

# 基于动态时空图卷积的多传感器路网交通流预测

孔文翔\*, 杨雪驰

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2023年4月4日; 录用日期: 2023年6月22日; 发布日期: 2023年6月29日

## 摘要

准确的交通预测为城市发展提供规划支持, 然而交通流预测精度取决于多重影响因素, 道路之间的复杂程度与时间逻辑上的变化规律, 现有方法无法合理分析路网交通模式的动态时空相关性, 为解决这些问题本文提出一种结合动态自适应(Dynamic Self-adapting)、时空注意力机制(TS Attention)、基准自适应机制(Benchmark Adaptive Mechanism)与空洞卷积(Dilated Convolution)的切比雪夫图卷积神经网络(GCN)。该模型采用时空注意力机制提取时间与空间动态相关性, 结合切比雪夫图卷积神经网络获取交通流数据空间依赖关系, 同时, 将GCN的输出作为输入, 该网络引入空洞卷积扩展感受野范围和提取时间和周期依赖关系, 增加残差模块以构建时空残差网络, 最后多模块融合预测。

## 关键词

交通流预测, 图卷积, 动态自适应, 时空注意力机制

# Multi-Sensor Road Network Traffic Flow Prediction Based on Dynamic Spatio-Temporal Graph Convolution

Wenxiang Kong\*, Xuechi Yang

School of Optical-Electrical Computer Engineering, University of Shanghai for Science & Technology, Shanghai

Received: Apr. 4<sup>th</sup>, 2023; accepted: Jun. 22<sup>nd</sup>, 2023; published: Jun. 29<sup>th</sup>, 2023

\*通讯作者。

文章引用: 孔文翔, 杨雪驰. 基于动态时空图卷积的多传感器路网交通流预测[J]. 运筹与模糊学, 2023, 13(3): 2339-2354. DOI: 10.12677/orf.2023.133234

## Abstract

Accurate traffic prediction provides planning support for urban development. However, the accuracy of traffic flow prediction depends on multiple influencing factors, the complexity between roads and the temporal logic change rule, and the existing methods cannot reasonably analyze the dynamic temporal and spatial correlation of road network traffic patterns. To solve these problems, this paper proposes a novel adapting method that combines Dynamic Self-adapting, TS Attention, Benchmark Adaptive Mechanism and Dilated convolution Convolution of Chebyshev graph convolutional neural networks (GCN). This model uses spatio-temporal attention mechanism to extract the dynamic correlation between time and space, and combines Chebyshev graph convolutional neural network to obtain the spatial dependency of traffic flow data. At the same time, the output of GCN is taken as the input, the network introduces the cavity convolution to extend the receptive field range and extract the time and period dependence, and adds the residual module to construct the spatiotemporal residual network. Finally, multiple modules are fused and predicted.

## Keywords

Traffic Flow Prediction, Graph Convolution, Dynamic Adaptive, Spatio-Temporal Attention Mechanism

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

近年来城市内的机动车数量随着城市的发展不断增长,一方面考验城市的道路资源需求能力,另一方面也影响到交通运输行业等的发展,道路交通资源与日益增长的交通需求之间不平衡的问题日渐凸显。交通流预测使用多传感器数据来预测道路的未来交通状况,作为智能交通领域中不可或缺的组成部分,交通流预测是复杂的时空预测问题,时间上,历史交通数据和不同时段事件促使预测存在动态变化;在空间上交通节点并非独立存在,节点之间产生相互作用,动态的变化在空间上同样存在。因此,交通流的变化在时间关系和空间关系上存在动态周期性特征,数据的线性、非线性以及周期性特征使得交通流精准预测具有挑战性。

交通流预测随时间而动态变化,早期研究者们对于交通流的预测仅基于历史交通流数据,历史平均模型(Historical Average, HA) [1]已经不能够适应动态变化,因而基于此学者们提出多种预测模型,主要可以分为传统统计学习方法、经典机器学习方法以及深度学习方法。传统统计学习方法例如 Box 和 Jenkins 提出自回归移动平均模型(Auto Regressive Integrated Moving Average, ARIMA) [2],该模型仅能够预测理想状态下线性变化情况,但不能够贴合实际,同时非线性特征表征能力不足,后期为了提高模型的预测精度提出关于 ARIMA 的变体季节性自回归综合移动平均模型(Seasonal Autoregressive Integrated Moving Average, SARIMA) [3],以应对季节性时序变化,但对于复杂、非线性数据的处理能力仍然有限。

跟上述方法相比较,经典机器学习方法具有处理一定复杂数据的能力,能够解决 ARIMA 所不能应对的非线性问题,经典机器学习方法例如支持向量机回归(Support Vector Regression, SVR) [4]等能够挖掘交通流数据中的非线性关系,通过核函数降低非线性数据维度,以便分析数据特征进行建模,其中核

函数直接影响到模型的性能, 而如贝叶斯网络(Bayesian Networks) [5]、K 邻近(K-Nearest Neighbor, KNN) [6]能够通过交通流数据构建特征向量, 但构建合理特征去提升预测精度方面仍具有一定难度。神经网络作为非参数机器学习网络, 在信息处理和非参数结构上更具有优势, 可以有效处理高速改变的交通流数据。Smith 等人使用 BP (Back Propagation) [7]神经网络处理动态信息, 相比较于传统统计学习模型以及非参数回归模型, 其在短时预测精度上有很大提高。

随着深度学习方法在自然语言处理领域取得了丰富的成果与经验, 应对具有时间和空间相关性的数据挖掘更具有优势, Yu 等人[8]提出长短期记忆网络(Long Short Term Memory, LSTM)的方法, 更好的捕捉交通数据的动态特征, 表征交通流的长期时间依赖性特征以及非线性变化; Koesdwiady 等人[9]将天气信息与交通信息相融合, 进一步提高了深度信念网络(DBN)的预测性能; Qi 等人[10]通过局部约束距离量度的方法检测道路交通拥堵程度, 随后 Qi 等人提出基于深度语义特征提取的鲁棒分层深度学习方法, 更好的使用拥堵的情况; Shi 等人[11]提出将卷积和循环神经网络相融合的预测模型 Conv-LSTM; Zhao 等人[12]考虑到交通网络的复杂拓扑结构和影响交通流因素的多样性, 交通模型通常是复杂多变的, 为捕捉空间和时间相关性, 提出时间图卷积网络 T-GCN, 该模型与图卷积网络和门控递归单元相结合; Hu 等人[13]提出了一种动态图生成模块, 该模块预先收集关于节点之间的地理邻近性和空间异质性信息; 谷远利等人[14]提出信息熵的灰色关联分析, 用于提取交通流中的空间依赖关系, 使用 Bi-LSTM 网络实现快速车道速度预测。

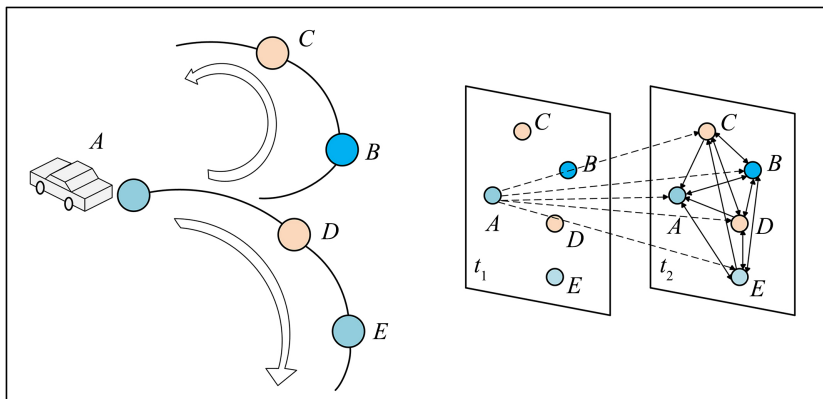
交通流的形成、扩散以及交通状态的演化过程与交通网络结构密切相关, 故对于路网的描述以及合理表达形式是交通流预测精度的重要性因素之一[15]。以往表征节点结构关系常以逻辑(0&1)关系表示, 这种权重和连通关系的设计对路网的空间结构刻画不足, 同时外部的道路事故等复杂情况制约也使得交通系统具有时空强耦合性, 在时空维度上任一节点会受到历史时刻相互连通节点的影响, 由于时间滞后性, 节点所受到的影响也是经过一段时间后才能传导并作用到该节点, 现有的研究没有考虑到时滞性的问题, 对交通流数据的时空特性挖掘不够充分。

本文采用无向图构建交通网络拓扑结构, 预测模型分别对交通流的三个时间分量建模, 包括近期周期分量、日周期分量和周周期分量以挖掘序列数据的时空相关性, 模型内包括三个重要的部分: 1) 采用时间和空间注意机制捕获数据中不同时间和位置之间的时间和空间动态相关性; 2) 采用切比雪夫图卷积神经网络获取交通流数据空间依赖关系, 不单独依靠节点之间的静态特性, 其中提出动态节点更新, 据此捕获隐藏的空间依赖关系; 3) 采用空洞卷积获取交通流数据时间依赖关系, 扩展感受野范围, 捕获多尺度时间变化特征。设计中为探索时间和空间的交互, 三个分量的输出被加权融合以生成最终预测结果。

## 2. 相关工作

### 2.1. 图卷积网络

图卷积网络(THOMAS KIPF, 2016)是卷积神经网络和图神经网络的衍生, 与传统的深度卷积神经网络(Convolutional Neural Network, CNN)不同, 图卷积神经网络(Graph Convolutional Networks, GCN)的计算核心是利用节点间的邻接关系进行邻域特征聚合, 现有对于图卷积网络的研究也主要关注利用拓扑图结构设计更合理的邻域聚合策略, 以交通流拓扑结构为例, 交通流数据在空间和时间上呈现相关性, 同时空间拓扑结构与时间连续性吻合图卷积网络的构建机制。如下图 1 所示为交通流的时空相关性示意图, 图中五个节点之间的关系共同构成拓扑结构, 根据节点 A 与其它节点之间的拓扑关系获得空间特征, 同时在时间上从  $t_1$  到  $t_2$ ,  $t_1$  时刻节点 A 与  $t_2$  时刻所有节点在时间上具有关联性,  $t_1$  时刻节点 A 会对  $t_2$  时刻所有节点产生不同程度的影响, 与此同时  $t_2$  时刻也存在节点之间的相互影响。



**Figure 1.** Schematic representation of spatio-temporal correlation of traffic flow  
**图 1.** 交通流时空相关性示意

一般的卷积以卷积核为核心, 在网格数据上滑动, 与此同时卷积核与其上的数据相乘相加后得到输出结果, 此外图神经网络中将数据通过图的形式进行记忆存储, 与卷积网络中数据的排布方式不同, 图神经网络中数据的关联性较强, 由于交通流数据具有一定周期性, 因而图神经网络的数据存储形式更佳。

图卷积网络分为基于频谱和基于空间的方法, 基于频谱方法从图信号的处理中引入放缩步骤, 傅里叶变换使时域信号转换为频域信号, 经过放缩后的频域信号通过傅里叶逆变换转为时域信号。

其中图卷积神经网络通式如式(1.1)所示:

$$H^{(l+1)} = f(H^{(l)}, A) \tag{1.1}$$

上式中  $l$  是指卷积层的层数, 初始的输入为  $H^0 = x$ , 特征矩阵  $x \in R^{N \times C}$  其中  $N$  是传感器数量即节点数量,  $C$  是特征向量的维数即每个节点采集到的特征数,  $A$  作为邻接矩阵(Adjacent Matrix)反应节点之间的相关性, 通过传感器采集到的数据特征形成多维数据, 每一个卷积层的输入都是邻接矩阵与节点的特征值。

如式(1.2)所示为逐层传播规则的形式, 其中  $W^{(l)}$  作为权重矩阵,  $\sigma(\cdot)$  作为非线性激活函数, 该模型考虑到邻接节点与节点特征之间的关系, 其中  $AH^{(l)}$  相乘等价于该节点与所有的邻接节点的特征相加,  $l$  层隐藏层叠加形成邻接节点的特征信息。

$$f(H^{(l)}, A) = \sigma(AH^{(l)}W^{(l)}) \tag{1.2}$$

考虑到以下几点, 首先对于每个节点, 所汇总的相邻节点所有特征向量, 而不是节点本身, 会忽略自身的影响, 其次邻接矩阵未经过归一化, 与特征矩阵相乘会改变原来分布, 产生一些不可预测的问题, 例如梯度消失等问题, 标准化的目的就是为了消除潜在的风险。因此对于前者在  $A$  的基础上加上单位矩阵得到  $\hat{A} = A + \lambda I$ , 当  $\lambda$  取值为 1 时, 意味着当前节点本身的特征与邻接节点特征具有相同的重要性, 此时  $\hat{A} = A + I$ ; 对于后者而言, 在式(1.3)中引入对称归一化的形式。

$$f(H^{(l)}, A) = \sigma\left(D^{-\frac{1}{2}}\hat{A}D^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \tag{1.3}$$

(1.3)式中  $\hat{A} = A + I$ , 其中  $I$  为单位矩阵(Unit Matrix),  $\hat{D}_{ii} = \sum_j \hat{A}_{ij}$ , 其中  $\hat{D}$  作为  $\hat{A}$  的度矩阵(Degree Matrix)。

$$f(H^{(l)}, A) = \sigma\left(\hat{D}^{-\frac{1}{2}}(D - A)\hat{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \tag{1.4}$$

## 2.2. 注意力机制

注意力机制是深度学习领域重要概念之一, 其灵感来源于人类的生物系统, 这些系统在处理大量信息时往往会专注于独特的部分。随着深度神经网络的深入发展, 注意力机制已经广泛应用于各类领域。注意力机制作为一种数据资源分配方案, 是解决信息过载问题的主要手段之一, 同时在计算能力有限的情况下, 可以用有限的计算资源处理有价值的信息。

注意力机制广泛使用在如图像标题生成、文本分类、机器翻译、动作识别、语音识别等领域, 在图像处理领域, 谷歌 Deep Mind [16]将注意力机制运用到循环神经网络中对图像分类处理, 该团队后使用两层的循环神经网络结合注意力机制, 这两项工作将注意力机制这一新思路引入该领域; 在语音识别领域, Hsiao 等人[17]将注意力机制融入到深度递归神经网络中, 用于发现更有价值的信号; 在图像标题生成领域, 注意力机制为该领域带来新的变革, Xu 等人[18]在解码器加入注意力机制, 目的是为了预测的信息更加关注图像中重要的部分, 而不是漫无目的关注整个图像信息。

在交通流预测方面, 注意力机制更多关注其中时间和空间上更加重要的数据信息, 提取交通流数据中的关键信息, 从而更好地预测交通状况。考虑到交通路网的拓扑结构以及时空动态关系, 本文中使用时间与空间注意力机制用于处理多维交通流数据, 精细化关注重要构成数据, 以提高计算效率与最终的预测精度。

## 3. 关键点分析

### 3.1. 问题定义

**定义 2.1** 根据交通拓网络进一步刻画节点之间的连通性以及互相影响程度, 构建表征时空关系空间图结构, 引入时间维度, 定义在  $t_1$  时刻无向图  $G_{t_1} = (V, E, A)$ , 其中  $V = \{v_1, v_2, \dots, v_N\}$  为无向图节点集合,  $N$  为节点数量,  $E = \{(v_i, v_j) | v_i, v_j \in V\}$  作为无向边集, 用于表明节点之间连通关系,  $A^d \in R^{N \times N}$  作为无向图  $G$  的邻接矩阵,  $d$  代表节点之间的欧氏距离, 如式(2.1)为无向图邻接矩阵的计算公式:

$$A_{i,j}^d = A_{j,i}^d \begin{cases} 0, & d_{i,j} = 0 \\ \frac{1}{d_{i,j}}, & d_{i,j} \neq 0 \end{cases} \quad (2.1)$$

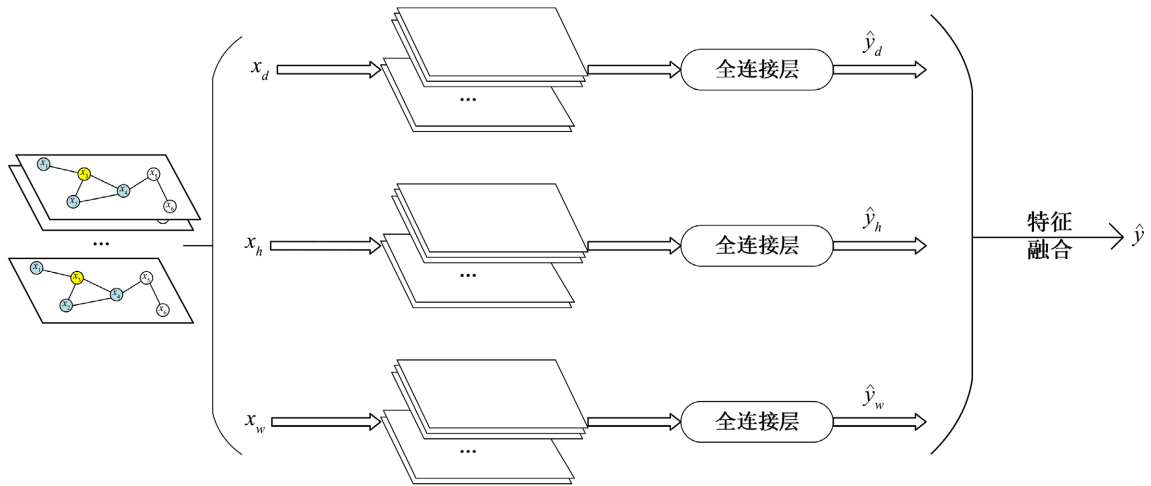
其中通过节点之间距离和连通性共同决定邻接矩阵, 0 表示节点之间无连通性,  $1/d_{i,j}$  表示节点之间的连通性, 距离越远则连通性越差, 符合实际的交通状况。

**定义 2.2** 节点属性特征矩阵, 表示为  $X \in R^{N \times F \times T}$ , 其中  $T$  表示过去历史时间长度,  $F$  表示交通路况特征的数量,  $X_i \in R^{N \times F}$  表示在过去时间  $t$  的第  $i$  个交通路况特征。

**定义 2.3** 交通流预测的目的是需要通过过去一段时间内的交通状况数据  $\{X_t | t=0, 1, \dots, n\}$ , 了解在未来时间  $(n + \Delta t)$  内交通状况信息  $X_{t+\Delta t} \in R^{N \times F}$ , 其中  $n$  是历史时间序列中最后一次观测到的时间点, 而  $\Delta t$  是需要预测的未来时间段。

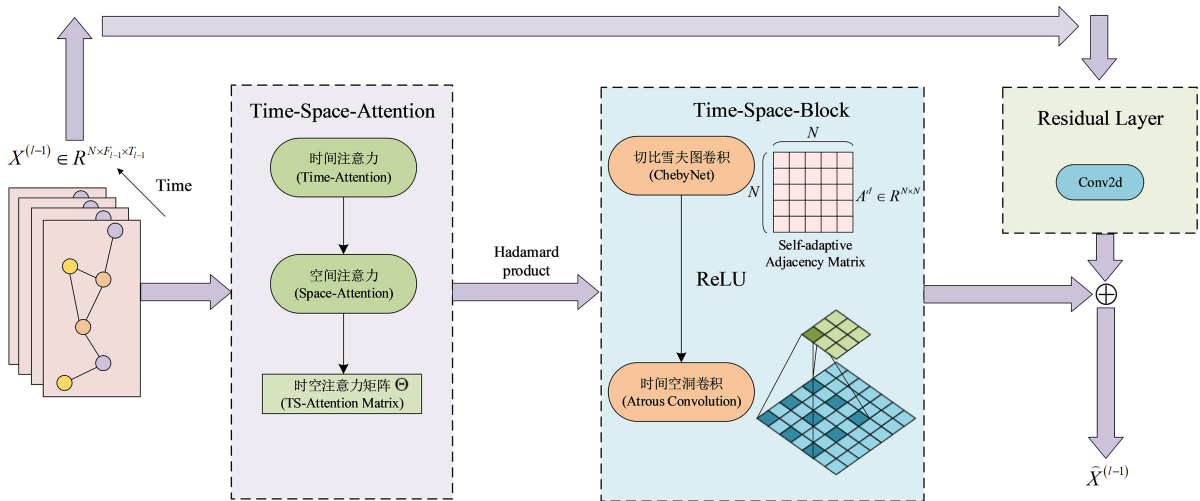
### 3.2. 模型总体框架

首先选取输入张量, 在本文中选取最近、日、周数据分量  $(x_h, x_d, x_w)$  三部分进行训练, 选取方法和选取过程在后续进行详细解释, 目的是获得周期性特征信息, 经过时空模块, 后续经过全连接层处理得到各部分的输出  $(\hat{y}_h, \hat{y}_d, \hat{y}_w)$ , 最终将每个部分的输出进行加权融合后得到最终的预测张量  $\hat{y}$ , 如图 2 所示为预测模型示意图。



**Figure 2.** Schematic of the prediction model  
**图 2.** 预测模型示意图

其中时空模块是模型主要组成部分, 如图 3 所示, 输入的交通流特征与时间特征应用时间和空间注意力机制, 以捕捉节点之间的时空动态特性, 处理好得到的时空注意力矩阵被利用到空间与时间关系建模中, 以获取时空依赖特性, 对于空间依赖, 以节点邻接静态关系为先验, 结合切比雪夫图卷积实现空间特征变化, 另参数化邻接矩阵, 实现动态更新效果进一步挖掘路网中的隐含空间关联; 之后基于时间空洞卷积对时间依赖建模, 提取有效的动态时序变化关系。时空注意力部分与时空模型部分共同构成一个整体, 多个块堆积进一步扩大时空相关性提取的范围, 经过多次时间和空间特征变化, 通过残差块连接, 残差块的作用是将浅层提取的特征, 通过跨层连接的方式传递给深层, 使得深层部分获得浅层信息, 起到了特征信息补充的作用以及防止梯度消失问题出现。



**Figure 3.** Spatio-temporal module structure  
**图 3.** 时空模块结构

### 3.3. 数据信息处理

#### 3.3.1. 数据初步处理

由于交通数据在宏观层面上具有一定的周期, 考虑数据周期性可以提高模型的预测精度。这里使用

每日周期数据作为每日周期分量的输入。然而, 每天时段的交通流量变化仍存在细微差异, 尤其是周末和工作日的交通流量趋势之间的差异。此外, 通过观察每周同一天的数据变化趋势, 可以发现数据中存在周周期的特征。因此, 为了弥补日周期的不足, 这里使用周周期数据作为周周期分量的输入。

对于日周期分量的步长选择, 一般来说多数的模型会采用过去两天的同时段数据作为日周期分量, 通过分析发现, 如下图 4 所示为 2021 年 1 月 2 日至 2021 年 1 月 8 日加利福尼亚州地区连续 24 小时内某交通传感器所采集到的当地交通流量变化情况, 通过图上粗略可以看出, 除了 DAY6(周六)与 DAY7(周日)和其他时间的交通流量在 4 时至 8 时存在较大差异, 周末的部分时段流量远少于工作日, 其余数据存在周期性特性。

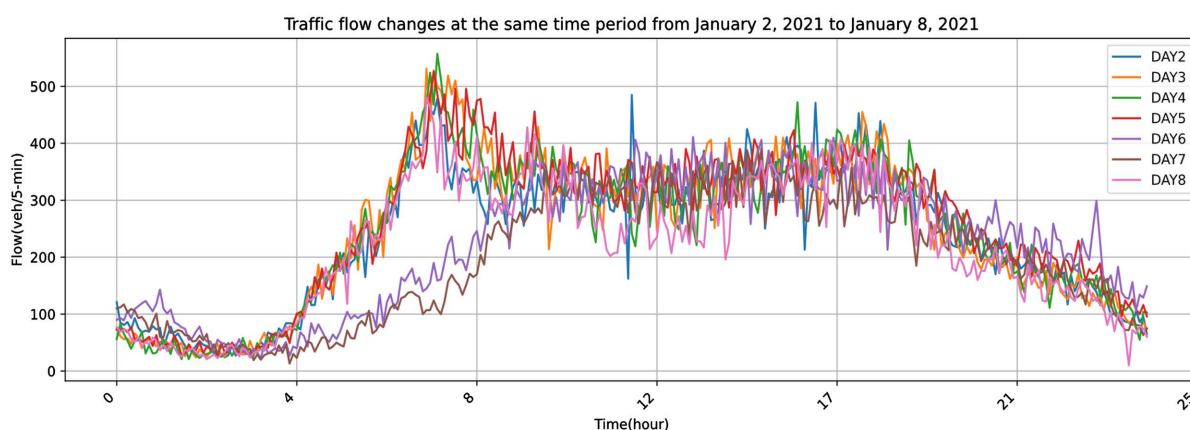


Figure 4. The 24-hour traffic flow changes at the same time and place on seven consecutive days

图 4. 连续七天同时刻、地点 24 小时交通流量变化

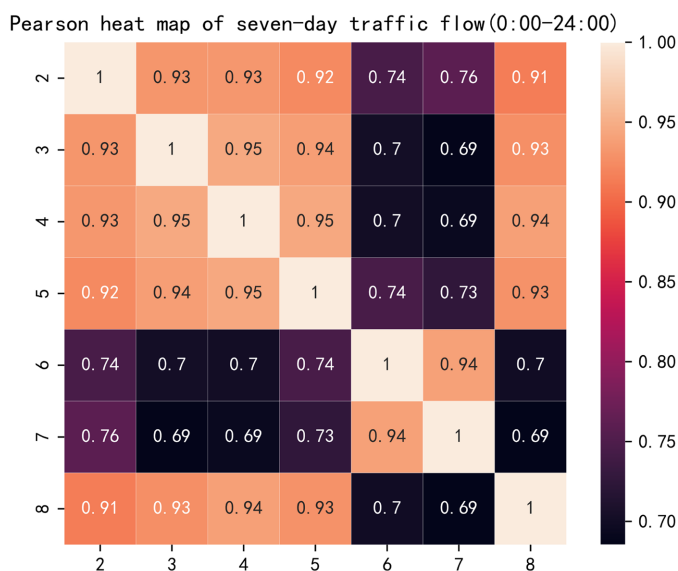


Figure 5. Pearson heat map of 24-hour traffic flow at the same time and place for seven consecutive days

图 5. 连续七天同时刻、地点 24 小时交通流量 Pearson 热力图

对上述的数据分析其中的相似度, 现有衡量向量的相似度方法多数基于欧氏空间, 无法考虑不同变量间取值的差异, 故在这里使用 Pearson 相关性系数度量数据之间的相似性, 其计算输出范围为-1 至+1,

0 代表无相关性, 正负值分别表示正相关与负相关。

如图 5 所示通过分析全天 24 小时交通流量的 Pearson 热力图发现, 周期内大多数数据存在一定的相关性, 周末的数据与工作日数据具有周期关联性, 而周末两天之间及工作日之间的数据关联程度更高。

在周期时间内, 早高峰时段(7:30~9:30)与晚高峰时段(17:00~19:00)的 Pearson 热力图中发现, 早高峰周末时段的关联性远低于全时段, 呈现负相关, 而晚高峰时段各数据之间呈现正相关, 数据中的 DAY6 与其他数据关联程度较低, 图 6 为早、晚高峰交通流量 Pearson 热力图。

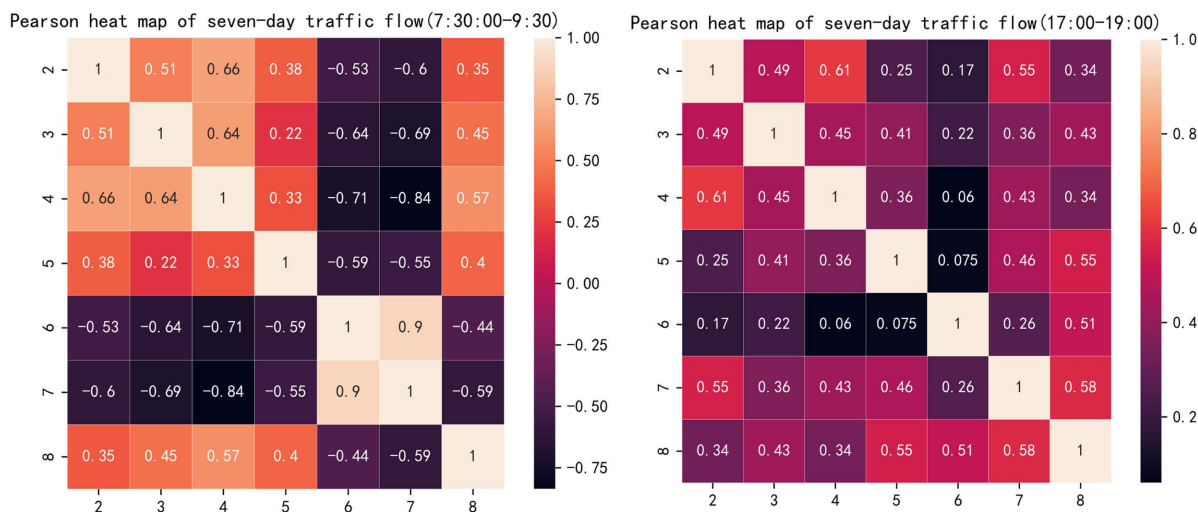


Figure 6. Pearson heat map of morning and evening peak traffic flow at the same time and place for seven consecutive days  
图 6. 连续七天同时刻、地点早、晚高峰交通流量 Pearson 热力图

综上所述, 根据连续七天全时段以及早晚高峰时段交通流量数据, 其中除早高峰时段外呈现周期性关联, 若仅考虑两天的数据可能因为非周期性的因素占比较大导致预测精度变低, 故在此选取预测时间段前 7 日数据作为历史周期数据。

在这里假设传感器采样的频率是每天  $z$  次, 那么每天所采集到的时间序列长度为  $z$ , 假设当前的时间点为  $t_0$ , 如下图 7 所示, 需要预测的未来时间窗口长度为  $T_p$ , 并且  $T_h$ 、 $T_d$ 、 $T_w$  作为最近、每日、每周历史时间数据分量的长度, 根据周期性的关系, 设定  $T_h = 2T_p$ 、 $T_d = 7T_p$ 、 $T_w = 2T_p$ 。

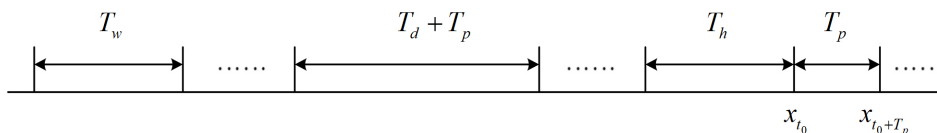


Figure 7. Schematic representation of segmented processing of input data  
图 7. 输入数据分段处理示意图

其中与预测区间相邻的时间序列  $X_h = (x_{t_0 - T_h + 1}, x_{t_0 - T_h + 2}, \dots, x_{t_0 - 1}, x_{t_0}) \in R^{N \times F \times T_h \times T_p}$ , 考虑到交通流量的变化趋势, 拥堵过程也是随时间的累积而形成, 直观上来看邻近时间段的交通状况始终会影响到未来的数据变化。

考虑到日常生活规律的影响, 例如工作日早晚高峰具有相似的交通数据特征; 同时每周同一天的交通模式也具有一定的周期性, 例如周末的交通数据与历史周末交通数据有较强的相似性, 所以考虑将日



周期分量、周周期分量作为补充, 以弥补最近时间分量的不足。

日周期分段数据可以表示为  $X_d = (x_1, x_2, \dots, x_{i-1}, x_i) \in R^{N \times F \times T_d \times T_p}$ , 其中  $x_i$  表示为式(2.2)

$$x_i = \left( x_{t_0 - \left(\frac{T_d}{T_p} - i - 1\right) * z + 1}, \dots, x_{t_0 - \left(\frac{T_d}{T_p} - i - 1\right) * z + T_p} \right) \quad (2.2)$$

周周期分段数据可以表示为  $X_w = (x_1, x_2, \dots, x_{i-1}, x_i) \in R^{N \times F \times T_w \times T_p}$ , 其中  $x_i$  表示为式(2.3)

$$x_i = \left( x_{t_0 - \left(\frac{T_w}{T_p} - i - 1\right) * z + 1}, \dots, x_{t_0 - \left(\frac{T_w}{T_p} - i - 1\right) * z + T_p} \right) \quad (2.3)$$

### 3.3.2. 滑动窗数据融合

滑动窗口构建训练集、测试集与验证集, 构建可变长度的窗口遍历指定长度的字串, 窗口每行进一个步长得到新的窗口数据, 直到最后一个窗口, 从而形成完整的数据集合。设定滑动窗口对应的大小分别为数据分量的长度  $T_w$ 、 $T_d$ 、 $T_h$ , 滑动步长为 1, 滑动过程中融合数据分量, 具体的融合过程如表 1 所示。

**Table 1.** Data fusion

**表 1.** 数据融合

#### Algorithm 1: 数据融合

**Input:** 数据集  $\text{Sample}[X_h, X_d, X_w] = \text{tar}$ ,  $\text{shape}(\text{Sample}) = (T, N, F)$   
**Output:**  $\text{Sample}[X'_h, X'_d, X'_w]$

1. **for**  $i = 1 \rightarrow i = S$  **do**
2. **while**  $\text{Sample}[X_h, X_d, X_w] \neq \Phi$  **do**
3.     **if**  $T_w > 0$  **then**
4.      $\text{shape}(\text{Sample}[X'_w]) = (1, N, F, T_w) \leftarrow$  转置、扩展数组维度
5.      $\text{Sample} \leftarrow$  append  $\text{Sample}[X'_w]$
6.     **if**  $T_d > 0$  **then**
7.      $\text{shape}(\text{Sample}[X'_d]) = (1, N, F, T_d) \leftarrow$  转置、扩展数组维度
8.      $\text{Sample} \leftarrow$  append  $\text{Sample}[X'_d]$
9.     **if**  $T_h > 0$  **then**
10.      $\text{shape}(\text{Sample}[X'_h]) = (1, N, F, T_h) \leftarrow$  转置、扩展数组维度
11.      $\text{Sample} \leftarrow$  append  $\text{Sample}[X'_h]$
12.      $\text{shape}(\text{tar}') = (1, N, T) \leftarrow$  转置、扩展数组维度
13.      $\text{Sample} \leftarrow$  append  $\text{tar}'$
14. **end while**
15. **end if**

### 3.3.3. 基准自适应机制

由于同一时间段下的交通模式并不是高度相关的, 某一群体的非周期性出行可能会引起当天交通状况波动变化, 例如图 4、图 5 所示, 周末的早高峰时段流量远低于工作日时段流量, 故为了方便模型学习, 设计数据替代方式以更好的引导训练。

在设计中添加基准自适应机制用于处理同一时间段交通模型的较大差异, 选择当前时间  $t_0$  之前的一小时数据作为关键时间段  $X_d$ , 以便于计算相关性系数  $r$ , 如式(2.5)所示计算  $x_h$  与

$X_d = (x_1, x_2, \dots, x_{i-1}, x_i) \in R^{N \times F \times T_d \times T_p}$  中  $x_i$  的相关性系数, 其中  $z$  是 24 小时内获取到的时间序列数据的长度。

$$X_h = \left( X_{t_0 - \frac{z}{24} + 1}, X_{t_0 - \frac{z}{24} + 2}, \dots, X_{t_0} \right) \in R^{N \times F \times \frac{z}{24}} \quad (2.4)$$

$$r = \frac{z \sum x_h x_i - \sum x_h \sum x_i}{\sqrt{z \sum x_h^2 - (\sum x_h)^2} \sqrt{z \sum x_i^2 - (\sum x_i)^2}} \quad (2.5)$$

通过计算得到的相关性系数得到平均值  $\bar{r}$ , 选取  $r$  值最高的一组数据以替代低于平均值的多组数据, 替换后的数据以作为后续模型的输入。

## 4. 多传感器时空图卷积网络

### 4.1. 注意力模块

注意力机制通过模拟人脑注意力的特点, 对其中重点的信息给予更多关注, 由于交通流的时序关联与空间交错特性, 呈现出复杂的非线性, 在不同时段和空间具有不同的依赖特性。某一局部路网的交通状态, 不仅和自身交通状态变化相关, 还会收到其他邻近节点的影响, 时序数据也会收到邻近时刻与较远时刻的影响。

为减少卷积层较少的影晌, 扩大捕捉数据窗口, 时序上考虑到较远时刻隐含的高阶时序特征, 高阶特征关系所呈现的不仅仅是单节点时序特征, 同样反应空间节点之间在时序变化中的动态影响, 而空间上存在需要进一步挖掘高阶动态空间特征, 以捕捉节点之间的动态时空特性, 故依次应用时间和空间注意力机制, 以提升模型性能。在时间与空间注意力机制中, 第  $l$  层注意力计算对用输出结果的计算过程如下。

$$Q^{l-1} = \left( X^{(l-1)} \right)^T W_h^I \quad (3.1)$$

$$K^{l-1} = W_h^{II} X^{(l-1)} \quad (3.2)$$

$$H^p = V_h \cdot \text{sigmoid} \left( b_h + Q^{l-1} W_h^{III} K^{l-1} \right) \quad (3.3)$$

$$H'_{i,j} = \frac{\exp(H_{i,j}^p)}{\sum_{j=1}^{T_{l-1}} \exp(H_{i,j}^p)} \quad (3.4)$$

上式中  $X^{(l-1)} = (x_1, x_2, \dots, x_{T_{l-1}}) \in R^{N \times F_{l-1} \times T_{l-1}}$  为  $l$  层输入特征,  $Q^{l-1}$ 、 $K^{l-1}$  分别作为中间向量,  $V_h \in R^{T_{l-1} \times T_{l-1}}$ 、 $b_h \in R^{T_{l-1} \times T_{l-1}}$ 、 $W_h^I$ 、 $W_h^{II}$ 、 $W_h^{III}$  为可学习参数,  $H^p$  为输入时空数据之间的注意力得分矩阵, 经过归一化后得到  $H'_{i,j}$ , 表征第  $i$  个时空节点与第  $j$  个时空节点之间的影响程度, 经过归一化的注意力得分矩阵与输入时序数据相乘得到  $\tilde{X}^{(l-1)}$ 。

$$\tilde{X}^{(l-1)} = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \dots, \tilde{x}_{T_{l-1}}) = (x_1, x_2, x_3, \dots, x_{T_{l-1}}) H' \quad (3.5)$$

经过时序注意力机制处理后得到的  $\tilde{X}^{(l-1)}$ , 作为空间注意力机制的数据输入, 经过时序与空间注意力机制处理得到的矩阵  $\Theta$ , 在后续建立良好的时空依赖关系中作为重要的组成部分, 丰富时空关联以提升

网络模型的整体时空关联性能。

## 4.2. 空间关系模块

对交通路网中不同位置空间关联状态存在相互影响, 包括车流以及平均车速等信息, 从空间拓扑结构看来交通特征存在空间相关性, 并且随着时序等多重因素的作用呈现复杂的动态变化趋势。而基于频域的图卷积神经网络能够借助图信号处理的方法捕获交通路网的空间依赖关系, 提取交通路网中的空间特性。

由于基于频域的图卷积神经网络自身复杂度较高, 傅里叶变换的存在让图卷积更为复杂, 通常使用切比雪夫多项式以降低其复杂度, 在多项式函数逼近邻域时, 切比雪夫多项式由于其数值稳定性以及计算高效性作为更加优先的选择。

在图理论中, 分析图的拉普拉斯矩阵的特征值以及特征向量来分析研究图的性质, 其中图的拉普拉斯矩阵表示为  $L = D - I$ , 在这里定义  $L = I_N - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \in R^{N \times N}$ , 拉普拉斯矩阵在频域被分解为  $L = U \Lambda U^T$ , 其中  $\Lambda = \text{diag}([\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_{N-1}]) \in R^{N \times N}$ , 由拉普拉斯矩阵的特征值组成确定,  $U$  是对称归一化的拉普拉斯算子。其中  $k$  阶切比雪夫多项式可以表示为:

$$\text{chabnet}(X) = \sigma \left( \sum_{k=0}^{k-1} \beta_k T_k(\tilde{L}) X \right) \quad (3.6)$$

$$T_k(\tilde{L}) = 2\tilde{L}T_{k-1}(\tilde{L}) - T_{k-2}(\tilde{L}) \quad (3.7)$$

$$\tilde{L} = \frac{2L}{\lambda_{\max}} - I_N \quad (3.8)$$

$$T_0(\tilde{L}) = 1, T_1(\tilde{L}) = \tilde{L} \quad (3.9)$$

切比雪夫图卷积能够通过调整  $k$  的值, 改变卷积感受野范围, 确定有效的空间特征影响范围, 以提高模型空间关系挖掘的能力。为了能够动态调节每个节点之间的影响程度, 将经过时间与空间注意力矩阵处理得到的矩阵  $\Theta$  与  $T_k(\tilde{L})$  相乘, 通过卷积变换得到  $\hat{G}$ 。

$$\hat{G} = \text{chabnet}(X)' = \sigma \left( \sum_{k=0}^{k-1} \beta_k (T_k(\tilde{L}) \odot \Theta) X \right) \quad (3.10)$$

### 动态邻接矩阵

通常的邻接矩阵以静态形式作为反映节点间关系的存在, 而真实情况下不同节点之间呈现错综复杂的层次关联, 随着时间而动态变化, 显然静态形式的邻接矩阵不能反映出节点空间真实的动态特性。而受到文献的启发, 利用自适应邻接矩阵, 以补充交通路网中动态空间依赖, 同时对卷积操作所丢失的空间特征信息进行补足。自适应邻接矩阵的构建, 使得模型结构更为可靠, 其矩阵计算过程如下:

$$M_1 = \tanh(\beta E_1 \theta_1) \quad (3.11)$$

$$M_2 = \tanh(\beta E_2 \theta_2) \quad (3.12)$$

$$A_{\text{adp}} = \text{SoftMax} \left( \text{ReLU} \left( \tanh \left( \beta (M_1 M_2^T - M_2 M_1^T) \right) \right) \right) \quad (3.13)$$

其中,  $E_1$  和  $E_2$  表示路网中节点表示向量,  $\theta_1$  和  $\theta_2$  为节点向量变换的参数,  $\beta$  为激活函数饱和率参数,  $M_1$

和  $M_2$  为中间变量矩阵, 计算得到结果  $A_{adp}$ , 结合定义(2.1)中邻接矩阵的定义, 与当前节点没有空间关系的节点在邻接矩阵中取值为 0, 进一步得到处理后的自适应邻接矩阵  $\tilde{A}_{adp}$ 。

### 4.3. 时序关系模块

交通流数据的本质是时间序列数据, 时间维度上道路交通状况呈现不同变换特点, 时序关系的建立旨在有效提取输入数据的动态时序变化趋势。现主流的处理方法多数为 RNN 的变体, 例如 LSTM、GRU 等时序关系处理方法, 为减少误差积累影响预测性能, 在这里采用时序卷积网络结合空洞卷积结构, 空洞卷积进一步增加感受野范围, 避免网络堆叠影响计算效率, 计算方式如下:

$$x^l = \text{relu}(\hat{G} * R + b + (x^{l-1} * T + c)) \quad (3.14)$$

$$x^l * f(t) = \sum_{s=0}^{K_f-1} f(s) x^l(t - d_f \times s) \quad (3.15)$$

其中 \* 表示标准卷积运算,  $R$ 、 $T$  分别代表不同大小的卷积核,  $d_f$  为空洞因数, 代表卷积核相邻单元间隔的距离, 最后添加二维卷积层作为残差连接。

### 4.4. 时空特征融合

设计中采用三个时间周期分量融合组成网络模型, 分别提取周周期分量、日周期分量、邻近周期分量的交通流数据, 经过多个 XX 块的堆叠进一步提取更大范围的动态时空相关性, 后采用全连接层确保三个输出具有与目标相同的维度和形状, 使用 ReLU 作为激活函数。

经过预测后得到的预测值, 融合三者进一步实现自适应控制不同的分量占比, 如式(3.16)所示, 其中  $\odot$  为哈达玛积,  $W_d \in R^{T_p \times T_p}$ ,  $W_h \in R^{T_p \times T_p}$ ,  $W_w \in R^{T_p \times T_p}$ 。

$$\hat{Y} = W_d \odot Y_d + W_h \odot Y_h + W_w \odot Y_w \quad (3.16)$$

## 5. 实验设计与分析

### 5.1. 数据集

设计中采用两个公开数据集作为性能评估, 分别是 PeMSD4 数据集和 PeMSD8 数据集, 这些数据来自 Caltrans Performance Measure System, 其中原始的数据被聚合成每 5 分钟的间隔, 其中 PeMSD4 数据集包含 307 个传感器, 时间跨度为 2018 年 1 月 1 日~2018 年 2 月 28 日, PeMSD4 数据集包含 170 个传感器, 时间跨度为 2016 年 7 月 1 日~2016 年 8 月 31 日。

因此, 每条数据包括流量、平均车速、平均占用率三个特征, 按照时间顺序将时空数据划分, 前 80% 的数据作为训练集, 10% 作为测试集, 10% 作为验证集。其中缺失值通过线性插值的方法填充, 将每个特征数据进行零均值化, 使得数据集平均值为 0, 在反向传播时加快网路中每层权重参数的收敛, 如式(4.1)所示。

$$X' = X - \text{average}(X) \quad (4.1)$$

### 5.2. 实验环境与对比基线

深度学习模型中超参数的选择, 直接影响到模型的学习效率与预测性能。本文使用网格搜索方法选择超参数, 即设置多种超参数组合并进行遍历, 并选择在验证集中性能最优的一组超参数。本文实验中, 批量大小(Batch Size)设置为 32, 学习率(Learning Rate)设置为 0.001, 图卷积阶数 K 设置为 3, 图卷积和空洞卷积设置为 64 层, 使用 Adam 优化器对模型参数进行优化。实验部分均在 Windows 操作系统下,

系统配置 Inter(R) Core(TM) i9-10900X @ 3.70GHz GPU: NVIDIA GeForce RTX 3090, 在 Pycharm 集成开发环境下编译, 基于 Pytorch 框架和 Python3.9 实现模型的预测。

实验中设置多组对比基线模型, 其中包括:

- 1) HA (历史平均值模型), 基于时间周期性对时间序列建模, 预测值为计算历史周期中的相同时序位置数据均值。
- 2) ARIMA (自回归移动平均模型), 非平稳时间序列经过差分处理后显示出同质性, 利用移动平均和回归方法实现预测。
- 3) SVR (支持向量机回归), 通过核函数降低非线性数据维度, 其中核函数直接影响到模型的性能。
- 4) LSTM (长短期记忆力网络): 一种循环神经网络模型的变种。
- 5) STGCN: 时空图卷积网络, 利用图卷积与纯卷积结构加以预测。
- 6) ASTGNN: 注意力机制与动态图卷积相结合, 捕捉空间异质性与周期性。
- 7) DCRNN: 扩散图卷积神经网络, 引入编码器-解码器结构, 通过神经网络捕捉时间依赖。
- 8) Graph WaveNet: 自适应依赖矩阵与卷积模块分别捕捉隐藏空间依赖与时间相关性。

### 5.3. 评价指标

为了能够评价模型的性能, 采用以下三种度量函数作为评测模型性能的依据。

平均绝对误差(Mean Absolute Error, MAE)

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (4.2)$$

均方根误差(Root Mean Square Error, RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (4.3)$$

平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (4.4)$$

其中  $\hat{y}_i$ 、 $y_i$  分别为预测值与观测值,  $n$  作为表示样本数量。

## 5.4. 实验数据分析

### 5.4.1. 预测性能分析

实验结果如表 2 和表 3 所示, 在两个数据集上对比不同基线与模型在多步预测下的实际性能, 性能评价指标采用 MAE、RMSE、MAPE。其中在 PEMS4 数据集上模型在预测效果大多高于其他基线, 在 PEMS8 数据集上模型具有较佳的预测效果。

综上所述, 文中提出的模型优于其他基线, 与近期的研究模型 ASTGCN、DCRNN、Graph WaveNet 相比较能够占据一定的优势。对于其他一些传统的预测方法, 例如 HA 等以及 LSTM 在长期预测中的表现并不佳, 其分别忽略了时间相关性与时空动态相关性, 而 STGCN、DCRNN、Graph WaveNet 并没有引入注意力机制以关注时空动态变化导致预测效果不佳, ASTGCN 中采用节点固定距离作为邻接矩阵, 未能考虑空间动态特性。因此, 本文模型在预测效果上优于现有模型, 具有良好的预测效果。

**Table 2.** Prediction effect of PeMSD4 dataset in different models  
**表 2.** PeMSD4 数据集使用不同模型的预测效果

Methods	15 min			30 min			60 min		
	MAE	RMSE	MAPE%	MAE	RMSE	MAPE%	MAE	RMSE	MAPE%
HA	38.03	59.24	27.88	38.03	59.24	27.88	38.03	59.24	27.88
ARIMA	29.24	50.11	19.87	33.82	56.57	25.71	43.67	68.69	29.83
SVR	22.36	35.81	14.63	24.34	38.35	16.37	27.51	42.32	18.79
LSTM	21.74	33.41	13.87	23.30	36.39	15.49	28.67	44.45	18.36
STGCN	20.27	32.79	13.03	22.72	35.68	14.83	27.08	41.34	17.85
ASTGCN	20.15	31.43	14.03	22.09	34.43	15.47	26.03	40.02	19.17
DCRNN	20.75	32.24	13.10	19.89	31.61	13.79	21.67	34.00	15.41
Graph WaveNet	20.45	<b>29.24</b>	12.27	<b>19.12</b>	30.62	13.28	20.69	33.02	<b>14.10</b>
MINE	<b>20.12</b>	29.54	<b>12.14</b>	19.15	<b>29.93</b>	<b>13.11</b>	<b>20.51</b>	<b>32.91</b>	14.52

\*注意: 每列最优值用粗体标注。

**Table 3.** Prediction effect of PeMSD8 dataset in different models  
**表 3.** PeMSD8 数据集使用不同模型的预测效果

Methods	15 min			30 min			60 min		
	MAE	RMSE	MAPE%	MAE	RMSE	MAPE%	MAE	RMSE	MAPE%
HA	33.64	51.95	23.71	33.64	51.95	23.71	33.64	51.95	23.71
ARIMA	24.24	42.79	17.54	30.78	48.26	21.26	38.67	56.10	24.93
SVR	16.84	25.19	11.63	19.82	29.31	13.37	24.22	35.13	14.92
LSTM	16.58	25.22	11.87	19.07	29.23	13.29	23.91	35.78	14.87
STGCN	16.14	24.33	10.33	17.89	27.92	11.17	21.22	32.95	12.86
ASTGCN	16.48	25.09	11.03	18.66	28.17	12.23	22.83	33.68	15.24
DCRNN	15.02	23.36	9.65	16.56	25.88	10.62	19.39	29.88	12.45
Graph WaveNet	13.93	22.76	<b>9.15</b>	15.78	25.22	10.27	17.92	28.54	11.29
MINE	<b>13.85</b>	<b>22.56</b>	9.22	<b>15.74</b>	<b>24.87</b>	<b>10.11</b>	<b>17.51</b>	<b>28.32</b>	<b>10.89</b>

\*注意: 每列最优值用粗体标注。

#### 5.4.2. 参数选择对比

此外实验还比较了超参数设定所给模型性能带来的变化情况, 其中包括 Batchsize 的值以及 K 值的比较选择, 其中 K 值为切比雪夫图卷积网络的重要参数, K 值大小影响拉普拉斯矩阵计算次数, 其表示图的 K 阶近邻关系, 另外比较了在常见的三种批量大小中模型的实际表现。实验结果如图 8 和图 9 所示, 其中图卷积的阶数以及 Batchsize 值对预测结果影响相对较大, 当 K = 3 以及 Batchsize = 32 时模型具有良好的预测性能。

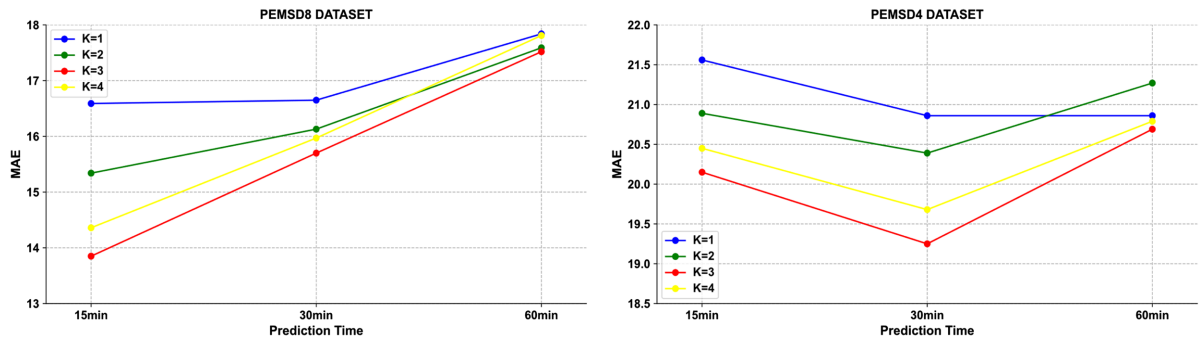


Figure 8. Prediction effect of different graph convolution order on datasets  
图 8. 不同图卷积阶数在数据集上的预测效果

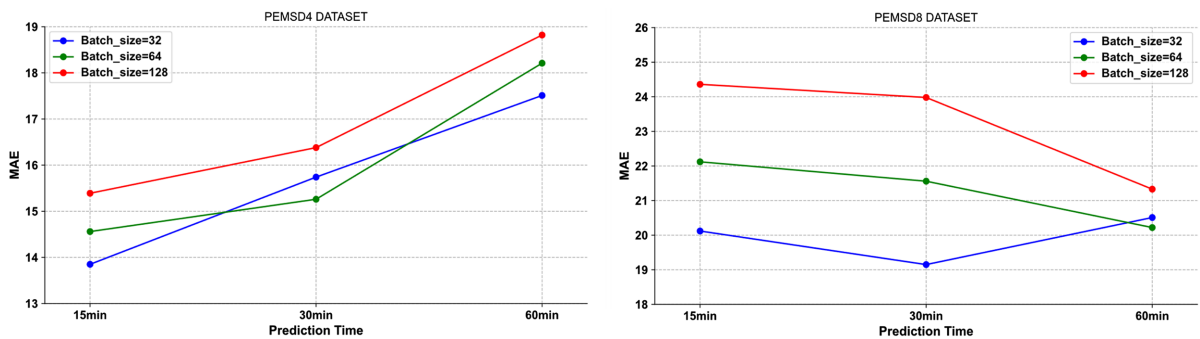


Figure 9. Prediction effect of different Batchsize on datasets  
图 9. 不同 Batchsize 值在数据集上的预测效果

### 5.4.3. 模型计算时间

前文提及到的个别模型与设计模型在同一数据集进行比较计算花销时间, 通过几种模型的平均计算时间的对比发现, 如表 4 所示在数据集 PEMS4 下, 该模型的训练速度较快, 其中训练阶段 Graph WaveNet 花费时间最长, 效率较低, STGCN 与 ASTGCN 花费时间比文中提出的模型长, 是因为在时空注意力分数计算阶段不断迭代更新; 在测试阶段, 由于 STGCN 需要多次迭代学习以生成多步长预测结果, 其他模型在一次运行后能够较快得出预测结果, 测试时间更短。

Table 4. The average calculation time of the models on the PeMSD4 dataset  
表 4. 模型在 PeMSD4 数据集上的平均计算时间

模型	计算时间	
	训练/s/epoch	测试/s
STGCN	50.25	96.56
ASTGCN	63.57	9.89
Graph WaveNet	180.3	5.44
MINE	60.21	7.43

## 6. 结语

本文中提出了一种用于交通流预测的模型, 其中模型结合了时空注意力机制、基准自适应机制以及

时空图卷积模块, 综合考虑到周期性数据之间的差异性, 提高输入数据的质量, 同时考虑到交通的时空动态特性以及节点之间的动态关系, 结合交通拓扑结构捕获时空动态相关性, 在两个真实数据集实验中取得良好的预测效果。未来的研究中将着眼于更多的外部扰动对于系统建模的影响, 围绕更加复杂的交通结构以实现更为精准的预测。

## 参考文献

- [1] Nagy, A.M. and Simon, V. (2018) Survey on Traffic Prediction in Smart Cities. *Pervasive and Mobile Computing*, **50**, 148-163. <https://doi.org/10.1016/j.pmcj.2018.07.004>
- [2] Van Der Voort, M., Dougherty, M. and Watson, S. (1996) Combining Kohonen Maps with ARIMA Time Series Models to Forecast Traffic Flow. *Transportation Research Part C: Emerging Technologies*, **4**, 307-318. [https://doi.org/10.1016/S0968-090X\(97\)82903-8](https://doi.org/10.1016/S0968-090X(97)82903-8)
- [3] Valipour, M. (2015) Long-Term Runoff Study Using SARIMA and ARIMA Models in the United States. *Meteorological Applications*, **22**, 592-598. <https://doi.org/10.1002/met.1491>
- [4] Kamarianakis, Y. and Prastacos, P. (2003) Forecasting Traffic Flow Conditions in an Urban Network: Comparison of Multivariate and Univariate Approaches. *Transportation Research Record: Journal of the Transportation Research Board*, **1857**, 74-84. <https://doi.org/10.3141/1857-09>
- [5] Lin, Z.S.C. and Zhaoming, C. (2013) Bayesian Network Model for Traffic Flow Estimation Using Prior Link Flows. *Journal of Southeast University (English Edition)*, **29**, 322-327.
- [6] Djenouri, Y., Belhadi, A., Lin, J.C.W. and Cano, A. (2019) Adapted K-Nearest Neighbors for Detecting Anomalies on Spatio-Temporal Traffic Flow. *IEEE Access*, **7**, 10015-10027. <https://doi.org/10.1109/ACCESS.2019.2891933>
- [7] Smith, B.L. and Demetsky, M.J. (1994) Short-Term Traffic Flow Prediction Models-A Comparison of Neural Network and Nonparametric Regression Approaches. *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, San Antonio, 2-5 October 1994, 1706-1709.
- [8] Yu, R., Li, Y., Shahabi, C., et al. (2017) Deep Learning: A Generic Approach for Extreme Condition Traffic Forecasting. *Proceedings of the 2017 SIAM international Conference on Data Mining*, Houston, 27-29 April 2017, 777-785. <https://doi.org/10.1137/1.9781611974973.87>
- [9] Koesdwiady, A., Soua, R. and Karray, F. (2016) Improving Traffic Flow Prediction with Weather Information in Connected Cars: A Deep Learning Approach. *IEEE Transactions on Vehicular Technology*, **65**, 9508-9517. <https://doi.org/10.1109/TVT.2016.2585575>
- [10] Wang, Q., Wan, J. and Yuan, Y. (2018) Locality Constraint Distance Metric Learning for Traffic Congestion Detection. *Pattern Recognition*, **75**, 272-281. <https://doi.org/10.1016/j.patcog.2017.03.030>
- [11] Shi, X.J., Chen, Z.R., Wang, H., et al. (2015) Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *Proceedings of the 28th International Conference on Neural Information Processing Systems*, **1**, 802-810.
- [12] Zhao, L., Song, Y., Zhang, C., et al. (2019) T-Gcn: A Temporal Graph Convolutional Network for Traffic Prediction. *IEEE Transactions on Intelligent Transportation Systems*, **21**, 3848-3858. <https://doi.org/10.1109/TITS.2019.2935152>
- [13] Hu, J., Lin, X. and Wang, C. (2022) DSTGCN: Dynamic Spatial-Temporal Graph Convolutional Network for Traffic Prediction. *IEEE Sensors Journal*, **22**, 13116-13124. <https://doi.org/10.1109/JSEN.2022.3176016>
- [14] 谷远利, 陆文琦, 李萌, 等. 基于组合深度学习的快速路车道级速度预测研究[J]. 交通运输系统工程与信息, 2019, 19(4): 79-86.
- [15] Kipf, T.N. and Welling, M. (2016) Semi-Supervised Classification with Graph Convolutional Networks. arXiv: 1609.02907.
- [16] 王婧娟, 陈庆奎. 一种时空周期性注意力网络的交通流量预测模型[J]. 小型微型计算机系统, 2022, 43(11): 2321-2327. <https://doi.org/10.20009/j.cnki.21-1106/TP.2021-0232>
- [17] Chen, C.C., Hasio, Y.T., Lin, C.Y., et al. (2017) Using Deep Learning to Predict and Optimize Hadoop Data Analytic Service in a Cloud Platform. *2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*, Orlando, 6-10 November 2017, 909-916. <https://doi.org/10.1109/DASC-PiCom-DataCom-CyberSciTec.2017.153>
- [18] Xu, H.L., Xia, W., Gao, Q.X., Han, J.G. and Gao, X.B. (2021) Graph Embedding Clustering: Graph Attention Auto-Encoder with Cluster-Specificity Distribution. *Neural Networks*, **142**, 221-230. <https://doi.org/10.1016/j.neunet.2021.05.008>