

符琦,巴冰,黄崇景,等. 面向短时交通流预测的动态时空图卷积网络[J]. 湖南科技大学学报(自然科学版), 2024, 39(1): 70-79. doi:10.13582/j.cnki.1672-9102.2024.01.009

FU Q, BA B, HUANG C J, et al. Dynamic Spatiotemporal Graph Convolutional Networks for Short-Term Traffic Flow Prediction [J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2024, 39(1): 70-79. doi: 10.13582/j.cnki.1672-9102.2024.01.009

面向短时交通流预测的动态时空图卷积网络

符琦*, 巴冰, 黄崇景, 蒋云霞

(湖南科技大学 计算机科学与工程学院, 湖南 湘潭 411201)

摘要: 针对现有的交通流量预测模型缺乏对交通数据动态时空相关性建模能力的问题, 提出一种新的基于深度学习的动态时空图卷积网络(DSTGCN)模型, 该模型无需给出道路网络信息, 从交通数据中即可建模时空相关性. 动态时空图卷积层包含2个主要部分: (1) 动态邻接矩阵生成模块: 使用时间自相关机制与空间注意力机制捕捉交通数据中的动态时空相关性; (2) 时空图卷积: 使用图卷积和标准二维卷积对信息进行高效聚合. DSTGCN通过堆叠动态时空图卷积层, 能够捕捉不同时间级别的时空依赖关系. 文章提出的方法在美国加利福尼亚州高速公路流量公开数据集上进行了实验, 结果表明: 所提出的DSTGCN模型在各项评价指标中均优于现有的基准方法. 在PeMSD04数据集上, 与当前较新的GeoMAN和ASTGCN模型相比, MAE分别降低了4.00和2.16, 验证了所提模型在交通流预测中的有效性.

关键词: 交通流预测; 时空图卷积; 自相关机制; 时空图建模

中图分类号: TP311 文献标志码: A 文章编号: 1672-9102(2024)01-0070-10

Dynamic Spatiotemporal Graph Convolutional Networks for Short-Term Traffic Flow Prediction

FU Qi, BA Bing, HUANG Chongjing, JIANG Yunxia

(School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, China)

Abstract: Aiming at the problem that the existing traffic flow prediction models lack the ability to model the dynamic spatiotemporal correlation of traffic data, a new deep learning-based dynamic spatiotemporal graph convolutional network (DSTGCN) model is proposed. The model can model spatiotemporal correlations from traffic data without giving road network information. The dynamic spatiotemporal graph convolutional layer consists of two main parts, one is the dynamic adjacency matrix generation module, which uses temporal auto-correlation mechanism and spatial attention mechanism to capture dynamic spatiotemporal correlations in traffic data; the other is the spatiotemporal graph convolution, which uses graph convolution and standard 2D convolution to efficiently aggregate information. By stacking dynamic spatiotemporal graph convolutional layers, DSTGCN is able to capture spatiotemporal dependencies at different temporal levels. The method proposed in this paper is tested on a public California freeway dataset. The results show that the proposed DSTGCN model outperforms the existing benchmark methods in every evaluation index. On the PeMSD04 dataset, compared with the current newer GeoMAN and ASTGCN models, the MAE is reduced by 4.00 and 2.16 respectively, verifying

收稿日期: 2023-02-13

基金项目: 湖南省自然科学基金资助项目(2021JJ30274); 湖南省自然科学基金资助项目(2021JJ30278)

* 通信作者, fuqi@hnust.edu.cn

the effectiveness of the proposed model in traffic flow prediction.

Keywords: traffic flow prediction; spatiotemporal graph convolution; auto-correlation mechanism; spatiotemporal graph modeling

随着经济的高速发展,全球汽车的持有量不断提高,与日俱增的汽车数量为各国的交通运输能力带来了挑战.为了解决交通问题,越来越多的国家开始投入智能交通系统(Intelligent Transportation System, ITS)的建设.交通预测问题是实现智能交通系统的重要一环,准确的交通预测能帮助交通管理部门更好地控制交通,减少交通拥堵^[1-3],能够为道路使用者提供更好的道路选择,指引司机更快地到达目的地.

交通预测是指在给出历史交通数据(通过传感器记录)和基础道路网络的情况下,预测未来一段时间的道路交通状况(如交通量、速度等)^[4].这项任务非常具有挑战性,主要是由于交通数据中复杂的空间依赖性和随着道路状况动态变化的非线性时间趋势.早期的方法大多基于统计学习方法或传统的机器学习方法.AHMED 和 COOK^[5]在 1979 年提出使用差分自回归移动平均模型(Autoregressive Integrated Moving Average, ARIMA)来预测高速公路交通数据,之后一些 ARIMA 的变体及其他统计学习方法^[6-11]也被广泛应用于该领域.传统的机器学习方法^[12-14],如 k 近邻算法(KNN)、支持向量机(SVM)和神经网络(NN)等也被用于交通预测,可以实现更高的预测精度.但这些方法一般都是针对单个节点的小数据集设计的,受限于对非线性建模的能力,无法有效地捕捉交通数据中动态的非线性时间趋势,忽略或几乎不考虑其中复杂的空间依赖关系.此外,祁其春等^[15]构建了适应于短时交通流预测的投影寻踪回归模型,使用灰色关联度分析方法确定交通流预测影响因子,并利用粒子群优化算法优化投影指标函数.

深度学习的最新进展,使得对交通预测中复杂的时空依赖性进行建模成为可能.一些研究^[16-17]将卷积神经网络(CNN)和循环神经网络(RNN)应用于交通预测.在这些研究中,使用 CNN 来捕捉交通网络中的空间相关性,使用循环神经网络来捕捉时间相关性.但是这些基于 CNN 的方法,只能处理规则网格结构的交通数据,而没有考虑由不规则道路网络主导的非欧几里得相关性.为了解决这个问题,研究人员应用图卷积来模拟交通预测的非欧几里得相关性.LI 等^[4]提出扩散卷积循环神经网络(DCRNN),用扩散卷积算子替换门控循环单元(GRU)中的全连接层,扩散卷积在给定的图上执行图卷积,并考虑流入和流出关系;WU 等^[18]提出用时空图建模的 Graph WaveNet,融合图卷积与扩张因果卷积,开发出一种新的自适应依赖矩阵;姜山等^[19]在 Graph WaveNet 的基础上提出基于图小波卷积的时空图建模方法,进一步提升了预测性能;YU 等^[20]提出 Spatial-Temporal GCN(STGCN),结合图卷积和一维因果卷积,在 STGCN 中使用图卷积捕捉空间相关性,而在时间轴上使用卷积捕获时间相关性,由于卷积的并行性更强,使得模型比 RNN 的计算效率高;冯宁等^[21]在 STGCN 的基础上提出 MCSTGCN 模型,通过 3 个组件分别对近期、日周期、周周期特性进行建模,获得了更好的预测效果.

上述基于 GCN 的方法大多是将道路网络中各节点之间的距离编码为表示空间依赖性的固定加权图.但对于交通预测问题而言,各节点之间的空间依赖关系是随时间动态变化的.图 1 展示了交通流量的时空相关性,可以观察到不同位置对 A 的影响是不同的,

且同一个位置在不同时间对 A 的影响也是不同的.另外,存在一些在道路网络中相距甚远的节点在某些时刻表现出相似的变化趋势.例如,在现实生活中,城东和城西 2 所相距甚远的学校在上学和放学时段车流量会表现出相同的变化趋势.因此,使用基于距离编码的固定加权图无法很好地捕捉交通数据中随时间变化的动态空间依赖关系,且这些方法需要给出道路网络信息.

针对上述问题,本文构建了一种新型的动态邻接矩阵生成模块,使用自相关机制将历史交通数据在时间维度进行信息聚合,然后送入空间注意力机制中计算出当前时段各节点之间的空间相关性.将动态邻接矩阵成模块地输出作为邻接矩阵与历史交通数据一并送入时空图卷积网络进行计算,最终得到交通流量预测模

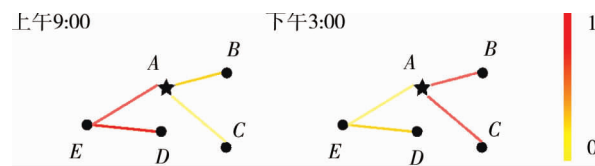


图 1 交通流量数据的动态时空相关性

型(Dynamic Spatial-Temporal Graph Convolution Network, DSTGCN).该模型在2个真实世界的城市交通速度数据集 PEMS4 和 PEMS8^[22] 上进行了广泛实验,在 15,30,60 min 的时间段内均达到了先进水平.

1 动态时空图卷积网络

在本节,首先给出交通预测问题的形式化定义,其次详细介绍使用时间自相关与空间自注意力机制生成动态邻接矩阵的动态邻接矩阵生成模块和时空图卷积模块,最后给出本文设计的模型总体架构.

1.1 问题定义

交通预测的目的是根据道路网络上 N 个相关传感器的历史数据,来预测道路网络的未来交通量(流量、速度等).我们把 N 个相关传感器构成的网络定义为一个加权有向图 $G = (V, E, A)$. 其中 V 是图 G 中的节点集合,其大小为 $|V|=N$, 使用 v_i 表示节点 i ; E 是图中的边集合,表示节点之间的连通性; $A \in \mathbf{R}^{N \times N}$ 是表示节点邻近度的加权邻接矩阵,其中 A_{v_i, v_j} 表示节点 v_i 到节点 v_j 的邻近度.

交通预测问题可以描述为在给定图 G 和其中 H 个时间步的历史观测数据,预测图 G 中未来 T 个时间步的特征信息, $X^t \in \mathbf{R}^{N \times D}$ 表示图 G 在 t 时刻的特征信息.故交通预测问题可形式化定义为求解一个映射函数 f 使其满足下列条件:

$$X^{t+1, \dots, t+T} = f(X^{t-H+1, \dots, t}, G). \quad (1)$$

式中: $X^{t+1, \dots, t+T} \in \mathbf{R}^{N \times D \times T}$; $X^{t-H+1, \dots, t} \in \mathbf{R}^{N \times D \times S}$.

1.2 动态邻接矩阵生成模块

动态邻接矩阵生成模块如图2所示,为了使生成的动态邻接矩阵既包含空间相关性,又包含时间相关性,首先采用自相关机制在时间维度对信息进行聚合,然后将聚合后的数据送入空间注意力机制中捕捉数据空间相关性,最后将其作为动态邻接矩阵输出.详情如式(2)所示.

$$G^{\text{dynamic}} = A_s(\text{reshape}(\widehat{A}(\widehat{X}))). \quad (2)$$

式中: A_s 为空间注意力机制; \widehat{A} 为多头自相关机制.具体来说,因为自相关机制是在二维空间上对序列自相关性进行计算并对信息进行聚合,所以首先输入特征 $X \in \mathbf{R}^{N \times D \times T}$, 通过 reshape 函数转换为二维数组 $\widehat{X} \in \mathbf{R}^{T \times M}$, 其中 $M = N \times D$, 将 \widehat{X} 送入自相关机制进行信息聚合,其次使用 reshape 函数将其转换到三维空间 $\mathbf{R}^{N \times D \times T}$, 最后将其送入空间注意力机制,计算得到最终的动态邻接矩阵.

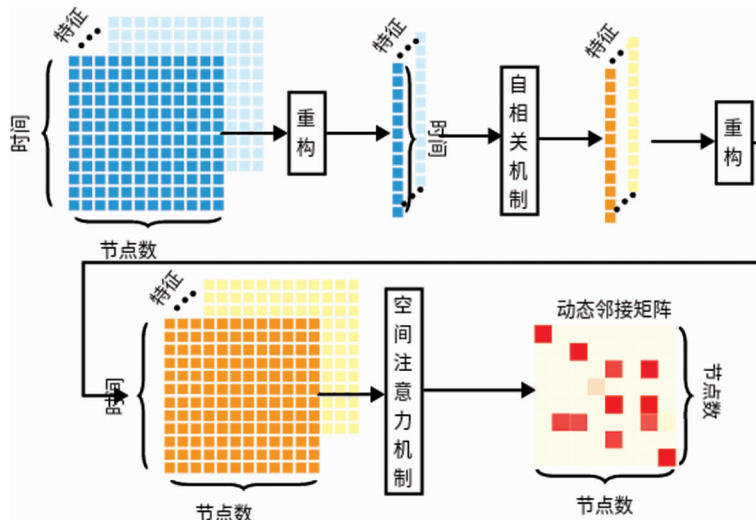


图2 动态邻接矩阵生成模块的结构

1.2.1 时间自相关机制

从时间维度上看,交通流量数据本质上是一个时间序列数据.对于时间序列,可以通过计算序列自相

关函数捕捉其内含的周期性特征^[23].受随机过程理论的启发,对于一个真实的离散时间序列,可以通过式(3)获得自相关系数 $R_{xx}(\tau)$.

$$R_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T X_t X_{t-\tau}. \quad (3)$$

式中: $\{X_t\}$ 为输入自相关机制的时间序列; $R_{xx}(\tau)$ 为序列 $\{X_t\}$ 与滞后 τ 个时间点的序列 $\{X_{t-\tau}\}$ 之间的时延相似性.在自相关机制中,将时延相似性看作置信度的周期估计,即周期长度为 τ 的置信度为 $R_{xx}(\tau)$.

基于 Wiener-Khinchin 理论,对于给定的时间序列 $\{X_t\}$, $R_{xx}(\tau)$ 可以通过快速傅里叶变换得到^[24],计算过程为

$$s_{xx}(f) = F(X_t) F^*(X_t) = \int_{-\infty}^{\infty} X_t e^{-i2\pi f t} dt \overline{\int_{-\infty}^{\infty} X_t e^{i2\pi f t} dt}; \quad (4)$$

$$R_{xx}(\tau) = F^{-1}(s_{xx}(f)) = \int_{-\infty}^{\infty} s_{xx}(f) e^{i2\pi f \tau} df. \quad (5)$$

式中: $\tau \in \{1, \dots, T\}$; F 为快速傅里叶变换 FFT; F^{-1} 为快速傅里叶变换 FFT 的逆变换; $*$ 为共轭运算.

在计算出自相关系数 $R_{xx}(\tau)$ 后,选出最可能的 k 个周期长度 τ_1, \dots, τ_k . 为了实现序列级连接,还需要将相似的子序列信息进行聚合.因此,提出了时延聚合模块如图 3b 所示,自相关机制依据估计出的周期长度,首先使用循环左移操作 Roll 进行信息对齐,然后通过 SoftMax 函数归一化置信度聚合子序列.对于单头情况和长度为 T 的时间序列 $\{X\}$,经过投影变换即可得到 query Q , key K 和 value V ,这样就可以无缝替换自注意力机制.自相关机制的完整计算过程为

$$\tau_1, \dots, \tau_k = \text{argTopk}(R_{Q,K}(\tau)), \tau \in \{1, \dots, T\}; \quad (6)$$

$$\hat{R}_{Q,K}(\tau_1), \dots, \hat{R}_{Q,K}(\tau_k) = \text{SoftMax}(R_{Q,K}(\tau_1), \dots, R_{Q,K}(\tau_k)); \quad (7)$$

$$A(Q, K, V) = \sum_{i=1}^k \text{Roll}(V, \tau_i) \hat{R}_{Q,K}(\tau_i). \quad (8)$$

式中: $\text{argTopk}(\cdot)$ 为计算最可能的 k 个周期长度的函数, $k = \lfloor c \log T \rfloor$, c 为一个超参数; $R_{Q,K}$ 为序列 Q 和 K 之间的自相关系数; $\text{Roll}(X, \tau)$ 为对序列 X 左移 τ 个位置,在此期间,超出第一个位置的元素被重新引入到最后一个位置.

对于多头自相关机制,可以很容易地由单头自相关机制推导得出.对于有 m 个头的自相关机制,计算过程可写为

$$\hat{A}(Q, K, V) = W_{\text{output}} \text{Concat}(h_1, \dots, h_m); \quad (9)$$

$$h_i = A(Q_i, K_i, v_i). \quad (10)$$

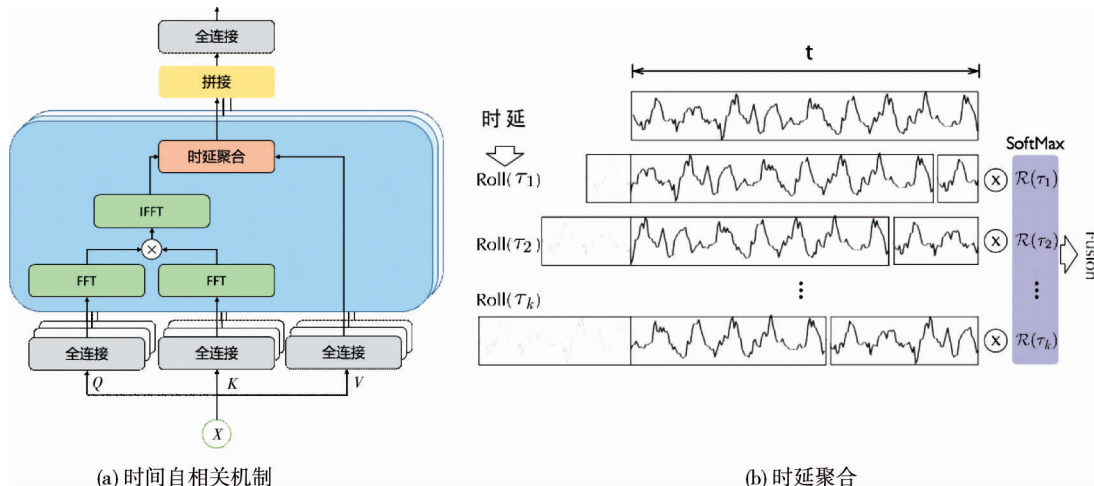


图 3 时间自相关机制与时延聚合示意图

1.2.2 空间注意力机制

在空间维度上,不同地点的交通状况相互影响,且这些相关性是随时间动态变化的.本文使用空间注意力机制^[25]来自适应地捕捉空间维度中的动态相关性.空间注意力可形式化地写为

$$\mathbf{S} = v_s \sigma[(XW_1)W_2(W_3X)]^T + b_s; \quad (11)$$

$$s'_{i,j} = \frac{\exp(s_{i,j})}{\sum_{j=1}^N \exp(s_{i,j})}. \quad (12)$$

式中: $X = (X_1, X_2, \dots, X_T) \in \mathbf{R}^{N \times D \times T}$ 为空间注意力机制的输入; D 为输入数据的通道数; T 为输入数据的时间维度; $v_s, b_s \in \mathbf{R}^{N \times N}$, $W_1 \in \mathbf{R}^T$, $W_2 \in \mathbf{R}^{D \times T}$, $W_3 \in \mathbf{R}^D$ 为可学习参数; σ 为激活函数,本文选择 sigmoid 函数,也可替换为 tanh 或其他激活函数.注意力矩阵 \mathbf{S} 是根据该层的当前输入动态计算的, \mathbf{S} 中元素 $s_{i,j}$ 的值在语义上表示节点 i 和节点 j 之间的相关程度.然后使用 SoftMax 函数来确保节点的注意力权重总和为 1.在进行时空图卷积时,使用时空注意力矩阵 $\mathbf{S}' \in \mathbf{R}^{N \times N}$ 作为邻接矩阵来聚合各节点之间的信息.

1.3 时空图卷积

这里提出的时空图卷积模块^[20]包括空间维度中的图卷积,用于捕获来自邻域的空间依赖关系和沿时间维度的 2 维卷积,利用来自附近时间的时间依赖关系.

图卷积已被证明可用于提取空间特征.KIPF 等^[26]提出切比雪夫谱滤波器^[27]的第一个近似值,它简化了卷积的定义,使其成为一个简单的邻域平均算子.在文献[26]中的图卷积层写为

$$Y = \hat{\mathbf{A}}XW. \quad (13)$$

式中: $X \in \mathbf{R}^{N \times D}$ 和 $Y \in \mathbf{R}^{N \times M}$ 为该层的输入和输出; $W \in \mathbf{R}^{D \times M}$ 为可训练的参数矩阵; $\hat{\mathbf{A}} \in \mathbf{R}^{N \times N}$ 为带自回路的归一化邻接矩阵.本文使用动态邻接矩阵 \mathbf{S}' 替换式(13)的邻接矩阵 $\hat{\mathbf{A}}$,将单个时间片上的数据视为图上的信号,然后直接在图上对图信号进行处理,来捕捉空间中有意义的模式和特征.最后再加入残差连接的空域图卷积的输出可以形式化地写为

$$Y^l = \text{Concat}(Y_1^l, \dots, Y_T^l); \quad (14)$$

$$Y_i^l = \text{Concat}(X_i^l, s'X_i^l W_i^l) \hat{W}_i^l. \quad (15)$$

式中: $X_i^{l-1} \in \mathbf{R}^{N \times M^{l-1}}$ 为第 l 层动态时空卷积层中第 i 个时间片上的数据; $W_i^l \in \mathbf{R}^{D \times M^l}$, $\hat{W}_i^l \in \mathbf{R}^{(M^{l-1}+M^l) \times M^l}$ 为可训练的参数矩阵; $Y_i^l \in \mathbf{R}^{N \times M^l}$ 为第 l 层动态时空卷积层中第 i 个时间片经过邻域信息聚合后的输出; $Y^l \in \mathbf{R}^{N \times M \times T}$ 为第 l 层动态时空卷积层中空域图卷积最终的输出.

通过空域图卷积操作对输入数据的空间特征进行建模之后,再用标准二维卷积捕捉时间维特征.使用 ReLU 函数作为激活函数,得到最终的时空图卷积层的输出 $X^l = \text{ReLU}(\Psi \text{ReLU}(Y^l) + \Psi X^{l-1}) \in \mathbf{R}^{N \times M^l}$.其中, Ψ 为时间维卷积核的参数, X^l 为第 l 层动态时空图卷积层的输出.

1.4 DSTGCN 总体架构

图 4 展示了本文提出的动态时空图卷积网络的模型架构,该架构由动态邻接矩阵生成模块和时空图卷积模块组成动态时空图卷积层,用以对交通流量预测问题的空间依赖性与时间相关性进行建模.其中,动态邻接矩阵生成模块由时间自相关机制和空间注意力机制组成,时空图卷积层由空域图卷积和二维标准卷积组成.通过堆叠多个动态时空图卷积层,DSTGCN 能够处理不同时间级别的时空依赖关系,使用平均绝对误差(Mean Absolute Error, MAE)为模型的损失函数,其定义为

$$L_{\text{MAE}} = \frac{1}{TND} \sum_{i=1}^T \sum_{j=1}^N \sum_{k=1}^D |\hat{Y}_{jk}^{t+i} - Y_{jk}^{t+i}|. \quad (16)$$

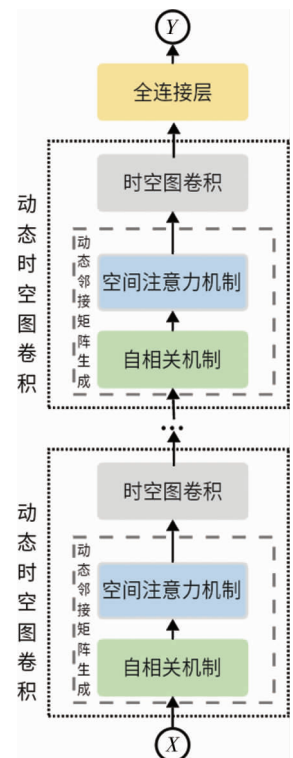


图 4 DSTGCN 模型架构

式中: $\hat{Y} \in \mathbf{R}^{N \times D \times T}$ 为预测的真实值.

2 实验

2.1 数据集

PeMSD4 和 PeMSD8 是 2 个用于交通流量预测的数据集,都是由美国加利福尼亚州的 Caltrans Performance Measurement System (PeMS) 实时采集的高速公路交通数据. PeMS 拥有超过 39 000 个传感器站,部署在加利福尼亚州高速公路系统的主要大都市区.这些传感器站每 30 s 采样一次原始数据,然后汇总成以 5 min 为时间间隔的样本.每个样本包含带时间戳的车流量、平均车速、平均车道占用率这 3 个维度的特征及采集这些信息的检测器的地理位置信息.

PeMSD4 数据集包含 2018 年 1 月 1 日—2018 年 2 月 28 日旧金山湾区 307 个环形检测器的交通数据,涵盖 29 条道路,总长度约为 137.8 km.

PeMSD8 数据集包含 2016 年 7 月 1 日—2016 年 8 月 31 日圣贝纳迪诺地区 170 个环形检测器收集的交通数据,涉及 8 条道路,总长度约为 19.2 km.

2.2 数据预处理

为了优化数据质量,去掉了一些不必要的检测器,使相邻检测器的距离超过 5.6 km,最终得到 2 个数据集(见表 1):PeMSD4 包含 307 个检测器,PeMSD8 包含 170 个检测器.交通数据从 30 s 的间隔聚合为 5 min 的样本,即每个检测器每天拥有 288 个数据点.此外,数据通过零均值归一化 $x' = x - \text{mean}(x)$ 处理,使平均值为 0.最后,本文实验部分将前 70% 的时间步划分为训练集,中间的 20% 为验证集,最后的 10% 作为测试集.

表 1 数据集的详细规模信息

数据集	节点数	时间步数
PeMSD4	307	16 992
PeMSD8	170	17 856

2.3 实验参数设置

本实验的运行环境为 Core(TM) i7-9750H CPU @ 2.60 GHz, NVIDIA GeForce GTX1660Ti GPU, 显存 6 GB. 本文基于深度学习框架 Pytorch 实现了 DSTGCN 模型.为了平衡模型的性能和效率,自相关机制中的超参数 c 选择在 1~3,动态时空图卷积层数为 2,空域图卷积使用 32 个相同大小的卷积核,时间卷积使用 16 个相同大小的卷积核,这些卷积核沿空间轴维度为 1,时间轴维度为 3.使用 Adam 优化器训练模型,初始学习率为 0.001,每过 10 次迭代,学习率衰减为原来的 1/3.将 $p=0.3$ 的 Dropout 应用于图卷积层的输出.本文选择的评估指标包括平均绝对误差 (MAE) 和均方根误差 (RMSE),使用这 2 个评价指标来评估真实交通流数据 X 和预测结果 Y 之间的差异,因此其值越小,预测效果越好.缺失值被排除在训练和测试之外.

2.4 基准方法

为了测试模型性能,将 DSTGCN 与以下模型进行比较:

IHA: 历史平均值法,使用最近交通状况的平均值作为预测值.在本文中,使用 12 个时间片的平均值来预测下一个时间片的值.

IARIMA^[4]: 自回归移动平均法,带有卡尔曼滤波器的自回归移动平均模型,属于时间序列预测模型.

IVAR^[28]: 向量自回归模型,是一种更高级的时间序列模型,可以捕捉所有交通流序列之间的成对关系.

ILSTM^[29]: 长短期记忆网络,一种特殊的循环神经网络模型,常用于预测时间序列.

IGRU^[30]: 门控循环单元网络,一种特殊的循环神经网络模型.

IGeoMAN^[31]: 为地理感知时间序列预测问题提出的基于多级注意力的循环神经网络模型.

ISTGCN^[20]: 基于空间方法的时空图卷积模型.

IASTGCN^[25]: 基于时空注意力机制的时空图卷积模型.

2.5 实验结果及分析

本文将所提出的模型 DSTGCN 在 PeMSD4 和 PeMSD8 这 2 个真实数据集上与其他 8 种基准方法进行了比较.表 2 显示了在未来 1 h 交通流预测性能的平均结果.

表 2 不同方法在数据集 PeMSD4 和 PeMSD8 上的性能比较

模型	PeMSD4		PeMSD8	
	MAE	RMSE	MAE	RMSE
HA	36.76	54.14	29.52	44.03
ARIMA	32.11	68.13	24.04	43.30
VAR	33.76	51.73	21.41	31.21
LSTM	29.45	45.82	23.18	36.96
GRU	28.65	45.11	23.18	36.96
STGCN	25.15	38.29	18.88	27.87
GeoMAN	23.64	37.84	17.84	28.91
ASTGCN	21.80	32.82	16.63	25.27
DSTGCN(Ours)	19.64	31.54	16.26	24.64

从表 2 可以明显地看出:在 2 个数据集上,本文提出的模型在所有评估指标上都取得了最佳性能.可以观察到,传统的时间序列分析方法的预测结果通常并不理想,这表明对于交通数据这类具有非线性时间趋势和复杂的空间依赖关系的时间序列,其建模能力有限.相对于传统时间序列预测方法,传统的深度学习模型(如 LSTM 和 GRU),得益于其强大的对非线性的时间趋势建模的能力,获得了更好的预测结果.此外,同时考虑时间和空间相关性的深度学习模型(STGCN 和 GeoMAN)进一步刷新了预测精度.在 STGCN 的基础上,ASTGCN 引入了时空注意力机制,提升了模型在时空相关性建模方面的能力.本文提出的结合自相关机制与空间注意力机制生成动态邻接矩阵的方法进一步减少了预测误差.图 5 为随机选择一个节点并绘制出 DSTGCN 在第 1 个时间步(即未来 5 min)的预测结果和真实值的对比图,从图 5 中可以直观地观察到 DSTGCN 模型很好地捕捉到了交通数据的变化趋势.

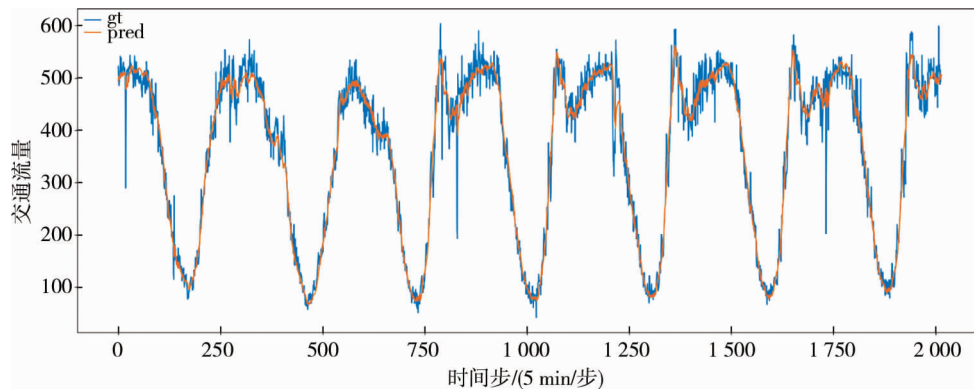


图 5 PEMSD8 数据集上的预测结果

图 6 显示了随着预测区间的增加,各种方法预测性能的变化.总体而言,随着预测区间的增大,相应的预测难度也越来越大,导致误差也随之增大.可以看到,传统的时间序列预测方法和只考虑时间依赖的深度学习方法在短期预测中取得了较好的预测结果,但随着预测区间的增大,其预测误差直线上升.而同时考虑时间与空间依赖关系的方法预测误差增长速度较为缓慢.本文提出的模型在短期预测和长期预测方面都取得了最好的结果,这是由于该模型结合时间自相关机制和空间注意力机制生成的动态邻接矩阵能更好地对当前时刻的时空关系进行建模.

最后,为了观察动态邻接矩阵生成模块在不同时刻生成出的时空图的差别,图 7 展示了在 PeMSD8 数据集不同时刻生成的动态时空图的热力图.为了使展示效果更加清晰,本文节选了所生成的动态邻接矩阵中的 50 个节点.可以从所生成的时空图发现,道路中的大多数节点在同一时刻只与少数的几个节点呈强相关,对剩下的多数节点呈弱相关.这表明,在同一时刻,少数的几个节点会影响整体的道路网络,剩下的

大部分节点在预测中仅仅起了一个很微弱的作用.另外,还可以看出,在相近时刻所生成出的时空图差别不大,这表明时空相关性具有短期不变的特性.

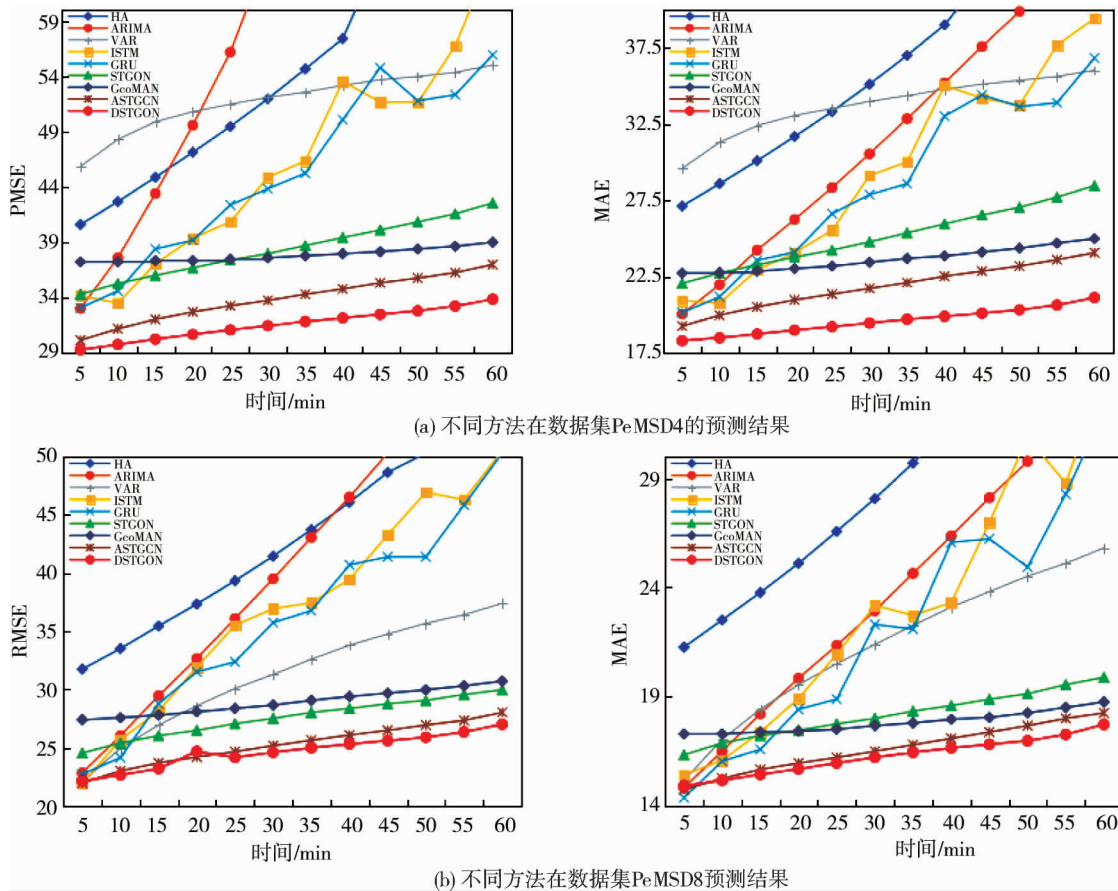


图 6 不同方法随着预测时长的性能变化情况

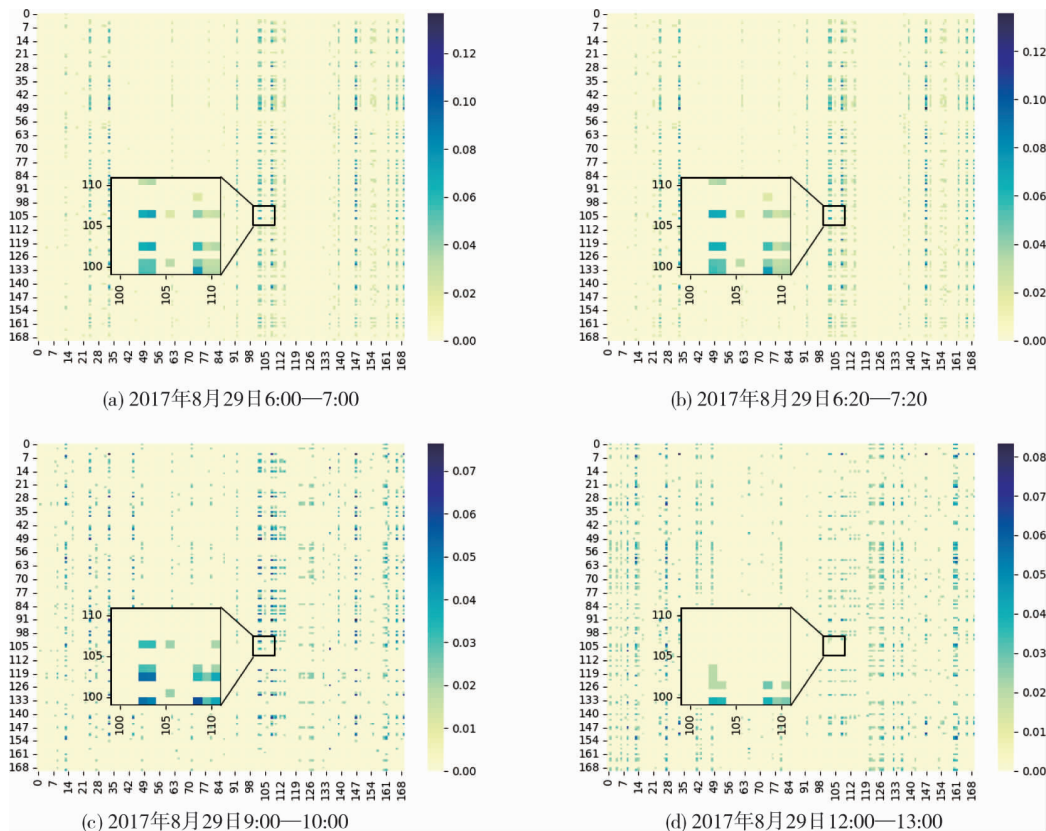


图 7 PeMSD8 数据集上不同时刻交通数据所生成的动态邻接矩阵

3 结论

1) 提出融合空间注意力机制和时间自相关机制来捕捉交通数据中的动态时空相关性的方法,基于该方法可以生成动态邻接矩阵.

2) 提出一种时空图卷积层,使用图卷积和标准二维卷积高效地聚合信息,并通过堆叠多个时空图卷积层来捕捉不同时间级别的时空依赖关系.

3) 在上述基础上,提出一种基于动态时空图卷积的交通流预测方法.

4) 所提出的 DSTGCN 模型在各项评价指标中均优于现有的基准方法,并且与当前较新的 GeoMAN 和 ASTGCN 模型相比,MAE 分别降低了 4.00 和 2.16.

参考文献:

- [1] 陈美林, 郑治豪, 郭宝, 等. 基于因果关联的交通拥堵传播分析[J]. 中南大学学报(自然科学版), 2020, 51(12): 3575-3583.
- [2] LV Z J, XU J J, ZHENG K, et al. LC-RNN: a deep learning model for traffic speed prediction[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. ACM, 2018: 3470-3476.
- [3] ZHENG C P, FAN X L, WEN C L, et al. DeepSTD: mining spatio-temporal disturbances of multiple context factors for citywide traffic flow prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(9): 3744-3755.
- [4] LI Y G, YU R, SHAHABI C, et al. Diffusion convolutional recurrent neural network: data-driven traffic forecasting[EB/OL]. 2017: arXiv: 1707.01926. <http://arxiv.org/abs/1707.01926.pdf>
- [5] Mohammed Shahgir Ahmed, Allen Rusty Cook. Analysis of freeway traffic time-series data by using box-jenkins techniques[J]. Transportation Research Record, 1979. URL: <http://onlinepubs.trb.org/Onlinepubs/trr/1979/722/722-001.pdf>
- [6] WILLIAMS B M, HOEL L A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results[J]. Journal of Transportation Engineering, 2003, 129(6): 664-672.
- [7] SHEKHAR S, WILLIAMS B M. Adaptive seasonal time series models for forecasting short-term traffic flow[J]. Transportation Research Record: Journal of the Transportation Research Board, 2007, 2024(1): 116-125.
- [8] LI X L, PAN G, WU Z H, et al. Prediction of urban human mobility using large-scale taxi traces and its applications[J]. Frontiers of Computer Science, 2012, 6(1): 111-121.
- [9] MOREIRA-MATIAS L, GAMA J, FERREIRA M, et al. Predicting taxi-passenger demand using streaming data[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(3): 1393-1402.
- [10] 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(2): 871-882.
- [11] WAGNER-MUNS I M, GUARDIOLA I G, SAMARANAYKE V A, et al. A functional data analysis approach to traffic volume forecasting[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(3): 878-888.
- [12] LI Z Y, SERGIN N D, YAN H, et al. Tensor completion for weakly-dependent data on graph for metro passenger flow prediction[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 4804-4810.
- [13] DUAN P B, MAO G Q, LIANG W F, et al. A unified spatio-temporal model for short-term traffic flow prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(9): 3212-3223.
- [14] SHIN J, SUNWOO M. Vehicle speed prediction using a Markov chain with speed constraints[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(9): 3201-3211.
- [15] 邴其春, 龚勃文, 林赐云, 等. 基于粒子群优化投影寻踪回归模型的短时交通流预测[J]. 中南大学学报(自然科学版), 2016, 47(12): 4277-4282.
- [16] MA X L, DAI Z, HE Z B, et al. Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction[J]. Sensors, 2017, 17(4): 818.
- [17] ZHANG J B, ZHENG Y, QI D K, et al. Predicting citywide crowd flows using deep spatio-temporal residual networks[J].

Artificial Intelligence, 2018, 259: 147–166.

- [18] WU Z H, PAN S R, LONG G D, et al. Graph wavenet for deep spatial-temporal graph modeling[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. ACM, 2019: 1907–1913.
- [19] 姜山, 丁治明, 朱美玲, 等. 面向时空图建模的图小波卷积神经网络模型[J]. 软件学报, 2021, 32(3): 726–741.
- [20] YU B, YIN H T, ZHU Z X. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting [C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. ACM, 2018: 3634–3640.
- [21] 冯宁, 郭晟楠, 宋超, 等. 面向交通流量预测的多组件时空图卷积网络[J]. 软件学报, 2019, 30(3): 759–769.
- [22] CHEN C, PETTY K, SKABARDONIS A, et al. Freeway performance measurement system: mining loop detector data[J]. Transportation Research Record: Journal of the Transportation Research Board, 2001, 1748(1): 96–102.
- [23] WU H X, XU J H, WANG J M, et al. Autoformer: decomposition transformers with auto-correlation for long-term series forecasting[J]. ArXiv e-Prints, 2021: arXiv: 2106.13008.
- [24] WIENER N. Generalized harmonic analysis[J]. Acta Mathematica, 1930, 55(1): 117–258.
- [25] GUO S N, LIN Y F, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 922–929.
- [26] KIPF T, WELING M. Semi-supervised classification with graph convolutional networks[C]// 5th International Conference on Learning Representations. Toulon, France: ICLR, 2016: 3144218
- [27] DEFFERRARD M, BRESSON X, VANDERGHEYNST P. Convolutional neural networks on graphs with fast localized spectral filtering[J]. ArXiv e-Prints, 2016: arXiv: 1606.09375.
- [28] Zivot E, Wang J. Vector autoregressive models for multivariate time series[M]//Modeling Financial Time Series with S-PLUS ®. New York: Springer-Verlag, 2006: 385–429.
- [29] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735–1780.
- [30] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. ArXiv e-Prints, 2014: arXiv: 1412.3555.
- [31] LIANG Y X, KE S Y, ZHANG J B, et al. GeoMAN: multi-level attention networks for geo-sensory time series prediction [C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. ACM, 2018: 3428–3434.