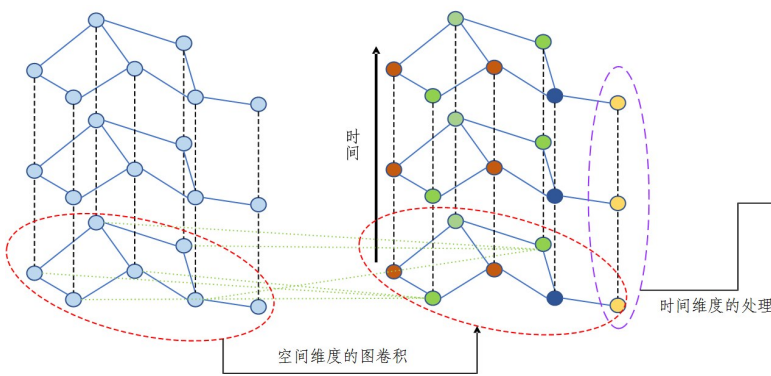


图7 时空相关性的卷积示意图
Fig. 7 Schematic diagram of convolution for spatiotemporal correlation

全局空间相关性的计算式如式(13)所示,该公式能够综合考虑所有节点之间的相互作用,从而准确地揭示它们之间的全局空间依赖关系。这一机制的引入有助于模型更精确地理解和预测时空数据的动态变化,提升了模型的预测性能和应用价值。

3.3.2 全局相关图卷积模块结构

本小节设计了一个全局相关图卷积模块。该模块结合了局部图卷积和带有残差连接的全局相关机制,旨在全面捕获复杂交通路网之间的局部和全局空间相关性。局部图卷积层负责提取相邻节点间的局部空间特征,而全局相关机制则通过残差连接的方式,有效地捕捉整个复杂交通路网中的全局空间关系。这种设计使得模型能够更深入地理解和利用交通路网中节点中的空间依赖性,进而提升模型的预测准确性和泛化能力。通过结合这两个关键组件,全局相关图卷积模块为处理复杂时空数据提供了有效的工具。

1) 局部图卷积

在图卷积网络(GCN)模型中,图卷积扮演着至关重要的角色,它作为一种基本的运算式,使得模型能够在图结构上有效地进行特征的学习与传播。通过图卷积,GCN 能够捕获节点间的连接关系,并在整个图网络中传递和融合节点的特征信息,从而实现了对图数据的深度学习和分析,具体过程如图6所示。这一过程有助于模型更好地理解 and 利用节点之间的关系,提高图上任务的性能。其计算式如下:

$$H^{(l+1)} = \sigma(D^{-\frac{1}{2}} \hat{A} D^{-\frac{1}{2}} H^{(l)} W^{(l)}) \quad (12)$$

其中, $H^{(l+1)}$ 表示经过一层图卷积操作后得到的新的节点特征矩阵, \hat{A} 表示图的邻接矩阵的一个变体, \hat{A} 是 \hat{A} 的度矩阵, σ 表示激活函数如 ReLU, $W^{(l)}$ 表示 l 层节点可学习权重矩阵。

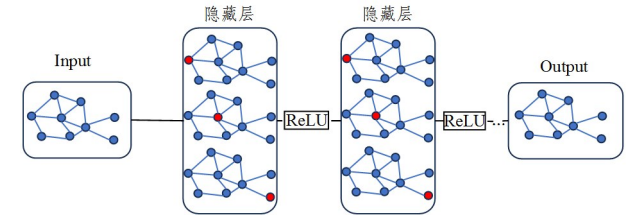
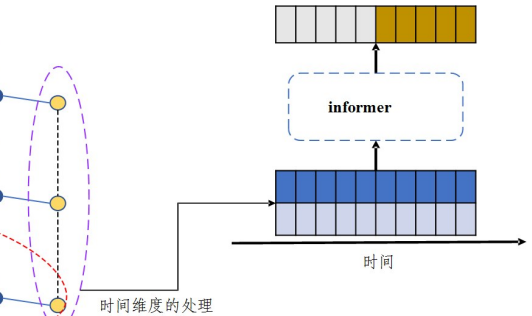


图6 局部图卷积结构
Fig. 6 Local graph convolutional structure

2) 全局空间相关机制

为了更全面地捕捉不同节点之间的全局空间相关性,本文设计了一种创新的全局空间相关机制,如图7所示。



$$y_{i,t} = \sum_{\forall v_j \neq v_i} s_{i,j} \cdot \Phi(x_{i,t}, x_{j,t}) \cdot \hat{x}_{j,t} W_g + \hat{x}_{i,t} W_r \quad (13)$$

其中, $y_{i,t}$ 表示当时间为 t 时,节点 i 的输出特征, $s_{i,j}$ 表示节点 i 和 j 之间存在边的关系, Φ 是高斯核函数,用来衡量两个节点 i 和 j 之间的编码表示之间的相关性,具体

公式为 $\Phi(x, y) = \exp(xWy)$, 其中 W 是可学习参数。

$\sum_{v_j \neq v_i} s_{i,j} \cdot \Phi(\hat{x}_{i,t}, \hat{x}_{j,t}) \cdot \hat{x}_{j,t} W_g$ 这部分用于计算节点 i 在时间 t 下的预测特征值, 通过将节点 i 与其相邻节点 j 的特征表示进行加权求和, 并乘以参数矩阵 W_g 。式(13)后面部分的 $\hat{x}_{i,t} W_r$ 表示残差连接, 它将局部图卷积之后的结果与之前的节点特征相加, 这有助于防止梯度消失问题, 并加快模型训练。 Y_t 表示时刻 t 的全局相关机制输出结果, 并且最后全局空间图卷积模块的输出特征矩阵为 $Y \in \mathbb{R}^{N \times T \times F}$, 其中 T 表示时间步长, F 表示特征的维度。这个输出特征矩阵包含了所有节点在每个时间步长下的预测特征值。

3.3.3 多维度时空图卷积融合

由于这 3 个时空组件分别建模近期、天周期和周周期的时空相关性, 它们对数据的影响在不同位置上有所不同。例如, 当预测周三上午 8:00 的数据时, 在因为早高峰产生明显拥堵的区域, 天周期模块和周周期模块两个模块的输出结果可能会对最后的预测结果产生一定影响, 进而影响最后的预测性能。另外, 如果一个区域没有明显的周期性变化, 那么这两个模块可能对最后预测结果的影响就相对有限。因此不同地理位置的数据会受到各种时间模式的混合影响, 而每种模式的影响可能会有显著差异。

为了解决该问题, 本文在 3 个时空组件输出结果后通过参数矩阵融合的方法影响 3 部分的比重。最终, 通过参数矩阵融合的方法得到输出结果 Y_{res} , 使模型更灵活地适应不同位置和时间段的数据变化, 提高预测性能。 Y_{res} 融合计算式如下:

$$Y_{res} = W_h \circ Y_h + W_d \circ Y_d + W_w \circ Y_w \quad (14)$$

其中, \circ 表示对应位置上的元素相乘, Y_{res} 表示通过融合 3 个时空组件输出后得到的结果。 W_h, W_d, W_w 对 3 个时空组件输出进行加权融合, 然后通过 \tanh 函数将 Y_{res} 映射到区间 $[-1, 1]$ 中, 以获得最终的预测结果 Y 。

GSTGCN 模型根据历史 T' 个时刻的数据来预测未来 T 个时刻的数据。该模型的损失函数使用的是式(15)定义的损失函数。

$$L(\hat{Y}_{t+1}, \dots, \hat{Y}_{t+T}; \Theta) = \sum_{t=1}^T \|\hat{Y}_{t+i} - Y_{t+i}\|^2 \quad (15)$$

整个损失函数的目的是通过最小化预测值 \hat{Y}_{t+i} 与真实值 Y_{t+i} 之间的差异来调整模型参数 Θ 。

4 实验部分

4.1 实验数据集

本文使用 PEMS 平台上的高速公路交通数据集 PEMS-BAY 和 METR-LA 来验证本文模型的有效性。METR-LA 数据集和 PEMS-BAY 数据集分别包含了美国洛杉矶高速公路上的交通信息和美国旧金山高速公路上的交通信息。PEMS-BAY 和 METR-LA 两个数据集的原始数据是由分布在高速公路上的检测器每隔 30 s 记录产生。在本文实验中选取了 METR-LA 中的 204 个检测器的数据, PEMS-BAY 则选取了 325 个检测器的数据。选取后使用的数据

是将数据集中的交通速度的源数据聚合到宽度大小为 5 min 的窗口后, 随后取平均值作为实验的输入, 每个检测器在一天之中则有 288 个数据点。经过处理后, 数据集按照 7:1:2 的比例选择训练数据、验证数据和测试数据。在数据预处理阶段, 在实验阶段本文将原始的速度数据进行归一化, 映射到 $[-1, 1]$ 中。在最后的预测结果评估时, 则会归一化映射处理后的速度, 重新映射回正常速度值, 以获取最后的预测结果。

4.2 评估指标

性能评估采用广泛使用的平均绝对误差 (Mean Absolute Error, MAE)、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) 来评估模型的预测结果。

4.3 实验基准方法

对比实验参考了领域内较有代表性的 8 种主流方法。

1) 自回归积分滑动平均模型 (AutoRegressive Integrated Moving Average, ARIMA): 常用于平稳时间序列的建模。它通过结合自回归 (AR)、差分 (I) 和滑动平均 (MA) 进行预测, 适用于单变量时间序列的短期预测。

2) 支持向量回归器 (Support Vector Regression, SVR): 专门用于回归任务。它通过构建一个超平面或一组超平面来预测连续值。SVR 在时间序列预测中能有效处理非线性关系。

3) 堆叠自编码器 (Stacked Autoencoder, SAE): 是一种深度学习模型, 使用多个自编码器层来学习高层次的特征表示。它主要用于降维、特征提取和时间序列的非线性预测。

4) 深度堆叠双向 LSTM 递归神经网络 (Subsequence-based LSTM, SUB-LSTM): 是 LSTM (长短期记忆网络) 的一种变体, 针对长时间序列数据进行分段处理。SUB-LSTM 将时间序列分成子序列并分别建模, 从而改善传统 LSTM 在长序列建模中的表现。

5) 扩散卷积递归神经网络 (Diffusion Convolutional Recurrent Neural Network, DCRNN): DCRNN 是一种结合了图卷积和递归神经网络 (RNN) 的时空数据建模方法。它通过图扩散卷积捕捉空间依赖性, 并使用 RNN 捕捉时间依赖性, 适用于时空预测问题, 如交通流量预测。

6) 时空图卷积网络 (Spatio-Temporal Graph Convolutional Networks, STGCN): 能够同时处理时空数据中的空间结构和时间动态。它广泛应用于交通、社交网络等有时空相关性的领域。

7) 时空元知识网络 (Spatio-Temporal Meta-Network, ST-MetaNet): 基于元学习的时空网络模型能够在不同的时空数据上进行泛化学习。ST-MetaNet 通过结合元学习框架和时空数据, 来提升在新环境下的适应能力。

8) 时空图波网络 (Graph WaveNet): 是一种时空图神经网络模型, 通过结合图卷积和一维卷积, 建模复杂的时空依赖性。它通过图上的学习捕捉节点之间的空间依赖, 并用卷积捕捉时间序列的依赖, 适用于交通预测、能耗预测等时空相关任务。

表2 GSTGCN和基准模型在数据集METR-LA上的预测效果对比

Table 2 Comparison of prediction performance between GSTGCN and baselines on dataset METR-LA

模型	15 min			30 min			60 min		
	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
ARIMA	3.99	8.21	9.60	5.15	10.45	12.70	6.90	13.23	17.40
SVR	3.87	8.19	9.34	5.05	10.22	12.19	6.67	12.97	16.77
SAE	3.74	8.02	9.06	4.97	10.04	11.94	6.53	12.76	15.43
SUB-LSTM	3.62	7.95	8.71	4.55	9.42	10.52	6.38	11.93	14.87
DCRNN	2.77	5.38	7.30	3.15	6.45	8.80	3.60	7.60	10.50
STGCN	2.88	5.74	7.62	3.47	7.24	9.57	4.59	9.40	12.70
ST-MetaNet	2.74	5.35	7.29	3.12	6.36	8.63	3.59	7.54	10.52
Graph WaveNet	2.69	5.15	6.90	3.07	6.22	8.37	3.53	7.37	10.01
GSTGCN	2.61	5.1	6.35	3.01	6.19	8.13	3.48	7.28	9.94

4.4 实验环境及设置

在实验中,基于PyTorch框架进行模型训练,使用的硬件配置为第11代Intel i7-11700K处理器(3.60GHz)和NVIDIA GeForce RTX 3060 GPU。通过网络搜索法确定最佳模型参数。

GSTGCN模型的配置如下:时间卷积模块采用Informer结构,卷积核数量为8,隐藏层特征维度为8,超参数 α 设置为2,构建静态邻接矩阵中的参数 k, ω_1, ω_2 均设置为0.5。模型训练共30个epoch, batch大小为256,选择Adam作为优化器类型,初始学习率为0.001。在测试阶段,用过去12个小时步预测接下来1h的交通速度。

在本研究中,为了比较不同模型的性能,配置了多种机器学习模型,具体如下。

ARIMA模型的差分系数设置为0、自回归项数设置为3、移动平均项数设置为1。这些设置使ARIMA能够有效地捕捉时间序列数据中的趋势和周期性变化。

SVR模型采用了惩罚参数为0.1,使用了5个历史观测值,并选择径向基函数(RBF)作为核函数。RBF核函数的使用有助于捕捉数据中的非线性关系,从而提升模型的预测能力。

在SAE模型中,设计了3层自编码器,每层包含400个神经元。自编码器的层数和每层的神经元数量旨在增强模型对输入数据的特征学习能力,从而改善重构误差。

SUB-LSTM模型包含两层双向LSTM,其中每层具有64个神经元。为了进一步优化模型性能,第一层和最后一层双向LSTM分别设有40个神经元,这种配置有助于捕捉数据中的长期依赖关系。

表3 GSTGCN和基准模型在数据集PEMS-BAY上的预测效果对比

Table 3 Comparison of prediction performance between GSTGCN and baselines on dataset PEMS-BAY

模型	15 min			30 min			60 min		
	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
ARIMA	1.62	3.30	3.51	2.36	4.79	5.55	3.38	6.58	8.32
SVR	1.85	3.60	3.81	2.48	5.18	5.66	3.27	7.13	8.30
SAE	1.79	3.47	3.56	2.38	5.02	5.54	3.21	7.02	8.09
SUB-LSTM	1.63	3.36	3.27	2.35	4.88	5.37	3.11	6.85	7.88
DCRNN	1.38	2.95	2.91	1.74	3.96	3.94	2.07	4.78	4.97
STGCN	1.36	2.94	2.88	1.81	4.27	4.17	2.49	5.69	5.79
ST-MetaNet	1.35	2.91	2.87	1.71	3.92	3.87	2.01	4.72	4.91
Graph WaveNet	1.30	2.74	2.73	1.63	3.71	3.69	1.94	4.52	4.63
GSTGCN	1.27	2.66	2.69	1.61	3.69	3.62	1.88	4.42	4.53

STGCN模型的时空卷积模块包括3层卷积层,卷积核的数量依次为64,16,64,卷积核大小为3。这样的设置使得STGCN能够有效地处理时空数据中的复杂模式。

DCRNN模型具有编码器-解码器结构,其中包含两层递归卷积层,每层有64个神经元,并设置随机游走步长 K 为3。该配置允许模型在捕捉空间依赖性的同时,增强对动态变化的适应能力。

ST-MetaNet模型的元学习器包括两层全连接层,每层有32个神经元,并且编码器-解码器单元配备了64个隐藏单元。这种设计旨在提升模型的泛化能力和学习效率。

最后,Graph WaveNet模型由6层堆叠的膨胀卷积层组成,膨胀因子依次为1,2,1,2,1,2,扩散步长 K 设置为2,节点嵌入维度为10。这种配置能够有效地捕捉图结构数据中的时序特征和空间依赖性。

4.5 实验结果

从表2和表3中可以清晰地看到,各种模型在不同预测时间尺度下的表现各不相同。对于简单的时间序列模型,如

ARIMA,它们在短期预测(如未来15min)中往往能够展现出相对良好的性能。然而,随着预测时间跨度的增加,这些模型的预测效果开始逐渐下降。这主要是因为时间序列模型主要关注时间序列本身的变化规律,而忽略了空间因素的影响,导致在长期预测中误差逐渐累积。与此相比,那些考虑了空间图结构的模型,如STGCN,在预测准确性上展现出了明显的优势。这些模型通过引入空间图结构,将节点之间的空间关系纳入考虑范围,从而更全面地捕捉数据的时空特性。这使得它们在长期预测中能够更好地应对误差累积的问题,保持相对稳定的预测性能。在所有考虑空间图结构的模型中,GSTGCN的表现尤为突出。随着预测任务的时间跨度增大,虽然所有模型的预测误差都在逐渐增加,但GSTGCN的误差增长率却相对较低。这表明GSTGCN在处理长期预测任务时具有更好的稳定性。此外,在数据集METR-LA和PEMS-BAY的各个指标上,GSTGCN均达到了最优预测,这进一步证明了其在预测任务中的有效性。

图8为在METR-LA数据集上,随着训练周期数增加,

未来 12 个时间步的均方根误差 (RMSE) 的变化图。实验结果表明,与其他模型相比,GSTGCN 在收敛速度和训练性能方面表现更为出色。

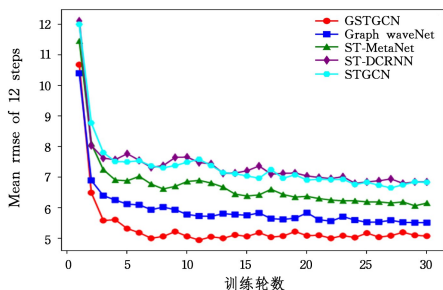


图 8 各模型在 METR-LA 数据集上随训练轮数增加预测未来 12 个时刻的平均 RMSE

Fig. 8 Mean RMSE of 12 steps versus the number of training epochs on METR-LA dataset

为了探究 GSTGCN 模型中不同子模块的效果,本文创建了 5 个变种模型:GSTGCN-g(无全局相关机制)、GSTGCN-hw(无天周期性时空模块)、GSTGCN-hd(无周周期性时空模块)、GSTGCN-s(图卷积模块替换为简单线性层)以及 GSTGCN-t(Informer 模块替换为简单线性层)。在数据集 METR-LA 上,对这些变种与原始 GSTGCN 进行了性能比较。实验结果如图 9 所示,缺少全局相关机制的 GSTGCN-g 模型在预测性能上不如原始 GSTGCN,凸显了全局相关机制在提升预测精度中的关键作用。而缺失天或周周期性时空模块的变种模型 (GSTGCN-hw 和 GSTGCN-hd) 在短期预测中与 GSTGCN 表现相近,但长期预测中性能下降,这进一步验证了融合多种周期模式对于捕捉长期时空依赖的重要性。此外,将图卷积模块和 Informer 模块替换为简单线性层后,得到的变种模型 GSTGCN-s 和 GSTGCN-t 预测误差显著上升,这证明了这两个模块在捕捉复杂时空依赖关系中的不可或缺性。进一步强调了解决时空数据预测问题时,精确捕捉时空依赖性的必要性。综上所述,通过对比实验,验证了 GSTGCN 模型中各个子模块的有效性,并强调了全局相关机制、多种周期模式的融合以及复杂时空依赖的捕捉在提升预测性能中的重要作用。

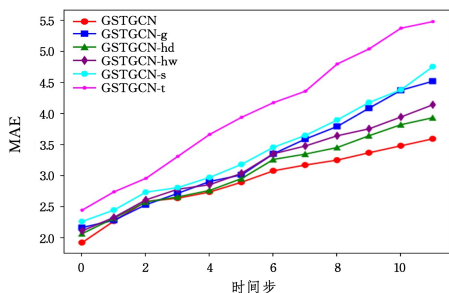


图 9 GSTGCN 模型 MAE 指标的消融实验结果

Fig. 9 Ablation study of GSTGCN model in term of MAE

结束语 本文提出了一种用于城市交通速度预测的全局时空图卷积网络 (GSTGCN) 模型。在空间维度上,该模型采用局部图卷积和全局相关机制,全面考虑了局部和全局的空间相关性。在构建表示道路网络结构的邻接矩阵时,GSTGCN 不仅考虑了传感器之间的距离,还考虑了传感器位置的相似性。在时间维度上,模型使用 3 个独立模块来建模近期、

每日周期和每周周期的时间依赖性,每个模块都由 Informer 块组成。实验结果表明,本文提出的 GSTGCN 模型在两个真实世界数据集上的预测精度显著优于现有模型。未来的工作中,计划进一步探索更复杂的空间相关性以提高预测精度。由于 GSTGCN 是一个适用于图结构数据时空预测问题的通用框架,它也可以广泛应用于其他涉及时间估计的实际应用。

参考文献

- [1] FANG Z X, HUANG S Q, SU R X, et al. Detecting Hierarchical Congestion Intervals Based on the Fusion of Multi-source Highway Data[J]. Geomatics and Information Science of Wuhan University, 2020, 45(5): 682-690.
- [2] HSUEH Y L, YANG Y R. A short-term traffic speed prediction model based on LSTM networks [J]. International Journal of Intelligent Transportation Systems Research, 2021, 19(3): 510-524.
- [3] MAKRIDAKIS S, HIBON M. ARMA models and the Box-Jenkins methodology[J]. Journal of Forecasting, 1997, 16: 147-163.
- [4] CHEN J, LI D, ZHANG G, et al. Localized space-time autoregressive parameters estimation for traffic flow prediction in urban road networks[J]. Applied Sciences, 2018, 8: 277.
- [5] ZIVOT E, WANG J. Vector autoregressive models for multivariate time series [C] // Modeling Financial Time Series with S-Plus ©; Springer: New York, NY, USA, 2006: 385-429.
- [6] ZHENG Z, SU D. Short-term traffic volume forecasting: A k-nearest neighbor approach enhanced by constrained linearly sewing principle component algorithm[J]. Transportation Research Part C Emerging Technologies, 2014, 43: 143-157.
- [7] WU C H, HO J M, LEE D T. Travel-time prediction with support vector regression [J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5: 276-281.
- [8] LIPPI M, BERTINI M, FRASCONI P. Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14: 871-882.
- [9] CHEN W, CHEN L, XIE Y, et al. Multi-Range Attentive Bi-component Graph Convolutional Network for Traffic Forecasting[J]. arXiv: 1911. 12093, 2019.
- [10] LV Y, DUAN Y, KANG W, et al. Traffic flow prediction with big data: A deep learning approach [J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 16: 865-873.
- [11] LUO X, LI D, YANG Y, et al. Spatiotemporal traffic flow prediction with KNN and LSTM [J]. Journal of Advanced Transportation, 2019, 2019: 4145353.
- [12] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computer, 1997, 9: 1735-1780.
- [13] YU R, LI Y, SHAHABI C, DEMIRYUREK U, et al. Deep learning: A generic approach for extreme condition traffic forecasting [C] // Proceedings of the 2017 SIAM International Conference on Data Mining. Houston, TX, USA, 2017: 777-785.
- [14] CUI Z, KE R, WANG Y. Deep stacked bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction [C] // Proceedings of the 6th International Workshop on Urban Computing (UrbComp 2017). Halifax, NS, Canada, 2017.

- [15] ZHANG J,ZHENG Y,QI D. Deep spatio-temporal residual networks for citywide crowd flows prediction[C]// Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 2017.
- [16] LIU P,ZHANG Y,KONG D, et al. Improved Spatio-Temporal Residual Networks for Bus Traffic Flow Prediction[J]. Applied Science, 2019, 9: 615.
- [17] DU S, LI T, GONG X, et al. A hybrid method for traffic flow forecasting using multimodal deep learning [J]. arXiv: 1803. 02099, 2018.
- [18] YAO H, TANG X, WEI H, et al. Modeling spatial-temporal dynamics for traffic prediction[J]. arXiv: 1803. 01254, 2018.
- [19] YAO H, WU F, KE J, et al. Deep multi-view spatial-temporal network for taxi demand prediction [C] // Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2018.
- [20] WU Z, PAN S, CHEN F, et al. A comprehensive survey on graph neural networks[J]. arXiv: 1901. 00596, 2019.
- [21] BRUNA J, ZAREMBA W, SZLAM A, et al. Spectral networks and locally connected networks on graphs[J]. arXiv: 1312. 6203, 2013.
- [22] DEFFERRARD M, BRESSON X, VANDERGHEYNST P. Convolutional neural networks on graphs with fast localized spectral filtering[C]// Advances in Neural Information Processing Systems; Neural Information Processing Systems (NIPS). Barcelona, Spain, 2016: 3844-3852.
- [23] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[J]. arXiv: 1609. 02907, 2016.
- [24] ATWOOD J, TOWSLEY D. Diffusion-convolutional neural networks[C]// Advances in Neural Information Processing Systems, Neural Information Processing Systems (NIPS). Barcelona, Spain, 2016: 1993-2001.
- [25] GILMER J, SCHOENHOLZ S S, RILEY P F, et al. Neural message passing for quantum chemistry[C]// Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 2017, 70: 1263-1272.
- [26] HAMILTON W, YING Z, LESKOVEC J. Inductive representation learning on large graphs[C]// Advances in Neural Information Processing Systems; Neural Information Processing Systems (NIPS). Long Beach, CA, USA, 2017: 1024-1034.
- [27] GROVER A, LESKOVEC J. node2vec: Scalable feature learning for networks[C]// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA, 2016: 855-864.
- [28] CUI P, WANG X, PEI J, et al. A survey on network embedding [J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 31: 833-852.
- [29] LI Y, YU R, SHAHABI C, LIU Y. Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting[C]// Proceedings of the International Conference on Learning Representations (ICLR'18). Vancouver, BC, Canada, 30 April-3 May 2018.
- [30] FANG S, ZHANG Q, MENG G, et al. Gstnet: Global spatial-temporal network for traffic flow prediction[C]// Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. Macao, China, 2019: 10-16.
- [31] YU B, YIN H, ZHU Z. Spatio-temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting [C]// Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAD). Stockholm, Sweden, 2018.
- [32] PAN Z, LIANG Y, WANG W, et al. Urban Traffic Prediction from Spatio-Temporal Data Using Deep Meta Learning [C] // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage, AK, USA 2019: 1720-1730.
- [33] WU Z, PAN S, LONG G, et al. Graph WaveNet for Deep Spatial-Temporal Graph Modeling[J]. arXiv: 1906. 00121, 2019.
- [34] OORD A V D, DIELEMAN S, ZEN H, et al. Wavenet: A generative model for raw audio[J]. arXiv: 1609. 03499, 2016.
- [35] CHEN W, CHEN L, XIE Y, et al. Multi-Range Attentive Bi-component Graph Convolutional Network for Traffic Forecasting[J]. arXiv: 1911. 12093, 2019.
- [36] YUAN J, ZHENG Y, XIE X. Discovering regions of different functions in a city using human mobility and POIs[C]// Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD). Beijing, China: ACM, 2012: 186-194.



WANG Jiahao, born in 1978, Ph.D, associate professor, is a member of CCF (No. 27769M). His main research interests include IoT, information security and data mining.



LI Wenbin, born in 1999, master. His main research interest is spatiotemporal data prediction.