

SCFNet:一种面向时空预测的外部空间特征融合框架

刘腾飞, 陈李越, 房江祎, 王乐业

引用本文

刘腾飞, 陈李越, 房江祎, 王乐业. SCFNet:一种面向时空预测的外部空间特征融合框架[J]. 计算机科学, 2025, 52(4): 110-118.

LIU Tengfei, CHEN Liyue, FANG Jiangyi, WANG Leye. SCFNet:Fusion Framework of External Spatial Features for Spatio-temporal Prediction [J]. Computer Science, 2025, 52(4): 110-118.

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)

[面向无人机航拍图像的YOLOv8目标检测改进算法](#)

An Improved YOLOv8 Object Detection Algorithm for UAV Aerial Images

计算机科学, 2025, 52(4): 202-211. <https://doi.org/10.11896/jsjx.240500042>

[基于虹膜纹理感知的自适应关联学习方法](#)

Adaptive Contextual Learning Method Based on Iris Texture Perception

计算机科学, 2025, 52(4): 185-193. <https://doi.org/10.11896/jsjx.250100022>

[基于边缘约束和改进Swin Unetr的复杂器官分割方法](#)

Complex Organ Segmentation Based on Edge Constraints and Enhanced Swin Unetr

计算机科学, 2025, 52(4): 177-184. <https://doi.org/10.11896/jsjx.240600007>

[基于群体投票的移动性数据驱动地点类别推测](#)

Mobility Data-driven Location Type Inference Based on Crowd Voting

计算机科学, 2025, 52(3): 169-179. <https://doi.org/10.11896/jsjx.240600164>

[基于注意力机制与对比损失的单视图草图三维重建](#)

3D Reconstruction of Single-view Sketches Based on Attention Mechanism and Contrastive Loss

计算机科学, 2025, 52(3): 77-85. <https://doi.org/10.11896/jsjx.240200102>

SCFNet:一种面向时空预测的外部空间特征融合框架

刘腾飞 陈李越 房江祎 王乐业

高可信软件技术教育部重点实验室(北京大学) 北京 100871

北京大学计算机学院 北京 100871

(2401112021@stu.pku.edu.cn)

摘要 道路信息与当前道路的流量模式息息相关,丰富的POI(Point of Interest)语义可以揭示一个地区的属性,人口数据可以揭示一个地区的人口流量趋势。在时空预测中考虑以上外部空间特征对流量带来的影响,可以帮助模型完成更精准的预测。现有的外部空间建模方法通常针对输入的外部空间特征,经过神经网络映射学得空间相关语义表示,再与最终的时空流量表示融合。然而,由于流量表示和空间特征之间具有异构性,已有的外部空间特征建模方法往往扩展性不高,只能针对特定外部空间特征或特定时空模型。为解决以上问题,提出了一种针对外部空间特征的通用建模框架SCFNet(Spatial Context Fusion Network for Traffic Forecasting)。具体而言,引入基于信息交互的注意力机制,在时空表示和外部空间特征之间计算注意力分数,从而实现外部空间特征和时空表示的高效融合;同时,设计了一种时间向量动态编码方式,以生成动态的空间特征语义。SCFNet采用模块化设计,能够与各类最新的时空流量预测网络结合。SCFNet支持区域人口数据、道路信息、POI等不同空间静态特征的混合输入。在3个真实交通数据集上进行了实验,实验结果表明,SCFNet可显著提高各类最新时空预测方法(如MTGNN,ASTGCN,GraphWaveNet)的预测精度。

关键词: 时空预测; 外部特征建模; 兴趣点; 道路特征; 人口数量; 注意力机制; 时变语义

中图分类号 TP183

SCFNet: Fusion Framework of External Spatial Features for Spatio-temporal Prediction

LIU Tengfei, CHEN Liyue, FANG Jiangyi and WANG Leye

Key Laboratory of High Confidence Software Technologies(Peking University), Ministry of Education, Beijing 100871, China

School of Computer Science, Peking University, Beijing 100871, China

Abstract Road information is closely related to the current traffic pattern of roads. Rich POI semantics can reveal the attributes of an area. Demographic data reveals the trend of population flow in an area. Considering the influence brought by the above external spatial features on the flow in spatio-temporal prediction can help the model accomplish more accurate prediction. Existing external spatial modeling methods usually focus on the input external spatial features, learn spatially relevant semantic representations through neural network mapping, and then fuse them with the final spatiotemporal flow representations. However, due to the heterogeneity between flow representations and spatial features, the existing external spatial feature modeling methods are often not highly scalable and can only target specific external spatial features or specific spatio-temporal models. To overcome the above problems, we propose a spatial context fusion network for traffic forecasting(SCFNet). Specifically, we introduce an attention mechanism based on information interaction to compute attention scores between spatio-temporal representations and external spatial features to achieve an efficient fusion of external spatial features and spatio-temporal representations, and we design a dynamic encoding method of time vectors to generate dynamic spatial feature semantics. SCFNet supports a mixture of different spatial static features such as regional demographic data, road information, and POI inputs. We conduct experiments on three real traffic datasets and demonstrated that SCFNet can significantly improve the prediction accuracy of various state-of-the-art spatio-temporal prediction methods such as(MTGNN, ASTGCN, and GraphWaveNet).

Keywords Spatio-temporal forecasting, External feature modeling, Point of interest, Road feature, Demographic, Attention mechanism, Time-varying semantics

1 引言

时空流量预测是一项重要的研究任务,旨在预测城市多个位置未来时刻的流量。通过结合历史流量数据、领域

知识(例如先验图结构、人口流量分布等)和外部特征(例如天气特征、地区属性特征),时空流量预测可以帮助城市交通管理者和驾驶者预测交通拥堵^[1]、规划路线^[2]以及优化交通信号^[3]等,对于构建智慧城市服务具有重要意义。

在时空流量预测中,外部空间信息,如道路信息、POI 和区域人口信息对于构建交通流之间的空间相关性都具有重要的作用。例如,人口数量和分布情况可以反映出区域内的交通需求。人口密集的地区通常会有更大的交通流量。道路级别(如主要道路、次级道路、高速公路等)通常与道路的设计标准及交通容量相关。了解道路级别,可以估计不同道路能够容纳的交通流量,从而辅助交通流量预测。POI 通常指代具有商业、文化、社交等各种吸引人们关注和兴趣的地点,其可以揭示不同地区的活动热点和人群流动趋势^[4],从而辅助时空流量预测^[5]。

近年来,外部空间特征作为重要的外部信息源被广泛地用于时空流量预测,以提升预测效果。Deng 等^[6]在道路网络上使用矩阵分解来捕获区域间的相关性,从而预测交通流量。Yuan 等^[7]根据道路级别的不同将地图划分为不同的区域,模型在不同区域内推理区域流量。Zhang^[5]等针对 POI 矩阵进行奇异值分解,将地区 POI 的特征向量作为外部特征,与模型结合进行预测。Zheng 等^[8]将 POI 频率、密度等统计结果作为外部特征,与模型结合进行预测。Xia 等^[9]计算区域 POI 的 TFIDF(Term Frequency Inverse Document Frequency),使用计算结果作为特征进行聚类。上述方法均是对空间特征进行统计和解析,学得固定的空间表示后用来辅助模型预测。然而,上述方法只考虑了外部空间特征的静态特性,即假设在不同时间段外部空间特征的语义相同。但实际上,与流量相关的空间信息语义是随着时间变化的,例如白天的商场人很多,但是凌晨的商场人很少。而关于时变的外部空间语义在现有工作中已经开始被考虑,例如 Zeng 等^[10]考虑到了时间动态的空间语义,通过计算每个网格的 POI 数量以及种类,得到不同时刻的兴趣点分布图,并通过卷积得到外部空间特征。Yuan 等^[11]对每个区域计算 POI 的 TF-IDF 值,以获取不同时刻 POI 的影响程度。Xing 等^[12]使用一个向量来表示每种 POI 分布,同时提取时间特征作为时间向量,将 POI 向量和时间向量拼接作为节点特征放入模型进行预测。以上模型都考虑了时间动态的外部空间语义,并将外部空间语义表示和时空表示直接融合,但忽略了外部空间特征和流量特征之间的交互。

此外,已有方法中的外部空间特征建模模块和流量的时空建模模块通常是耦合的,不便于扩展^[12-15]。如果外部空间建模模块与时空建模模块高度耦合,当时空模型中需要添加 POI 或道路网络等外部信息作为输入时,需要对当前时空模型进行重构。但随着深度学习技术的发展,高耦合的建模方式难以受益于最新的技术。相较而言,若能实现模块化的时空框架,当引入新的输入或改变模型的结构时,只需修改相应的模块而不必重新设计整个模型,更加高效和灵活。因此,如何提出一种独立于时空模型之外、针对外部空间特征进行建模的可插拔通用模块,对于提升模型的预测效果具有重要的研究和实践意义。

针对上述问题,本文设计了一个外部空间特征融合框架 SCFNet,其包括外部时变空间语义生成模块和交互注意力空间表征模块。在外部时变空间语义生成模块中,通过所设计的时间向量动态编码方式可以生成动态的外部空间语义;

同时交互注意力空间表征模块可以使外部空间信息的空间表征和时空流量表征进行交互,从而习得更有效的外部空间语义。SCFNet 采用模块化设计,通过与现有时空网络相融合,能够增强各类已有的时空预测方法的性能。具体而言,本文的主要贡献如下:

1)提出了一种面向时空预测的外部空间特征融合建模框架——SCFNet。在 SCFNet 中,我们设计了交互注意力机制在外部空间语义和时空表示之间高效融合,从而提取与流量相关的有效外部空间语义表示。同时,我们设计了一种时间向量动态编码方式,其能够有效地生成动态外部空间语义,获取空间语义的时变特征。

2)SCFNet 仅需与时空模型最后一层的时空表示进行融合,因此可应用于大多数时空预测模型,具备良好的扩展性,在无需重构模型的前提下就能够提高预测效果。SCFNet 也支持区域人口数量、道路级别划分等多种空间外部特征的各种组合输入。

3)在 3 个数据集上进行了实验,对每个数据集,我们选择了多种时空模型作为基准。实验表明,加入 SCFNet 之后,时空模型的 RMSE(Root Mean Square Error)最高可降低 10% 以上。不同数据集上稳定提升的效果证明了 SCFNet 良好的泛化性。

2 相关工作

2.1 基于深度学习的时空预测模型

近年来,随着神经网络的兴起,各种深度学习技术也被广泛应用于时空预测。例如 Wu 等^[16]提出了 GraphWaveNet,并使用堆叠的膨胀 1D 卷积分量来捕获长期序列之间的相关性。Ma 等^[17]将交通网络视为一个图像,并使用卷积神经网络(Convolutional Neural Network, CNN)来预测道路中的交通速度,同时评估拥堵程度。除了 CNN 外,循环神经网络(Recurrent Neural Network, RNN)能够捕捉和建模时间序列数据中的时序依赖性^[18-19]。然而, RNN 在长序列数据训练中容易产生梯度消失问题。而 GRU(Gate Recurrent Unit)和 LSTM(Long Short Term Memory)作为特殊的 RNN,通过引入记忆单元和门控机制,拥有了捕获长期依赖的能力,避免了梯度信息的消失。Do 等^[20]训练了 GRU-NN 单元进行时空特征提取,同时为了捕捉不规则网格的空间拓扑关系,将图神经网络用于空间建模。除 CNN 和 GRU 等技术外,注意力机制也被广泛用于时空预测任务。注意力机制的核心是使用一组动态的权重参数,对感兴趣的区域进行加权。ASTGCN^[21]是一种基于注意力机制的时空图卷积网络模型,通过时空注意力机制捕获动态时空相关性。在 GMAN^[22]中, Zheng 等提出了一种多注意力图神经网络,对未来和历史时间步之间进行建模,缓解了长期预测误差过大的问题。Self-Attention^[23]能够关注输入序列的全局相关特征,使模型更高效,很多工作都将 Self-Attention 应用于时序预测^[24]。以上基于深度学习的时空模型对常规的知识有良好的预测效果,但对外部空间特征的考虑还不足。

2.2 考虑外部空间因素的时空预测模型

各种外部空间因素(如道路信息、POI、人口区域数据)对

于构建交通流之间的空间相关性具有重要的作用。Zheng 等提出的 DeepSTD^[8] 模型, 将 POI 频率和密度的统计结果作为外部特征, 与模型结合进行预测。Deng 等^[6] 在道路网络上使用矩阵分解来捕获区域间的相关性, 从而预测交通流量。Wang 等^[25] 提出了 KXGBoost, 对区域人口数据、交通服务数据等潜在因素进行统计, 识别出地铁需求量的决定性特征, 从而对乘客的出行需求进行预测。Xia 等^[9] 提出的 3DGCN 模型通过计算区域内 POI 的 TD-IDF, 使用计算结果作为特征进行聚类, 划分不同区域。然而上述工作只是对外部空间特征进行了统计和分析, 忽略了外部空间信息的语义是随着时间发生变化的。因此, Zeng 等^[10] 提出的 DeepSTN+ 模型通过计算每个网格的 POI 数量以及种类, 得到不同时刻的 POI 分布图, 并通过卷积得到外部特征。在 STRN^[26] 中, Liang 等通过时间信息和道路信息特征矩阵相乘的方式为 POI 添加时间依赖。Yuan 等^[11] 提出了 STLSN, 该模型对每个区域计算 POI 的 TF-IDF 值, 从而获得不同时刻兴趣点的影响程度。以上工作在预测过程中都生成了动态的外部空间特征时空表示, 但它们都是将模型得到的外部特征表示与时空流量表示直接融合。实际上, 直接融合的方式忽略了外部空间语义表示和交通流量表示的交互。

Zeng 等^[14] 提出的 AttDeepstn+ 采用一种新的融合机制将 POI 信息和注意力模块融合, 用于表现位置属性对人流的影响。Xing 等^[12] 提出的 STGs 使用空间信息作为节点特征, 通过图卷积学习空间知识对流量的影响。Pan 等^[27] 提出了 ST-MetaNet, 其使用 POI 信息作为边和节点的元信息, 从而学习得到边和节点的相关性。以上方式虽然将空间特征有效地嵌入到模型中, 但是空间特征建模模块与时空模型的耦合

程度较高, 难以应用于其他时空模型中, 不具有灵活的可移植性。因此, 我们提出了一种面向时空预测的框架——SCFNet。SCFNet 基于交互注意力模块, 可以完成时空表示和外部空间语义的交互, 提取与流量相关的有效外部空间语义, 并且可以生成动态外部空间语义, 捕获时变空间特征。同时, 该模型具有良好的可移植性, 可与大多数时空模型结合来提升模型的预测效果。

3 问题定义

对时空预测问题的定义, 与文献[28]中的定义相同。首先定义一个站点集合 $\mathcal{L} = \{l_1, l_2, \dots, l_n\}$, n 代表预测站点个数。对每一个站点 l_i 来说, 都存在一个长度为 k 的历史流量序列 \mathbf{F}_i 。 $\mathbf{F}_i = \{f_{i1}, f_{i2}, \dots, f_{ik}\}$, 而我们要预测第 $k+1$ 个时刻的流量 $f_{i(k+1)}$:

$$\text{error}(\hat{f}_{i(k+1)}, f_{i(k+1)}), \forall l_i \in \mathcal{L} \quad (1)$$

其中, $\hat{f}_{i(k+1)}$ 是第 i 个站点在 $k+1$ 时刻的预测值, $f_{i(k+1)}$ 代表第 i 个站点在 $k+1$ 时刻的流量真值。 error 代表预测值和真值之间的预测误差。本文分别采用均方根误差 (RMSE) 和对称平均绝对百分比误差 (SMAPE) 来衡量预测值和真值之间的误差。

4 本文方法

4.1 概述

本文所提出的 SCFNet 是一种面向时空预测的外部空间特征通用融合框架。其包含两个模块: 交互注意力空间表征模块和外部时变空间语义生成模块。模型框架如图 1 所示。

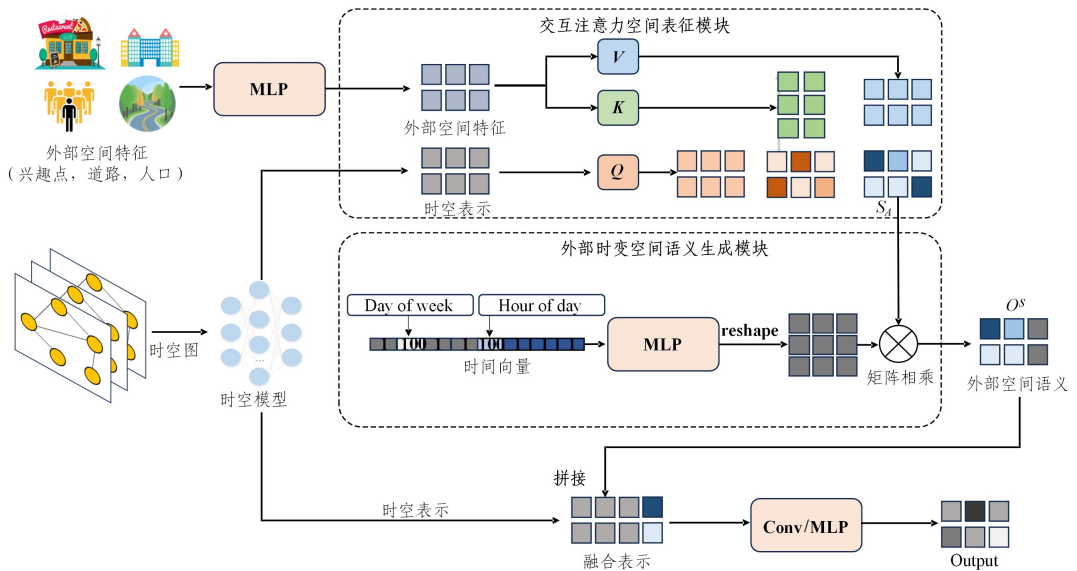


图 1 SCFNet 模型框架

Fig. 1 Structure of SCFNet

交互注意力空间表征模块是为了提取更全面的外部空间语义表示。具体而言, 我们使用 MLP 对输入的外部空间语义进行提取, 接着使用跨模态注意力, 将时空表示作为 Query, 外部空间语义表示分别作为 Key 和 Value, 使得外部空间语义表示和时空表示之间进行交互, 从而保留更多与时空表

示相关的外部空间表征。外部时变空间语义生成模块是为了生成外部空间特征的动态语义表示。具体而言, 该模块中我们提出时间向量动态编码方式将时间向量作为输入, 通过全连接神经网络学习将时间向量转化为动态的时间特征, 然后采用类似 STRN^[26] 中的融合方式, 将时间特征与学到的空间

表征做矩阵乘法,最终获得全面的动态外部空间语义表示。最后,将最终学到的外部空间语义表示与时空模型学到的时空表示相结合,通过卷积层或全连接层得到预测结果。

4.2 交互注意力空间表征模块

外部空间特征包含丰富的空间语义,我们可以从中学到外部空间区域表示,从而辅助模型进行更精准的流量预测。首先,在交互注意力空间表征模块中,我们使用全连接神经网络提取对输入的外部空间特征进行学习:

$$S_o = MLP(S) \quad (2)$$

其中, $S \in \mathbf{R}^{N \times d_1}$ 指输入的外部空间初始特征, $S_o \in \mathbf{R}^{N \times d_o}$ 是经过 MLP(Multilayer Perceptron)学到的空间特征。

提取到空间特征之后,使用交互注意力,使得时空模型的时空表示和外部空间语义特征进行交互并融合,即在时空表示和空间表征之间进行注意力分数计算。具体而言,我们针对查询向量 Q 、键向量 K 和值向量 V 进行设计:

$$\begin{aligned} Q &= XW^Q \\ K &= S_o W^K \\ V &= S_o W^V \end{aligned} \quad (3)$$

其中, W^Q, W^K 和 W^V 是可学习的参数矩阵。 $X \in \mathbf{R}^{N \times d_x}$ 是时空模型学到某时刻的时空表示,其中 d_x 代表时空表示特征数, N 表示站点个数。

因此,这里 V 和 K 是外部空间特征 S_o 的 Value 和 Key 的映射, Q 是时空表示 X 的 Query 映射。

接着,使用注意力分数来加权向量 V :

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

这样设计可以使得外部特征 S_o 和时空表示 X 之间通过计算相似度得到一组注意力权重,并将注意力权重与外部空间语义相乘得到带权重的外部空间表征,从而达到两种特征的交互作用。在模型中,实际上为了提高模型的表达能力,这里使用多头注意力(Multi-Head Attention):

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (5)$$

$$MultiHead(Q, K, V) = Concat(head_1, head_i)W_i^O \quad (6)$$

$$S_A = MultiHead(Q, K, V) \quad (7)$$

其中, $W_i^Q, W_i^K, W_i^V \in \mathbf{R}^{d_{model} \times d_k}$, 而 i 表示第 i 个头,每个头通过分组计算注意力分数之后拼接在一起,并通过一个可训练权重 $W_i^O \in \mathbf{R}^{(h \times d_k) \times d_{model}}$, 得到输出 $S_A \in \mathbf{R}^{N \times d_{model}}$ 。其中 h 是注意力头数; d_{model} 是注意力层内的特征数,一般设置为 512。

多头注意力通过创建多组 Q, K 和 V 权重矩阵,对 Q, K, V 进行多次不同的线性变换,并行计算多个注意力分数和输出,然后将这些输出连接在一起,再通过线性变换得到最终的输出。这样可以使模型在不同表示子空间学习到信息。最终在该模块,我们完成了流量表示和外部空间特征的交互,得到了带有权重的外部空间表征。

4.3 外部时变空间语义生成模块

因为一天中不同时间段的外部空间特征代表的语义不总

是相同的,所以时间静态的外部空间特征在细粒度的时空预测中很难充分发挥其作用。为了生成动态的外部空间语义,我们设计了一种在模型中能够有效加入时间信息的编码方式-时间向量动态编码,即对预测时刻的时间进行编码,且时间向量 T_0 由两部分矢量 t_1 和 t_2 组成。

$$T_0 = concatenate(t_1, t_2) \quad (8)$$

矢量 t_1 表示一天中的第几个预测时间段, $t_1 \in \mathbf{R}^{d_1}$, $d_1 = 24/fitness$, 其中 $fitness$ 为预测粒度,单位是小时(h); 矢量 t_2 表示一周中的第几天, $t_2 \in \mathbf{R}^{d_2}$, $d_2 = 7$ 。拼接后得到的时间向量 $T_0 \in \mathbf{R}^{d_1+d_2}$ 。实验发现,向量的数值用 1 或 100 填充时,便于模型从时间向量中更好地提取到时间信息。因此,如果预测时刻属于这一周中第 i 天第 j 个时刻,则 t_1 上第 j 个位置的值为 100,其余位置的值为 1; t_2 上第 i 个位置的值为 100,其余位置的值为 1。

编码好时间向量 T_0 后,将 T_0 通过两层的 MLP 得到 P_0 , 然后 reshape 得到 P :

$$P_0 = MLP(T_0) \quad (9)$$

$$P = reshape(P_0) \quad (10)$$

其中, $T_0 \in \mathbf{R}^{d_1+d_2}$, $P_0 \in \mathbf{R}^{d_p}$, 其中 $d_p = d_{model} * d_t$ 。在得到 P_0 后,将 P_0 reshape 为 $P \in \mathbf{R}^{d_{model} * d_t}$, 然后将外部空间语义 S_A 与 P 相乘:

$$O = Matmul(S_A, P) \quad (11)$$

得到与时空表示预测任务相关的全面外部空间语义表示 $O \in \mathbf{R}^{N \times d}$ 。

得到输出 O 之后,将时空模型的流量表示 X 与 O 拼接,得到拼接结果后通过最后一层卷积层或全连接层,得到最终预测输出 $Output$ 。

$$O_i = Concat(X; O) \in \mathbf{R}^{N \times (d+d_{model})} \quad (12)$$

$$Output = f_{fusion}(O_i) \quad (13)$$

5 本文实验

5.1 数据集

本次实验涉及 3 个数据集,分别为 Bike NYC、Metro NYC 和 Taxi Chicago 数据集。

实验所使用的数据集包含了网约车数据集、共享单车数据集和地铁数据集 3 种类型,实验涉及预测 30 min, 60 min, 120 min 间隔的流量。

BikeNYC¹⁾ 数据集是从纽约市自行车共享系统收集的共享单车流量预测数据集。其记录了 820 个站点,包括行程编号、起止时间、经纬度、自行车 ID 等信息。时间为 2013 年 7 月 1 日至 2017 年 9 月 30 日,主要任务为单车流量预测。

Metro NYC²⁾ 数据集由纽约市大都会交通管理局提供,其提供纽约市每小时不同地铁站的地铁乘客量估算。该数据集包括 426 个地铁站点,包含 2022 年 2 月 1 日至 2023 年 12 月 21 日的流量,主要任务为地铁客流量预测。

Taxi Chicago³⁾ 数据集是从芝加哥市网约车共享系统收

¹⁾ <https://www.citibikenyc.com/system-data>

²⁾ <https://github.com/uctb/Urban-Dataset>

³⁾ <https://data.cityofchicago.org/Transportation/Taxi-Trips>

集的数据。该数据集记录了 77 个站点,时间范围为 2013 年 1 月 1 日至 2018 年 1 月 1 日,主要任务为网约车需求预测。

5.2 测试场景划分

为了详细分析不同区域加入外部空间特征时的预测情况,我们将评测场景划分为中心区域和非中心区域。通常中心区域的站点人口更密集,交通流量更大,非中心区域的交通流量相对更小。如图 2 所示,红色区域为城市的中心区域,蓝色和绿色区域为非中心区域。对于纽约市,我们将曼哈顿行政区作为中心区,其余行政区作为非中心区。对于芝加哥市,我们将中心商务区(Central Business District, CBD)作为中心区域,其余地区作为非中心区。地图边界数据来源于谷歌地图。



图 2 不同城市测试场景划分可视化(电子版为彩图)

Fig. 2 Visualization of test scenario partitioning in different cities

5.3 POI 数据获取

本文收集了所使用数据集相应年份的 POI 数据,所有的 POI 数据收集自开源网站 OpenStreetMap。通过对每个城市中重要的 POI 种类进行筛选,得到了重要的 POI 种类,同时将芝加哥市和纽约市两个城市的 POI 种类合并,最终选择了 35 种 POI 作为本文实验的外部特征。表 1 列出了所筛选出的部分 POI 种类,图 3 展示了所收集的纽约市和芝加哥市 2013 年的 POI 分布。

表 1 部分 POI 种类

Table 1 Partial POI categories

POI 种类
Fast food
Bank
School
Kindergarten
Bar
Bench
Bicycle parking
Nightclub
Car sharing
Bench
Theatre
Pub

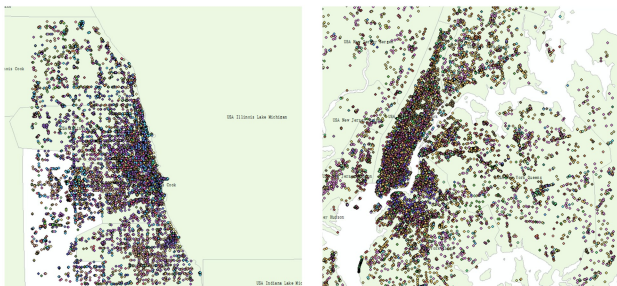


图 3 纽约市和芝加哥市的 POI 可视化

Fig. 3 Visualization of POI distribution in NYC and Chicago

5.4 区域人口数据和道路级别数据收集

我们收集的区域人口数据来自多个城市、不同年份。由于官方大规模的人口普查在美国 10 年举行一次,考虑到使用数据集的年份,我们收集了 2010 年和 2020 年不同城市的人口普查数据。其中,芝加哥市的人口数据来自芝加哥政府官网提供的 2010 年的人口普查数据;纽约市的人口数据来自于纽约政府官网提供的 2020 年的人口普查数据。不同城市的人口普查区域划分如图 4 所示。

道路级别划分数据收集于 OpenStreetMap 网站,路网数据由于整体在短时间内变化不大,因此我们使用当年 1 月 1 日的路网表示当年的路网情况。比如使用 2013 年 1 月 1 日收集到的路网数据作为 Metro NYC 2013 年的路网信息,路网数据在 Arcgis 软件上可视化如图 4 所示。

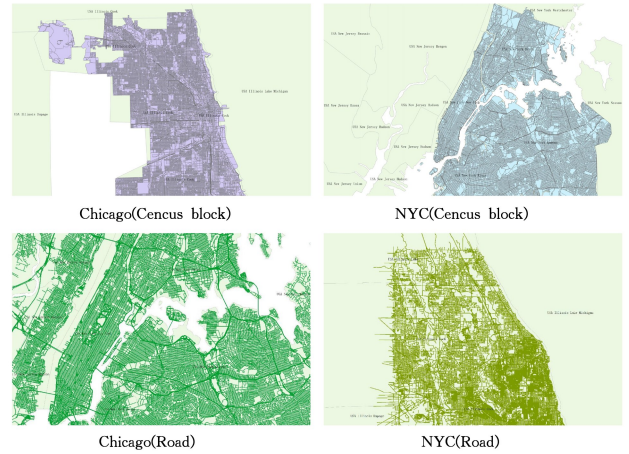


图 4 不同城市区域人口信息和道路信息可视化

Fig. 4 Visualization of population and road networks in different areas

5.5 评价指标

本文采用两种评价指标来评估模型的性能,分别是均方根误差(RMSE)和平均绝对百分比误差(SMAPE)。

1)均方根误差 RMSE:

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (14)$$

其中, y_i 代表真值, \hat{y}_i 代表预测值, N 代表站点个数。RMSE 衡量的是预测值和真实值之间的偏差,因此 RMSE 越小越好。

2)平均绝对百分比误差(SMAPE):

$$SMAPE(y, \hat{y}) = \frac{100\%}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{\frac{|\hat{y}_i| + |y_i|}{2}} \quad (15)$$

其中, y_i 代表真值, \hat{y}_i 代表预测值。但是使用 SMAPE 时,当真实值和预测值均等于 0 时,计算结果会出现除 0 现象,因此实际应用时为了避免除 0 现象的发生,我们将设置阈值 θ ,在预测值或真值小于 θ 时,将值截断为 θ 。

5.6 基准模型

本文提出的 SCFNet 可与时空模型结合来提高预测精度,在实验中将其与以下 3 种已有的时空流量预测模型进行融合和比较。

1) ASTGCN^[21]是一种基于注意力机制的时空图卷积网络模型,其通过时空注意力机制来捕获动态时空相关性,在空间和时间维度上进行卷积操作,实现对复杂时空数据的建模与预测。

2) MTGNN^[29]是一种针对多变量时序预测问题的通用框架,其通过图学习模块学习到变量间的单向关系构图,然后通过图卷积模块和时空卷积模块来捕获时间序列间的时空依赖特征。

3) GWN^[16]是一种基于自适应图的时空模型。该模型使用节点数据构建自适应图,在数据中捕捉隐藏的空间依赖性,并进一步预测结果,是一种端到端的模型。

5.7 实验设置

我们基于 Pytorch1.10 框架实现了 SCFNet。所有实验

均在 GeForce RTX 4070Ti 上进行训练,训练时均采用 Adam 优化器,初始学习率设置为 1×10^{-5} ,Batchsize 设为 32。Attention 的注意力头数为 4,注意力层数为 6。最终输出外部空间语义表示的特征数为 10。在数据集方面,我们使用 Bike NYC 和 Taxi Chicago 数据集 2013 年的数据,以及 Metro NYC 数据集 2022 年的数据进行训练和测试,同时对比模型和 SCFNet 均采用从头训练方式,训练集、验证集和测试集按时间从前往后划分,比例为 7:1:2。

5.8 实验结果及分析

5.8.1 主实验结果

我们将 SCFNet 模块分别添加在 3 种模型上,并与原始模型进行对比。表 2 列出了不同模型在未来 60 min 内流量的预测结果,粗体代表最佳结果。

表 2 添加 SCFNet 前后的模型预测结果

Table 2 Prediction results of the models before and after adding SCFNet

	Bike NYC(60 min)				Metro NYC(60 min)				Taxi Chicago(60 min)			
	中心区域		非中心区域		中心区域		非中心区域		中心区域		非中心区域	
	RMSE	SMAPE	RMSE	SMAPE	RMSE	SMAPE	RMSE	SMAPE	RMSE	SMAPE	RMSE	SMAPE
MTGNN	4.0401	0.3492	1.7482	0.2658	155.2700	0.2131	78.487	0.2254	111.8100	0.18847	15.1370	0.0496
MTGNN(+SCFNet)	3.8453	0.3346	1.7021	0.2626	137.1600	0.1993	75.509	0.2016	95.7110	0.1649	13.5690	0.0476
GWN	3.8705	0.3435	1.7184	0.2573	136.6600	0.1911	78.789	0.1552	115.3000	0.1854	13.8120	0.0441
GWN(+SCFNet)	3.7888	0.3270	1.7008	0.2505	121.5600	0.1428	74.094	0.1491	83.6320	0.1519	11.4760	0.0424
ASTGCN	3.9572	0.3476	1.8050	0.2659	151.4000	0.2140	83.080	0.2124	108.6700	0.2109	14.7850	0.0618
ASTGCN(+SCFNet)	3.9157	0.3512	1.7314	0.2630	147.4300	0.2122	81.238	0.2147	99.8150	0.1599	12.9360	0.0476

从表 2 中可以看出,在 3 个模型中加入 SCFNet 之后,相较于未使用 SCFNet 的模型均有明显提升。3 个数据集上加入 SCFNet 的 GWN 整体预测效果最好。

根据以上结果,可得出结论:

1) 通过纵向对比模型针对不同数据集的预测情况,在 SCFNet 引入动态外部空间语义后,在不同类型的预测任务中都可以辅助提高模型的预测性能。

2) 横向对比 3 个时空模型在同一个数据集上的预测情况,可以看到不同模型虽然对同一个任务的预测情况不同,但是加入 SCFNet 之后对最终结果都有不同程度的提升。

3) 对比 3 个数据集可以发现,不同的数据集的提升效果也是不同的。整体而言,我们发现数据集提升的程度与数据集的站点间流量水平的波动程度呈正相关性。本文用离散系数来衡量数据集中不同站点流量水平的波动程度,如表 3 所列。

从表 2 和表 3 可以发现,Metro NYC 数据集和 Taxi NYC 数据集站点流量的波动程度明显高于 Bike NYC,而 Metro NYC 和 Taxi Chicago 中心区域的 RMSE 提升整体在 10% 以上,但 Bike NYC 中心区域的 RMSE 提升整体只在 5% 左右。

表 3 不同数据集站点间流量水平的离散系数

Table 3 Coefficients of variation about site traffic in different

datasets	
数据集	离散系数(C)
Bike NYC	0.89
Metro NYC	1.31
Taxi Chicago	3.42

5.8.2 不同时间粒度 SCFNet 对预测的影响

表 4 展示了在不同时间粒度(30 min/60 min/120 min)下,Bike NYC 数据集上模型的预测效果。可以看出,在不同时间

粒度下加入 SCFNet 对中心区域的预测结果均有不同程度的提升。在较短的时间粒度下(例如 30 min/60 min)对非中心区域也有不同程度的提升。然而在 120 min 的时间粒度下,发现非中心区域的 RMSE 不降反升,这说明在更长跨度的时间粒度的预测中,加入外部特征并不会对非中心区域的预测结果起到提升效果。

表 4 不同时间粒度下对比模型的效果

Table 4 Comparing model performance at different granularities

时空模型	Bike NYC			
	中心区域		非中心区域	
	RMSE	SMAPE	RMSE	SMAPE
时间粒度=30 min				
MTGNN	2.3794	0.3066	1.1072	0.1681
MTGNN(+SCFNet)	2.3283	0.2981	1.1021	0.1634
时间粒度=60 min				
MTGNN	4.0401	0.3492	1.7482	0.2658
MTGNN(+SCFNet)	3.8453	0.3346	1.7021	0.2626
时间粒度=120 min				
MTGNN	6.8632	0.3125	2.7887	0.2269
MTGNN(+SCFNet)	6.5894	0.2976	2.8431	0.2267

5.8.3 消融实验

为了验证 SCFNet 模块设计的合理性和有效性,从 SCFNet 中移除各个模块来比较模型效果。变体模型如下。结果如表 5 所列。

MTGNN:即模型不使用外部特征,仅使用时空模型 MTGNN 进行预测。

MTGNN(+S):加入交互注意力空间表征模块。

MTGNN(+T):加入外部时变空间语义生成模块,使用时间向量动态编码。

MTGNN(+SCFNet):时空模型 MTGNN 加入 SCFNet 框架。

表5 消融实验

Table 5 Ablation experiment

时空模型	BikeNYC(60 min)			
	中心区域		非中心区域	
	RMSE	SMAPE	RMSE	SMAPE
MTGNN	4.0401	0.3492	1.7482	0.2658
MTGNN(+S)	3.9881	0.3468	1.7295	0.2701
MTGNN(+T)	3.8372	0.3422	1.7162	0.2871
MTGNN(+SCFNet)	3.8487	0.3362	1.6941	0.2650

根据表5得出实验结论:

1) 当仅使用交互注意力空间表征模块 MTGNN(+S) 时,无论是中心区域还是非中心区域, RMSE 均降低 1% 左右; 同样, 当仅使用外部时变空间语义生成模块时, 所有区域的 RMSE 明显降低, 说明该模块可以降低站点整体的预测误差。

2) 单独使用外部时变空间语义生成模块和交互注意力

表6 不同外部特征输入对比

Table 6 Comparison of different external feature inputs

Metro NYC(60 min)							
外部空间特征	时空模型	非中心区域		非中心区域		整座城市	
		RMSE	SMAPE	RMSE	SMAPE	RMSE	SMAPE
无外部空间特征	MTGNN	155.27	0.2131	78.487	0.2254	105.90	0.2220
区域人口数量	MTGNN(+SCFNet)	136.05	0.2045	76.065	0.2172	96.714	0.2136
POI	MTGNN(+SCFNet)	134.72	0.2071	76.836	0.2076	96.714	0.2075
道路级别	MTGNN(+SCFNet)	136.37	0.2129	77.075	0.2167	97.498	0.2156
POI&道路&人口	MTGNN(+SCFNet)	137.16	0.1993	75.509	0.2016	96.927	0.2010
Bike NYC(60 min)							
无外部空间特征	MTGNN	4.0401	0.3492	1.7482	0.2658	3.2307	0.3122
区域人口数量	MTGNN(+SCFNet)	3.8480	0.3371	1.6959	0.2681	3.0845	0.3065
POI	MTGNN(+SCFNet)	3.8487	0.3362	1.6941	0.2650	3.0846	0.3046
道路级别	MTGNN(+SCFNet)	3.8369	0.3385	1.7036	0.2759	3.0787	0.3133
POI&道路&人口	MTGNN(+SCFNet)	3.8453	0.3346	1.7021	0.2626	3.0842	0.3026

在 Metro NYC 数据集上, 将添加不同的空间特征结果进行对比。从表6中可以发现, 添加不同的空间特征对数据集整体的预测结果都有明显提升。将所有空间特征均输入模型中时, 模型的预测效果最好。相比于不使用 SCFNet, 加入所有空间特征时, 模型在整个数据集上 RMSE 降低了 8.57%, SMAPE 降低了 9.45%。这表明本文模型可以在不同的空间特征中提取各自有效的空间信息表示, 并与模型结合辅助进行预测。

同样地, 表6展示了在 Bike NYC 数据集上分别将区域人口数量(Demographic)和道路级别(Road Level)加入 MTGNN 模型中, 并将所有空间特征融合加入到 MTGNN 模型中和未加入外部特征的结果进行对比的结果。从表6中可以发现, 加入其他空间特征对预测结果均有提升, 但是当加入道路级别时, 虽然中心区域的各项指标有所降低, 但明显非中心区域的 SMAPE 比不加入空间特征时更差。而非中心区域的站点通常是低流量站点, 中心区域多为高流量站点。针对这一情况, 绘制整体站点的预测情况和流量均值较低的最后 40 个站点情况并进行分析, 如图5所示。

如图5所示, 其中横轴是测试集连续预测时间片段, 纵轴是该时刻站点流量的均值。可以发现, MTGNN(+Context)站点的整体预测误差很小, 峰值的预测结果也优于不加入空间特征的 MTGNN。然而在低流量站点, MTGNN

空间表征模块时, 对非中心区域的预测 SMAPE 变差, 说明单独加入哪一个模块对低流量站点的预测都不是很好。

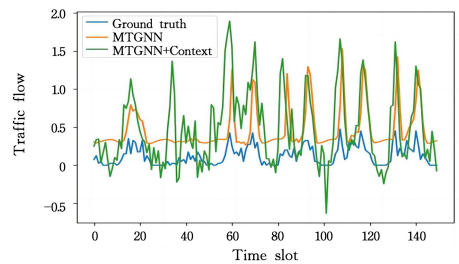
3) 同时使用两个模块时, 模型可以捕捉到更全面有效的外部空间语义, 对中心区域和非中心区域的站点预测都有提升, RMSE 和 SMAPE 都明显降低。

5.8.4 SCFNet 中加入各类外部特征的探究

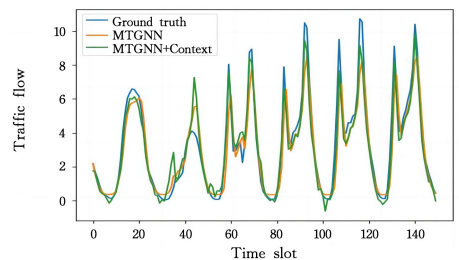
SCFNet 实际上是一个外部空间特征通用的框架, 它支持多种类型的空间特征加入到模型中, 同时支持多空间特征输入。这里我们使用 Metro NYC 和 Bike NYC 数据集来探究不同外部特征输入的预测结果。

表6展示了在 Metro NYC 数据集上分别将区域人口数量(Demographic)和道路级别(Road Level)加入到 MTGNN 模型中, 并将所有空间特征融合加入到 MTGNN 模型中和未加入外部特征的结果进行对比。

(+Context)的预测结果常常出现过高的情况, 而 MTGNN 较为保守的预测反而更符合低流量站点的趋势, 误差更小。



(a) 流量低站点真值/预测值对比



(b) 所有站点真值/预测值对比

图5 在只加入道路信息时, 流量低站点/所有站点真值和预测值对比

Fig. 5 Comparison of true and predicted values at low flow station pairs and all station pairs when road information is added

虽然在加入道路级别信息时,非中心区域的 SMAPE 增加,但是当所有空间特征融合输入时,可以发现模型比不使用空间特征的结果更好,并且比加入任何单个空间特征的结果更好。这证明了 SCFNet 可以在多种空间表示中减小无效的空间表示对预测结果的影响,提取有效的空间表示并融合,从而辅助模型进行更精准的预测。

5.8.5 SCFNet 与其他外部空间建模技术的比较

我们将 SCFNet 与其他外部空间特征建模技术相比较,这里使用的是 STRN^[26] 中提出的 Meta-learner 建模方式,以及 Semantic-GAT^[13] 中嵌入 POI 空间信息的方法。为了公平比较,我们将 3 种特征通过全连接层后拼接,然后放入模型中进行学习。最终的对比结果如表 7 所列。

表 7 其他外部空间特征建模技术与 SCFNet 的效果对比

Table 7 Comparison of other external spatial feature modeling techniques and SCFNet

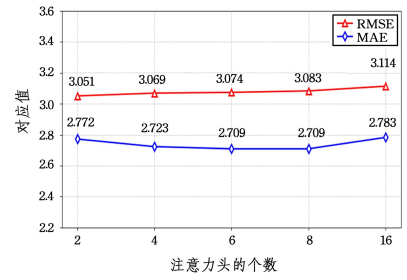
时空模型	中心区域		非中心区域	
	RMSE	SMAPE	RMSE	SMAPE
Bike NYC(60 min)				
MTGNN	4.0401	0.3492	1.7482	0.2658
MTGNN(+SCFNet)	3.8453	0.3346	1.7021	0.2626
MTGNN(+STRN)	3.8810	0.3462	1.7282	0.2765
MTGNN(+Semantic-GAT)	4.0029	0.3495	1.7363	0.2736
Metro NYC(60 min)				
MTGNN	155.2700	0.2131	78.487	0.2254
MTGNN(+SCFNet)	137.1600	0.1993	75.5090	0.2016
MTGNN(+STRN)	137.8500	0.2453	76.513	0.2619
MTGNN(+Semantic-GAT)	154.6600	0.2145	78.2580	0.2041
Taxi Chicago(60 min)				
MTGNN	111.8100	0.1884	15.1370	0.0496
MTGNN(+SCFNet)	95.7110	0.1649	13.5690	0.0476
MTGNN(+STRN)	92.0180	0.1926	14.0420	0.0572
MTGNN(+Semantic-GAT)	108.8800	0.1909	14.984	0.0510

从多个数据集的实验结果可以看出,SCFNet 相对于其他外部特征建模方式具有优越性。虽然在 Taxi Chicago 数据集上,使用 STRN 方法时中心区域的 RMSE 略低于 SCFNet,但 SMAPE 却远高于 SCFNet 方法,说明 SCFNet 对最终的预测结果提升更全面,对不同流量水平的站点整体提升更好。

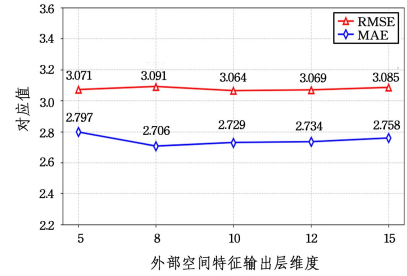
5.8.6 SCFNet 参数敏感性实验

超参数对于模型的拟合与优化具有重要作用,但如果模型在一个数据集上的性能过于依赖超参数的设置,则说明该模型的泛化能力较弱,不具有鲁棒性。因此本文选择 Bike NYC 作为验证数据集,使用 MTGNN(+SCFNet)作为时空模型,在该数据集上进行参数敏感性实验。本文分别针对 3 个重要的超参数,即注意力头数、注意力层数和输出层维度进行实验。具体而言,在固定其他超参数的前提下,仅使一个超参数在一定范围内变化,然后观察不同指标的变化情况,如图 6 所示。

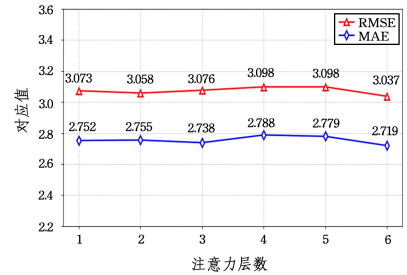
从图 6 中可以看出,无论是注意力头数、注意力层数,还是外部空间特征输出层维度,当该超参数发生一定程度的变化时,MAE 和 RMSE 的波动程度都在 2% 以内,说明 SCFNet 模型的拟合能力不依赖于超参数的设置而变化,对超参数不具有敏感性。



(a)



(b)



(c)

图 6 MTGNN(+SCFNet)模型在不同超参数上的敏感性实验
Fig. 6 Sensitivity experiments of MTGNN(+SCFNet) on different hyperparameters

结束语 本文设计了一种外部空间特征融合框架——SCFNet。针对不同类型的时空预测任务,该模型以外部空间特征作为输入,高效地生成动态的外部空间语义,并且在动态的外部空间语义和时空表示之间进行交互,获得全面的外部空间语义,从而辅助预测在当前时刻不同区域之间的流量情况。SCFNet 采用模块化设计,使其可以与多种时空模型结合。本文使用 3 个数据集,并在多个时空预测模型中添加 SCFNet,结果均优于基准方法,证明了 SCFNet 可以成功地聚合外部空间特征和时间特征,提升时空流量预测的效果。同时,SCFNet 支持多种类型的外部空间特征输入,例如区域人口数量、道路级别划分、POI 等特征。实验证明当输入多种空间特征时,SCFNet 预测效果要优于单个空间特征输入时的效果,证明了 SCFNet 可以在多种特征中学习并融合到有效的空间表示,具有良好的泛化性和实用性。

参考文献

- [1] CHEN Y, CHEN C, WU Q, et al. Spatial-temporal traffic congestion identification and correlation extraction using floating car data [J]. Journal of Intelligent Transportation Systems, 2021, 25(3): 263-280.
- [2] DRAGOMIR G. A spatial-temporal data model for choosing optimal multimodal routes in urban areas [C] // 2012 IEEE 8th In-

- ternational Conference on Intelligent Computer Communication and Processing. IEEE, 2012.
- [3] WU L, WANG M, WU D, et al. DynSTGAT: Dynamic spatial-temporal graph attention network for traffic signal control[C]// Proceedings of the 30th ACM International Conference on Information & Knowledge Management. 2021.
- [4] WU F, ZHU M, WANG Q, et al. Spatial-temporal visualization of city-wide crowd movement[J]. Journal of Visualization, 2017, 20:183-194.
- [5] FANG S, ZHANG Q, MENG G, et al. GSTNet: Global Spatial-Temporal Network for Traffic Flow Prediction [C] // IJCAI. 2019.
- [6] DENG D, SHAHABI C, DEMIRYUREK U, et al. Latent space model for road networks to predict time-varying traffic[C]// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016: 1525-1534.
- [7] YUAN J, ZHENG Y, XIE X. Discovering regions of different functions in a city using human mobility and POIs[C]// Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2012:186-194.
- [8] ZHENG C, FAN X, WEN C, et al. DeepSTD: Mining spatio-temporal disturbances of multiple context factors for citywide traffic flow prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(9):3744-3755.
- [9] XIA T, LIN J, LI Y, et al. 3dgcnn: 3-dimensional dynamic graph convolutional network for citywide crowd flow prediction[J]. ACM Transactions on Knowledge Discovery from Data, 2021, 15(6):1-21.
- [10] LIN Z Q, JI FENG J, LU Z Y, et al. Deepstn+: Context-aware spatial temporal neural network for crowd flow prediction in metropolis[C]// AAAI. 2019:1020-1027.
- [11] YUAN X M, HAN J C, WANG X, et al. A novel learning approach for citywide crowd flow prediction [C] // IEEE. 2019: 341-346.
- [12] XING J, KONG X, XING W, et al. STGs: Construct spatial and temporal graphs for citywide crowd flow prediction[J]. Applied Intelligence, 2022, 52(11):12272-12281.
- [13] LI F, FENG J, YAN H, et al. Crowd flow prediction for irregular regions with semantic graph attention network [J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2022, 13(5):1-14.
- [14] ZENG H, PENG Z, HUANG X H, et al. Deep spatio-temporal neural network based on interactive attention for traffic flow prediction[J]. Applied Intelligence, 2022, 15(9):1-12.
- [15] WANG K, LIU L B, LIU Y, et al. Urban regional function guided traffic flow prediction [J]. Information Sciences, 2023, 634:308-320.
- [16] WU Z H, PAN S R, LONG G D, et al. Graph wavenet for deep spatial-temporal graph modeling[J]. arXiv:1906.00121, 2019.
- [17] MA X, DAI Z, HE Z, et al. Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction[J]. Sensors, 2017, 17(4):818.
- [18] JIE L, YONG F. Prediction of time series data based on multi-time scale rnn [J]. Computer Applications and Software, 2018, 35(7):33-37.
- [19] LV Z, XU J, ZHENG K, et al. Lc-rnn: A deep learning model for traffic speed prediction[C]//IJCAI. 2018.
- [20] DO L N N, VU H L, VO B Q, et al. An effective spatial-temporal attention based neural network for traffic flow prediction [J]. Transportation Research Part C: Emerging Technologies, 2019, 108:12-28.
- [21] ZHU J, WANG Q, TAO C, et al. AST-GCN: Attribute-augmented spatiotemporal graph convolutional network for traffic forecasting[J]. IEEE Access, 2021, 9:35973-35983.
- [22] ZHENG C, FAN X, WANG C, et al. Gman: A graph multi-attention network for traffic prediction[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2020:1234-1241.
- [23] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems December. 2017: 6000-6010.
- [24] LIU X, YU H F, DHILON I, et al. Learning to encode position for transformer with continuous dynamical model[C]// International Conference on Machine Learning. PMLR, 2020: 6327-6335.
- [25] WANG K, WANG P, HUANG Z, et al. A two-step model for predicting travel demand in expanding subways [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(10):19534-19543.
- [26] LIANG Y, OUYANG K, SUN J, et al. Fine-grained urban flow prediction[C]// Proceedings of the Web Conference 2021. 2021: 1833-1845.
- [27] 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. 2019:1720-1730.
- [28] WANG L, CHAI D, LIU X, et al. Exploring the generalizability of spatio-temporal traffic prediction: Meta-modeling and an analytic framework[J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 35(4):3870-3884.
- [29] WU Z, PAN S, LONG G, et al. Connecting the dots: Multivariate time series forecasting with graph neural networks [C] // Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2020:753-763.



LIU Tengfei, born in 2002, is a member of CCF (No. U9218G). His main research interests include urban computing and urban traffic.



WANG Leye, born in 1987, associate professor, Ph.D supervisor, is a member of CCF (No. C9319S). His main research interests include ubiquitous computing, mobile crowdsensing and urban computing.