

## 基于双路先验自适应图神经常微分方程的交通流预测

袁蓉, 彭莉兰, 李天瑞, 李崇寿

### 引用本文

袁蓉, 彭莉兰, 李天瑞, 李崇寿. 基于双路先验自适应图神经常微分方程的交通流预测[J]. 计算机科学, 2024, 51(4): 151-157.

YUAN Rong, PENG Lilan, LI Tianrui, LI Chongshou. Traffic Flow Prediction Model Based on Dual Prior-adaptive Graph Neural ODE Network [J]. Computer Science, 2024, 51(4): 151-157.

---

### 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3D骨架相似性的自适应移位图卷积神经网络人体行为识别算法](#)

Human Action Recognition Algorithm Based on Adaptive Shifted Graph Convolutional Neural Network with 3D Skeleton Similarity

计算机科学, 2024, 51(4): 236-242. <https://doi.org/10.11896/jsjcx.221200120>

#### [基于图卷积神经网络的节点分类方法研究综述](#)

Review of Node Classification Methods Based on Graph Convolutional Neural Networks

计算机科学, 2024, 51(4): 95-105. <https://doi.org/10.11896/jsjcx.230600071>

#### [基于标签信息融合与多任务学习的中文命名实体识别](#)

Chinese Named Entity Recognition Based on Label Information Fusion and Multi-task Learning

计算机科学, 2024, 51(3): 198-204. <https://doi.org/10.11896/jsjcx.230200114>

#### [基于双分支串行混合注意力的输电线路缺陷检测深度神经网络模型](#)

Deep Neural Network Model for Transmission Line Defect Detection Based on Dual-branch Sequential Mixed Attention

计算机科学, 2024, 51(3): 135-140. <https://doi.org/10.11896/jsjcx.230600109>

#### [命题逻辑中一类正则标准矛盾体的构造与复合](#)

Construction and Compounding of a Class of Regular Standard Contradictions in Propositional Logic

计算机科学, 2024, 51(1): 295-300. <https://doi.org/10.11896/jsjcx.230600009>

# 基于双路先验自适应图神经常微分方程的交通流预测

袁蓉 彭莉兰 李天瑞 李崇寿

西南交通大学计算机与人工智能学院 成都 611756

可持续城市交通智能化教育部工程研究中心 成都 611756

(yrong@my.swjtu.edu.cn)

**摘要** 准确的交通流量预测是智能交通系统不可或缺的组成部分。近年来,图神经网络在交通流预测任务中取得了较好的预测结果。然而,图神经网络的信息传递是不连续的潜在状态传播,且随着网络层数的增加存在过平滑的问题,这限制了模型捕获远距离节点的空间依赖关系的能力。同时,在表示道路网络的空间关系时,现有方法大多仅使用先验知识构建的预定义图或仅使用路网状况构建的自适应图,忽略了两类图结合的方式。针对上述问题,提出了一种基于双路先验自适应图神经常微分方程的交通流预测模型。利用时间卷积网络捕获序列的时间相关性,使用先验自适应图融合模块表示道路网络的空间关系,并通过基于张量乘法的神经常微分方程以连续的方式传播复杂的时空特征。最后,在美国加利福尼亚州4个公开的高速公路流量数据集上进行对比实验,结果表明所提模型的预测效果优于现有的10种对比方法。

**关键词:** 交通预测;先验自适应图;图卷积神经网络;神经常微分方程;张量乘法

**中图分类号** TP181

## Traffic Flow Prediction Model Based on Dual Prior-adaptive Graph Neural ODE Network

YUAN Rong, PENG Lilan, LI Tianrui and LI Chongshou

School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China

Engineering Research Center of Sustainable Urban Intelligent Transportation, Ministry of Education, Chengdu 611756, China

**Abstract** Accurate traffic flow prediction is an indispensable part of intelligent transportation system. In recent years, graph neural networks have generated effective results in traffic flow prediction tasks. However, the information transfer of graph neural network is discontinuous latent state propagation, and there is an over-smoothing problem as the number of network layers increases, which limits the ability of the model to capture the spatial dependencies of distant nodes. At the same time, when representing the spatial relationship of the road network, most of the existing methods only use the predefined graph constructed by prior knowledge or the adaptive graph constructed only by the road network conditions, ignoring the combination of those two graphs. Aiming at solving the above problems, this paper proposes a traffic flow prediction model based on a dual prior adaptive graph neural ordinary differential equation. Temporal convolutional network are utilized to capture the temporal correlation of sequences, a priori adaptive graph fusion module is used to represent the road network, and complex spatio-temporal features are propagated in a continuous manner through tensor multiplication-based neural ODEs. Finally, experiments are carried out on four public data sets of highway traffic in California, USA. Experimental results show that the prediction performance of the model is better than that of the existing ten methods.

**Keywords** Traffic forecasting, Prior adaptive graph, Graph convolutional network, Neural ordinary differential equations, Tensor multiplication

到稿日期:2023-01-16 返修日期:2023-05-29

基金项目:国家自然科学基金(62202395, 62176221);四川省自然科学基金(2022NSFSC0930);中央高校基本科研业务费专项资金(2682022CX067)

This work was supported by the National Natural Science Foundation of China(62202395, 62176221), Natural Science Foundation of Sichuan Province, China (2022NSFSC0930) and Fundamental Research Funds for the Central Universities of Ministry of Education of China (2682022CX067).

通信作者:李崇寿(lics@swjtu.edu.cn)

## 1 引言

随着人民生活质量的不断提升,我国汽车保有量迅速增长,但交通道路建设缓慢,加之城市受限于空间难以修建更多的道路,交通拥堵问题愈发严峻。车流量是反映道路运行状态的主要参数之一,准确的交通流预测对于交通管理<sup>[1]</sup>、个人出行规划<sup>[2]</sup>和辅助规划交通基础设施的建设<sup>[3]</sup>具有重要价值。交通流预测是典型的时空预测问题,其既包含传感器采集的随时间变化的流量特征,也包含交通路网构成的空间结构信息。交通流预测任务是通过给定的历史交通条件和基础道路网络去预测未来道路的交通情况。

近年来,由于图卷积神经网络(GCN)的兴起和不断发展,加之交通网络的非欧几里得特性是天然的图结构,因此,时空图建模受到了广泛关注。已有大量的基于图卷积神经网络的方法被应用于交通流预测任务中,例如 STGCN<sup>[4]</sup>, DCRNN<sup>[5]</sup>, GraphWaveNet<sup>[6]</sup>等均使用图卷积神经网络捕获交通网络的空间特征,并且取得了较好的结果。但这些方法忽略了以下两点问题:

1)先验知识构建的预定义图和自适应图结合的空间关系表示方法。STGCN<sup>[4]</sup>和 DCRNN<sup>[5]</sup>均仅使用路网中节点间的距离构建邻接矩阵,并传入图卷积神经网络来建模空间特征。由于交通路网中存在位置相近的节点之间的流量模式可能不同,而远距离的节点之间的流量模式可能相似的情况,预定义图无法很好地表示节点间的空间依赖关系。GraphWaveNet<sup>[6]</sup>提出了自动推断节点相关性的自适应邻接矩阵,其忽视了交通路网中自然的物理连接关系。因此,考虑基于先验知识构建的预定义图和自适应图的融合非常必要。

2)使用 GCN 来捕获空间依赖时,其信息传递是不连续的潜在状态传播,且随着网络层数的增加会造成过平滑问题。因此,使 GCN 以连续的潜在状态传递信息同时解决过平滑的问题也是有必要的。

为解决上述问题,本文提出了基于双路先验自适应图神经微分方程的交通流预测模型。模型框架的骨干部分是堆叠的 2 个双路先验自适应图神经微分方程层(Dual Path Prior Adaptive Graph Neural Ordinary Differential Equation Layer, D-PAGNODE Layer)。双路先验自适应图神经微分方程层由先验自适应图(Prior Adaptive Graph, PAG)、基于张量乘法的连续图传播机制(Continuous Graph Propagation, CGP)以及时间卷积网络<sup>[7]</sup>(Temporal Convolutional Network, TCN)组成,其目的是使模型能同时捕获静态和动态的空间关系与远距离的时间相关性,并且以连续的潜在状态传递复杂的时空交互关系,进而提升模型的预测性能。本文的主要贡献概括如下:

1)提出了先验自适应图融合模块,该模块首先使用自适应图学习模块得到自适应邻接矩阵 $A_{adp}$ ,再分别与基于先验知识构建空间邻接矩阵 $A_{sp}$ 和语义邻接矩阵 $A_{sc}$ 的值融合得到先验自适应矩阵 $A_{pag}$ 以综合考虑节点的静态和动态空间关系,为模型捕获空间依赖提供更准确和丰富的空间关系。

2)设计了双路先验自适应图神经微分方程层,一路是

基于先验知识构建的空间邻接矩阵与自适应邻接矩阵值相加融合,另一路是基于先验知识构建的语义邻接矩阵与自适应邻接矩阵值相加融合,进而使模型能同时捕获不同的静态和动态的空间关系。每路均使用时间卷积网络和连续图传播层建模复杂的时空依赖。

3)在 4 个真实的高速公路流量数据集上进行了对比实验,验证了本文提出的模型在交通预测方面的有效性。

## 2 相关工作

### 2.1 交通流预测

交通流量预测对于交通管理、道路建设、公共安全等具有重要作用。早期对于交通预测的研究,主要使用的是自回归积分滑动平均法(Autoregressive Integrated Moving Average model, ARIMA)<sup>[8]</sup>、卡尔曼滤波<sup>[9]</sup>等统计学方法,而后  $K$  近邻<sup>[10]</sup>、支持向量机<sup>[11]</sup>等经典的机器学习算法也被用于交通流量预测。然而,基于统计学习的方法对复杂非线性的交通数据处理能力不足,基于机器学习的方法虽然能对复杂的数据进行建模,但面临着需要大量的特征工程的问题,且这些方法大多仅捕获了序列在时间上的相关性,忽略了节点空间信息的影响,而这对准确的交通流预测非常重要。

随着深度学习的发展,许多学者借助深度学习强大的非线性数据挖掘能力,提出了基于深度学习的交通流预测模型,并且为了既能捕获时间依赖,又能捕获空间依赖,学者们开始整合多个网络模型来进行交通预测<sup>[12-18]</sup>。具体而言,他们使用循环神经网络(RNN)及其变种来捕获时间依赖关系;使用卷积神经网络(CNN)来捕获空间依赖关系。基于 CNN 的方法提取空间特征时,需要将交通数据转换为图像网格数据,破坏了路网的自然连接。

随着图神经网络的兴起和不断发展,加之交通网络的非欧几里得特性,越来越多的学者将基于图神经网络的方法<sup>[19-20]</sup>用于交通流预测,并取得了较好的预测结果。

### 2.2 图卷积神经网络

图卷积神经网络能够实现强大的非结构化数据(如图数据)学习的能力。Yu 等<sup>[4]</sup>将图卷积网络应用到交通预测问题上,在时间和空间视图上构建卷积结构,从而能够以更少的参数实现更快的训练速度。Li 等<sup>[5]</sup>则将交通流视为扩散过程,并在有向图上使用双向随机游走捕获空间依赖关系。Guo 等<sup>[21]</sup>使用注意力机制捕获时间和空间的动态依赖;使用 3 个独立重复的组件考虑临近性、日周期性和周周期性的影响。上述方法主要使用独立的组件分别捕获时间和空间的依赖关系。

Song 等<sup>[22]</sup>为了将所有节点与其前一时刻和后一时刻相连,构建了局部时空图,并为时空网络序列添加位置嵌入,使得模型可以同时考虑时空信息;Li 等<sup>[23]</sup>通过对不同时间段的不同空间和时间图进行融合运算,可以有效地学习隐藏的时空相关性,以此来达到同时提取时空依赖的效果,是一种动态建模方式。

然而,图卷积神经网络已被证明随着网络层数的增加会出现过平滑的问题,加之离散的网络架构在信息传递时是不

连续的潜在状态传播,这限制了模型建模远距离节点的空间相关性的能力,导致模型产生更大的数值误差。为了打破这一局限,进一步提升建模性能,神经常微分方程受到了关注。

### 2.3 神经常微分方程

在神经网络中,隐藏状态是由一系列离散的变换表示的:

$$h_{(l+1)} = f(h_{(l)}, \theta_{(l)}, l) \quad (1)$$

其中,  $f$  是参数为  $\theta_{(l)}$  的第  $l$  层神经网络。Chen 等<sup>[24]</sup> 提出能否通过插入更多的层并采取更小的积分操作,进而直接参数化和近似潜在状态的连续演化,从而避免不连续的潜在状态轨迹的影响。基于此,神经常微分方程的基本思想如下:

$$\begin{cases} \frac{dh(l)}{dt} = f(h_{(l)}, \theta_{(l)}, l) \\ h(l) = \text{ODESolve}(h(0), f, l_0, l_1, \theta) \end{cases} \quad (2)$$

近年来,神经常微分方程已被应用于交通流预测任务。STGODE<sup>[25]</sup> 作为目前在交通流预测任务上首次应用神经常微分方程的方法,其使用神经常微分方程结合张量乘法以连续的方式传递复杂的时空关系,并证明了该神经常微分方程有解析解。但该模型仅考虑了预定义图结构上的连续图传播的影响,没有考虑自适应图的影响以及自适应图和预定义图的融合影响。

本文提出的方法受神经常微分方程和 STGODE<sup>[25]</sup> 的启发,但不同的是,首先我们构建了一个自适应图学习模块,来动态地学习节点间的相关性;其次我们提出了一种先验自适应图融合模块,使用自适应图分别与基于先验知识构建的预定义图的值相加融合;并设计了双路结构,每路传入基于不同先验知识学到的预定义图与自适应图的融合,以综合全面地考虑节点间的关系,进而为模型建模空间依赖提供更准确、丰富的关系,并通过实验证明了本文提出的模型对性能提升的有效性。

## 3 问题定义

交通流量预测是指利用交通路网中的传感器、车载 GPS 等采集到的丰富数据结合路网中传感器的位置信息构成的历史时空数据去预测未来交通网络中各个传感器的流量信息。我们用  $\mathbf{V}$  表示路网中传感器  $N$  的集合,即  $|\mathbf{V}| = N$ 。用  $\mathbf{E}$  表示传感器之间连接关系,用  $\mathbf{A}$  表示由  $N$  和  $\mathbf{E}$  构成的邻接矩阵,则交通路网可以由图  $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{A})$  表示。

假设在时刻  $t$ ,  $N$  个传感器采集到的信息为  $\mathbf{X}^t \in R^{N \times D}$ , 其中  $D$  为传感器采集到的数据,如流量、车道占用率、车速等,其表示特征维度。因此交通流预测任务可以表示为:

$$\mathbf{X}^{t+1}, \mathbf{X}^{t+2}, \dots, \mathbf{X}^{t+T} = f(\mathbf{X}^{t-T+1}, \mathbf{X}^{t-T+2}, \dots, \mathbf{X}^t, \mathbf{G}) \quad (3)$$

即学习一个映射函数  $f$ , 从历史观测量  $T$  去预测未来  $T'$  时刻的交通情况。

## 4 基于双路先验自适应图神经常微分方程的交通流预测

本文提出的模型的总体框架如图 1(a) 所示,主要由两层双路先验自适应图神经常微分方程层堆叠而成。输入层接收到历史交通流数据后传入堆叠的双路先验自适应图神经常微分

方程层中,最后再经过一层池化输出得到最终预测值。

图 1(b) 所示为双路先验自适应图神经常微分方程层(D-PAGNODE);一路是基于先验知识构建的空间邻接矩阵和自适应邻接矩阵融合得到的先验自适应矩阵,另一路是基于先验知识构建的语义邻接矩阵和自适应邻接矩阵融合得到的先验自适应矩阵,且每路由两个 TCN 和一个 CGP 组成。时间卷积网络用于建模长期的时间相关性,连续图传播使用基于张量乘法的神经常微分方程以连续的潜在状态传播方式传播复杂的时空依赖关系信息,先验自适应图的主要目的是融合先验知识构建的预定义图与自适应图相加得到先验自适应图,以综合全面地考虑节点的空间关系,使模型能同时捕获静态和动态的空间关系,并将其和时间卷积网络的输出一起作为连续图传播机制的输入。

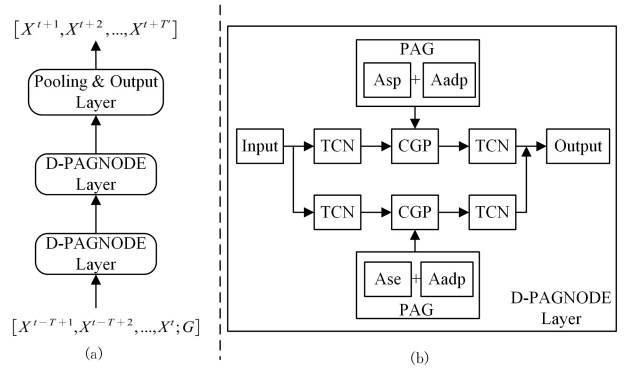


图 1 模型框架图

Fig. 1 Model architecture diagram

### 4.1 先验自适应邻接矩阵的构造

在交通路网中,传感器的位置分布提供了直观的空间位置信息,距离近的节点间的相关性可能更大,距离远的节点间的相关性可能较小;其次,不同节点之间的流量趋势的相似性也说明这些节点可能具有相似的位置功能信息,例如,不同区域的办公区有着相似的交通模式。基于这两种先验知识,我们构建了两种先验图,分别是基于路网结构的空间邻接矩阵和基于动态时间规整算法(Dynamic Time Warping, DTW)<sup>[26]</sup> 的语义邻接矩阵。

DTW 算法相较于逐点时间相似性的度量方法,优势在于其考虑了时间轴上的弯曲和变形,对于形状更敏感,能更好地捕获序列间的相似性。给定两条时间序列  $\mathbf{X} = (x_1, x_2, \dots, x_m)$  和  $\mathbf{Y} = (y_1, y_2, \dots, y_n)$ , DTW 算法定义如下:

$$D(i, j) = \text{dist}(x_i, y_j) + \min(D(i-1, j), D(i, j-1), D(i-1, j-1)) \quad (4)$$

其中,  $D(i, j)$  表示子序列  $\mathbf{X} = (x_1, x_2, \dots, x_i)$  和  $\mathbf{Y} = (y_1, y_2, \dots, y_j)$  的最短距离,  $\text{dist}(x_i, y_j)$  是  $x_i$  和  $y_j$  的绝对距离。

本文的空间邻接矩阵定义如下:

$$A_{ij}^p = \begin{cases} \exp\left(\frac{-d_{ij}^2}{\sigma^2}\right), & \text{if } \exp\left(\frac{-d_{ij}^2}{\sigma^2}\right) \geq \epsilon \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

其中,  $d_{ij}$  表示节点  $i$  到节点  $j$  的距离,  $\sigma^2$  和  $\epsilon$  是控制矩阵稀疏程度的超参数。

同时,为了捕获节点流量模式的相似性,受 DTW<sup>[26]</sup> 算法

的启发,我们将语义邻接矩阵定义为:

$$\mathbf{A}_{ij}^s = \begin{cases} 1, & \text{if } D(\mathbf{X}^i, \mathbf{Y}^j) < \epsilon \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

其中,  $\epsilon$  是控制语义邻接矩阵稀疏程度的超参数,  $\mathbf{X}^i$  表示第  $i$  个节点的时间序列。

此外,由于上述两种节点之间的连接关系是固定且静态的,不能反映节点之间的动态空间关系。受 Wu 等<sup>[6]</sup>的启发,我们通过自适应图学习模块学习节点之间动态的空间相关性。

首先随机初始化一个可学习的节点嵌入矩阵  $\mathbf{E}_A \in R^{N \times e}$ , 其中每一行是节点的嵌入表示,  $e$  是节点的嵌入维度,在模型训练过程中  $\mathbf{E}_A$  将被动态更新。将  $\mathbf{E}_A$  与其转置矩阵  $\mathbf{E}_A^T$  相乘得到的结果即为节点之间的空间相关性。通过式(7)得到自适应邻接矩阵:

$$\mathbf{A}_{\text{adp}} = \text{softmax}(\text{ReLU}(\mathbf{E}_A \cdot \mathbf{E}_A^T)) \quad (7)$$

为了更好地利用构建的图,使模型能同时捕获不同的静态和动态的空间关系,本文设计了一个双路先验自适应图融合机制,一路是空间邻接矩阵  $\mathbf{A}_{\text{sp}}$  与自适应邻接矩阵  $\mathbf{A}_{\text{adp}}$  的融合,另一路是语义邻接矩阵  $\mathbf{A}_{\text{se}}$  与自适应邻接矩阵  $\mathbf{A}_{\text{adp}}$  的融合。

## 4.2 连续图传播

图卷积神经网络通过图卷积操作聚合其自身和节点嵌入来更新节点表示。经典的图卷积运算公式如下:

$$\mathbf{H}^{(l+1)} = \delta(\mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{\frac{1}{2}}) \mathbf{H}^{(l)} \mathbf{W}^{(l)} \quad (8)$$

其中,  $\mathbf{H}^{(l)} \in R^{N \times C}$  表示图卷积第  $l$  层的输入,  $\mathbf{A} \in R^{N \times N}$  是图构建的邻接矩阵,  $\mathbf{D}$  是度矩阵,  $\mathbf{W} \in R^{C \times C}$  是可学习的参数矩阵。

尽管 GCN 为图数据带来了一种全新的建模方式,然而 Li 等<sup>[27]</sup>和 Xu 等<sup>[28]</sup>都明确指出 GCN 模型无法像视觉任务中的 CNN 模型一样进行深度堆叠,一旦使用多层 GCN,就会出现过平滑的问题,进而导致相关任务的性能急剧下降,限制了 GCN 的建模能力。加之此类方法大多仅对空间进行聚合,忽略了复杂的时空交互。

为了将仅考虑空间建模的离散图传播过程转变为建模复杂时空交互的连续图传播过程,同时解决 GCN 随着层数的增加遇到的过平滑问题。本文使用基于张量乘法的神经微分方程,公式如下:

$$\frac{d\mathbf{H}(l)}{dl} = \mathbf{H}(l) \times_1 (\mathbf{A}_{\text{pag}} - \mathbf{I}) + \mathbf{H}(l) \times_2 (\mathbf{U} - \mathbf{I}) + \mathbf{H}(l) \times_3 (\mathbf{W} - \mathbf{I}) + \mathbf{H}_0 \quad (9)$$

其中,  $\mathbf{H}_0$  表示基于张量乘法的神经微分方程的初始输入;  $\mathbf{H}(l)$  表示第  $l$  层时空张量的隐藏表示,  $\times_i$  表示按第  $i$  个模态展开进行张量乘法;  $\mathbf{U}$  是时间变换矩阵,为时间卷积网络的输出;  $\mathbf{A}_{\text{pag}}$  是先验自适应矩阵;  $\mathbf{I}$  为单位矩阵;  $\mathbf{W}$  表示特征变换矩阵。使用神经微分方程求解器回传梯度:

$$\mathbf{H}(l) = \text{ODESolve}\left(\frac{d\mathbf{H}(l)}{dl}, \mathbf{H}_0, l\right) \quad (10)$$

## 4.3 时间卷积模块

基于 RNN 及其变体的模型和基于注意力机制的模型在捕获序列的时间依赖中得到了广泛的应用,但基于递归的网络模型存在耗时迭代、不稳定梯度和对动态变化延迟响应

等问题,基于注意力机制的网络模型计算复杂度较高。

因此,本文使用 TCN 来捕获时间依赖,其定义如下:

$$\mathbf{H}_{\text{tcn}}^l = \begin{cases} \mathbf{X}, & l=0 \\ \delta(\mathbf{W}_{d'}^{l'} \mathbf{H}_{\text{tcn}}^{l-1}), & l=1, 2, \dots, L \end{cases} \quad (11)$$

其中,  $\mathbf{X} \in R^{N \times T \times F}$  表示 TCN 的输入;  $\mathbf{H}_{\text{tcn}}^l \in R^{N \times T \times F}$  表示第  $l$  层 TCN 的输出;  $\mathbf{W}^{l'}$  表示第  $l$  层的卷积核大小;  $d'$  为扩展因子,为了获得更大的感受野,设置  $d' = 2^{l-1}$ 。

## 5 实验与结果分析

### 5.1 数据集

本文使用美国加利福尼亚州公开的真实高速公路数据集 PEMS03, PEMS04, PEMS07 和 PEMS08 来验证本文模型的性能。PEMS 数据集是由 Caltrans Performance Measurement System 实时采集的高速公路交通数据,该系统拥有超过 39000 个传感器站,这些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。

该数据集是基于每次 30 s 的频率采样得到的原始数据汇总成以 5 min 为间隔的时间样本,本文按照时间顺序将这些数据集的 60% 作为训练数据集,20% 作为验证集,20% 作为测试集。输入序列长度为 12,目标序列长度为 12,即使用过去 1 h 的道路交通历史数据,预测未来 1 h 的道路交通情况。表 1 列出了 PEMS 不同区域数据集的详细信息。

表 1 PEMS 不同区域数据集的统计信息

Table 1 Statistics for different regional dataset in PEMS

数据集	节点数	边数	时间步长
PEMS03	358	547	26 208
PEMS04	307	340	16 992
PEMS07	883	866	28 224
PEMS08	170	295	17 856

### 5.2 实验参数设置

本文提出的框架中,每个 D-PAGNODE 层包含 3 个 D-PAGNODE 块,D-PAGNODE 块的时间卷积网络的隐藏层维度设置为 64,32,64,控制空间邻接矩阵稀疏程度的超参数  $\sigma$  和  $\epsilon$  的阈值分别为 10 和 0.5;控制语义邻接矩阵稀疏程度的超参数  $\epsilon$  的阈值为 0.6;动态关系图学习模块的节点嵌入维度  $e$  设置为 10。模型通过 Adam 优化器进行训练,学习率为 0.002, batch\_size 为 16。每个时间卷积层的输出使用 0.2 的 Dropout。实验均在 Linux 服务器上进行,其 CPU 为 Intel (R) Xeon(R) Silver 4210R CPU @ 2.40 GHz, GPU 为 NVIDIA GeForce GTX3090。

### 5.3 对比方法

本文将所提模型与以下 10 种已有的应用于交通预测任务上的方法进行比较。这些方法可大致分为 5 类:1) 1 个传统的时间序列模型 ARIMA<sup>[8]</sup>; 2) 2 个简单的深度学习模型 TCN<sup>[7]</sup> 和 FC-LSTM<sup>[14]</sup>; 3) 4 个使用独立组件分别捕获时间和空间依赖的模型 STGCN<sup>[4]</sup>, DCRNN<sup>[5]</sup>, GraphWaveNet<sup>[6]</sup> 和 ASTGCN(r)<sup>[21]</sup>; 4) 2 个考虑时间和空间依赖复杂交互的模型 STSGCN<sup>[22]</sup> 和 STFGCN<sup>[23]</sup>; 5) 1 个以连续形式进行时空

信息传递的模型 STGODE<sup>[25]</sup>。这些模型的详细信息描述如下:

ARIMA<sup>[8]</sup>:自回归积分移动平均法,该方法是一个经典的基于统计模型的时间序列分析模型。

TCN<sup>[7]</sup>:时间卷积网络,其使用卷积网络以并行的方式对时间序列进行建模。

FC-LSTM<sup>[14]</sup>:LSTM 隐藏单元全连接的循环神经网络,以全连接的方式解决了 LSTM 仅考虑了时序没考虑空间相关性的问题。

STGCN<sup>[4]</sup>:一种时空图卷积神经网络,其将图卷积神经网络应用于交通预测问题上。

DCRNN<sup>[5]</sup>:使用双向随机游走捕获空间相关性,使用 GRU 捕获时间相关性,并使用了编码器-解码器架构。

GraphWaveNet<sup>[6]</sup>:一种集成了扩散图卷积和一维扩展卷积的时空图卷积网络。

ASTGCN(r)<sup>[21]</sup>:一种基于注意力机制的时空图卷积网络。其使用注意力机制捕获时间和空间的动态依赖,使用 3 个独立重复的组件考虑临近性、日周期性和周周期性的影响。

STSGCN<sup>[22]</sup>:一种时空同步图卷积网络模型。通过将所有节点与其前一时刻和后一时刻相连,构建了局部时空图,并为时空网络序列添加位置嵌入,使得模型可以同时考虑时空信息。

STFGCN<sup>[23]</sup>:一种时空融合图神经网络模型,其使用数据驱动的方式产生时间图,补足空间图可能没有体现

出来但存在的关联。

STGODE<sup>[25]</sup>:一种基于张量乘法的时空图神经微分方程网络。其使用基于张量形式的连续图表示突破了网络深度的限制,提高了提取长距离的时空依赖的能力。

本文采用 huber 损失函数作为损失函数,采用平均绝对误差(MAE)、均方误差(RMSE)和平均绝对百分比误差(MAPE)作为评价指标,指标公式定义如下:

$$L_{\delta}(Y, \hat{Y}) = \begin{cases} \frac{1}{2}(Y - \hat{Y})^2, & |Y - \hat{Y}| \leq \delta \\ \delta|Y - \hat{Y}| - \frac{1}{2}\delta^2, & \text{otherwise} \end{cases} \quad (12)$$

$$MAE(y_i, \hat{y}_i) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (13)$$

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

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

其中, $\delta$ 为控制异常值敏感性的超参数, $y_i$ 表示第  $i$  个样本的真实值, $\hat{y}_i$ 表示第  $i$  个样本的预测值, $n$ 表示样本数量。

#### 5.4 实验结果分析

将本文提出的基于双路先验自适应图神经微分方程的交通流预测模型在数据集 PEMS03, PEMS04, PEMS07 和 PEMS08 上与上述 10 种对比模型进行了比较,本文引用了相关文献中公开的上述 10 种方法的结果。表 2 列出了对未来 1 h 的道路交通流量进行预测的结果。

表 2 不同方法在 PEMS03, PEMS04, PEMS07 和 PEMS08 数据集上的性能比较

Table 2 Performance comparison of different approaches on PEMS03, PEMS04, PEMS07 and PEMS08 datasets

Models	PEMS03			PEMS04			PEMS07			PEMS08		
	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
ARIMA	35.41	47.59	33.78	33.73	48.8	24.18	38.17	59.27	19.46	31.09	44.32	22.73
TCN	19.32	33.55	19.93	23.22	37.26	15.59	32.72	42.23	14.26	22.72	35.79	14.03
FC-LSTM	21.33	35.11	23.33	26.77	40.65	18.23	29.98	45.94	13.20	23.09	35.17	14.99
STGCN	17.55	30.42	17.34	21.16	34.89	13.83	25.33	39.34	11.21	17.50	27.09	11.29
DCRNN	17.99	30.31	18.34	21.22	33.44	14.17	25.22	38.61	11.82	16.82	26.36	10.92
GraphWaveNet	19.12	32.77	18.89	24.89	39.66	17.29	26.39	41.50	11.97	18.28	30.05	12.15
ASTGCN(r)	17.34	29.56	17.21	22.93	35.22	16.56	24.01	37.87	10.73	18.25	28.06	11.64
STSGCN	17.48	29.21	16.78	21.19	33.65	13.90	24.26	39.03	10.21	17.13	26.80	10.96
STFGNN	16.77	28.34	<u>16.30</u>	<u>20.48</u>	<u>32.51</u>	16.77	23.46	<u>36.60</u>	<u>9.21</u>	16.94	26.25	<u>10.60</u>
STGODE	<u>16.50</u>	<u>27.84</u>	16.69	<u>20.84</u>	32.82	<u>13.77</u>	22.99	37.54	10.14	<u>16.81</u>	<u>25.97</u>	10.62
Ours	<b>15.91</b>	<b>27.11</b>	<b>15.63</b>	<b>19.91</b>	<b>31.42</b>	<b>13.64</b>	<b>21.89</b>	<b>34.87</b>	<b>9.21</b>	<b>15.76</b>	<b>24.87</b>	<b>10.23</b>
Improvements/%	3.58	2.62	4.11	2.78	3.35	0.94	4.76	4.73	0	6.25	4.24	3.49

从表 2 的结果可以看出,本文模型在 4 个数据集的所有评价指标上都取得了最优的结果,且基于深度学习的方法在 4 个数据集上的所有评估指标比经典的统计学方法 ARIMA<sup>[8]</sup>的预测误差更小。在深度学习方法中,STGCN<sup>[4]</sup>, DCRNN<sup>[5]</sup>, GraphWaveNet<sup>[6]</sup>, ASTGCN(r)<sup>[21]</sup>, STSGCN<sup>[22]</sup>, STFGCN<sup>[23]</sup>, STGODE<sup>[25]</sup> 因为同时考虑对时间和空间维度进行建模,所以性能普遍优于 TCN<sup>[7]</sup>和 FC-LSTM<sup>[14]</sup>。且 STSGCN<sup>[22]</sup>, STFGCN<sup>[23]</sup>, STGODE<sup>[25]</sup> 由于考虑了时间维度和空间维度的复杂交互问题,性能又普遍优于分别对时间和空间维度进行建模再加权融合的模式。

此外,实验分析了使用不同的图对模型性能的影响,并且在数据集 PEMS03, PEMS04, PEMS07 和 PEMS08 上进行了

实验结果的对比。其中 only-prior 表示仅使用基于先验知识构建的空间邻接矩阵和语义邻接矩阵,仍为双路结构;only-adaptive 表示仅使用自适应图,由于只有自适应图,所以为单路结构。图 2 给出了在 PEMS03, PEMS04, PEMS07 和 PEMS08 数据集上使用不同图的 MAE, RMSE, MAPE 值的对比结果。从图中可以直观地看出本文提出的融合先验自适应图在各评价指标上均有了较明显的提升,说明将自适应图与先验图融合对模型性能的提升具有不可忽视的作用。

最后,为了进一步验证使用连续的图传播机制和常规的 GCN 对模型性能的影响,我们在数据集 PEMS03, PEMS04, PEMS07 和 PEMS08 上进行了实验对比,其结果如表 3 所列。

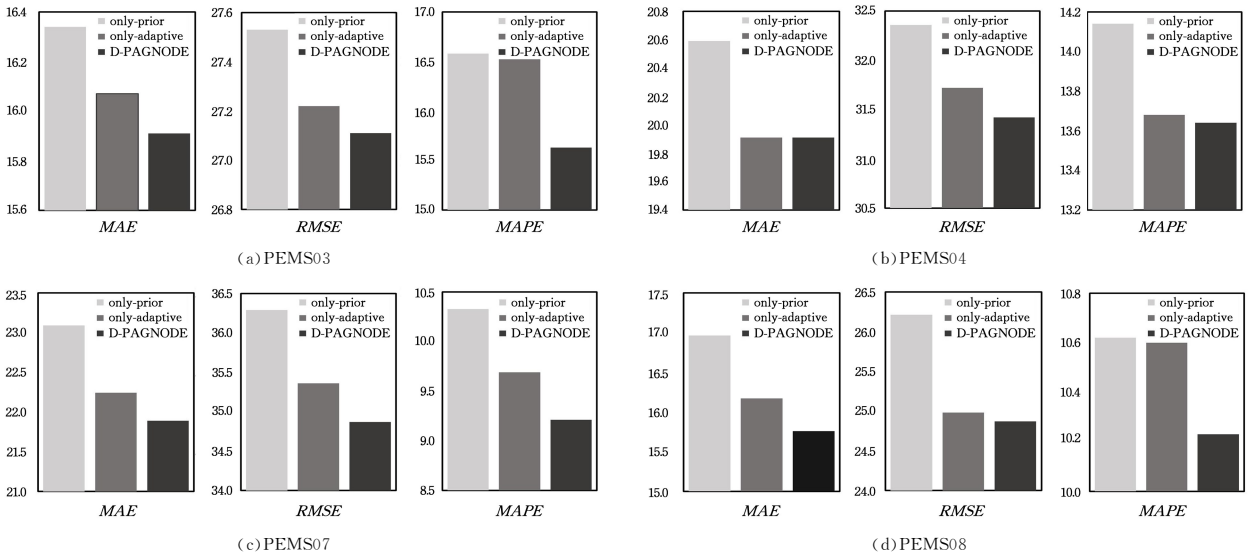


图 2 在 PEMS03, PEMS04, PEMS07 和 PEMS08 数据集上使用不同图的实验结果

Fig. 2 Experimental results using different Graphs on PEMS03, PEMS04, PEMS07 and PEMS08 datasets

表 3 常规 GCN 和 CGP 在数据集 PEMS03, PEMS04, PEMS07 和 PEMS08 上的性能比较

Table 3 Performance comparison of classic GCN and CGP on PEMS03, PEMS04, PEMS07 and PEMS08 datasets

Models	PEMS03			PEMS04			PEMS07			PEMS08		
	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
D-PAG-gcn	17.15	28.79	17.51	20.63	33.22	14.42	25.25	40.90	11.23	16.58	25.84	11.97
D-PAGNODE	<b>15.91</b>	<b>27.11</b>	<b>15.63</b>	<b>19.91</b>	<b>31.42</b>	<b>13.64</b>	<b>21.89</b>	<b>34.87</b>	<b>9.21</b>	<b>15.76</b>	<b>24.87</b>	<b>10.23</b>

从表 3 可以看出,使用连续的图传播机制对模型的性能提升有很大帮助,说明使 GCN 以连续的潜在状态传递信息是有必要的。

**结束语** 本文提出了一种基于双路先验自适应图神经常微分方程的交通流预测模型,该模型通过时间卷积网络、先验自适应图分别对时间依赖和节点间关系进行建模,利用基于张量乘法的神经常微分方程以连续的潜在状态传播形式挖掘复杂的时空依赖,并将整个模型设置成双路结构以更好地捕获交通数据中的时空信息,获得更低的预测误差。最后,在真实的高速公路流量数据集上进行了对比实验,实验结果表明本文提出的模型在交通预测方面的性能更好。未来,随着神经常微分方程的不断发展,我们将考虑在时间依赖的捕获上也引入神经常微分方程的思想,以进一步提升模型的预测能力,实现更精准的交通预测。

## 参考文献

- [1] CHEN Q, SONG X, YAMADA H, et al. Learning deep representation from big and heterogeneous data for traffic accident inference[C]// Proceedings of the 30th AAAI Conference on Artificial Intelligence. 2016.
- [2] GONG Y, LI Z, ZHANG J, et al. Potential passenger flow prediction: A novel study for urban transportation development [C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2020:4020-4027.
- [3] JAYARAJAH K, TAN A, MISRA A. Understanding the interdependency of land use and mobility for urban planning[C]// Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous

- Computing and Wearable Computers. 2018:1079-1087.
- [4] 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. 2018:3634-3640.
- [5] LI Y, YU R, SHAHABI C, et al. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting[J]. arXiv: 1707.01926, 2017.
- [6] WU Z, PAN S, LONG G, et al. Graph WaveNet for Deep Spatial-Temporal Graph Modeling [C]// The 28th International Joint Conference on Artificial Intelligence (IJCAI). 2019.
- [7] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[J]. arXiv:1803.01271, 2018.
- [8] WILLIAMS B M, DURVASULA P K, BROWN D E. Urban freeway traffic flow prediction: application of seasonal autoregressive integrated moving average and exponential smoothing models[J]. Transportation Research Record, 1998, 1644(1): 132-141.
- [9] OKUTANI I, STEPHANEDES Y J. Dynamic prediction of traffic volume through Kalman filtering theory[J]. Transportation Research Part B: Methodological, 1984, 18(1): 1-11.
- [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] HABTEMICHAEL F G, CETIN M. Short-term traffic flow rate forecasting based on identifying similar traffic patterns[J]. Transportation Research Part C: Emerging Technologies, 2016,

- 66:61-78.
- [12] ZHAO Z, CHEN W, WU X, et al. LSTM network: a deep learning approach for short-term traffic forecast[J]. IET Intelligent Transport Systems, 2017, 11(2): 68-75.
- [13] FU R, ZHANG Z, LI L. Using LSTM and GRU neural network methods for traffic flow prediction[C]// Proceedings of the 31st Youth Academic Annual Conference of Chinese Association of Automation(YAC). IEEE, 2016: 324-328.
- [14] SUTSKEVER I, VINYALS O, LE Q V. Sequence to Sequence Learning with Neural Networks[J]. arXiv: 1409. 3215, 2014.
- [15] ZHANG J, ZHENG Y, QI D. Deep spatio-temporal residual networks for citywide crowd flows prediction[C]// Proceedings of the 31th AAAI Conference on Artificial Intelligence. 2017.
- [16] JIANG W, ZHANG L. Geospatial data to images: A deep-learning framework for traffic forecasting[J]. Tsinghua Science and Technology, 2018, 24(1): 52-64.
- [17] YU H, WU Z, WANG S, et al. Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks[J]. Sensors, 2017, 17(7): 1501.
- [18] REN J H, ZHU Y, MENG X F, et al. Predicting Citywide Traffic Flow Using Dynamic Spatial-temporal Neural Networks[J]. Journal of Chinese Computer Systems, 2023, 44(3): 529-535.
- [19] ZHAO L, SONG Y, ZHANG C, et al. T-gcn: A temporal graph convolutional network for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(9): 3848-3858.
- [20] YU B, LEE Y, SOHN K. Forecasting road traffic speeds by considering area-wide spatio-temporal dependencies based on a graph convolutional neural network(GCN)[J]. Transportation Research Part C: Emerging Technologies, 2020, 114: 189-204.
- [21] GUO S, LIN Y, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2019, 33(1): 922-929.
- [22] SONG C, LIN Y, GUO S, et al. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(1): 914-921.
- [23] LI M, ZHU Z. Spatial-temporal fusion graph neural networks for traffic flow forecasting[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2021, 35(5): 4189-4196.
- [24] CHEN R T Q, RUBANOVA Y, BETTENCOURT J, et al. Neural ordinary differential equations[J]. arXiv: 1806. 07366, 2018.
- [25] FANG Z, LONG Q, SONG G, et al. Spatial-temporal graph ode networks for traffic flow forecasting[C]// Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021: 364-373.
- [26] BERNDT D J, CLIFFORD J. Using dynamic time warping to find patterns in time series[C]// KDD workshop. 1994, 10(16): 359-370.
- [27] LI Q, HAN Z, WU X M. Deeper insights into graph convolutional networks for semi-supervised learning[C]// Proceedings of the 32th AAAI Conference on Artificial Intelligence. 2018.
- [28] XU K, LIC, TIAN Y, et al. Representation learning on graphs with jumping knowledge networks[C]// International Conference on Machine Learning. PMLR, 2018: 5453-5462.



**YUAN Rong**, born in 1999, postgraduate, is a member of CCF (No. O8080G). Her main research interests include spatio-temporal data mining and traffic forecasting.



**LI Chongshou**, born in 1988, Ph.D, associate professor, is a member of CCF (No. J8308M). His main research interests include intelligent transportation, data analysis and AI.

(责任编辑:何杨)