

基于延迟时空依赖的非平稳时间序列 交通流量预测模型

李小龙^{1,2}, 李 曦¹, 刘 洋¹, 李柄廷¹, 易畅言³, 曾宁俊^{1*}

(1. 湖南工商大学计算机学院, 湖南长沙 410205; 2. 湘江实验室, 湖南长沙 410205;
3. 南京航空航天大学计算机科学与技术学院, 江苏南京 211106)

摘 要: 建立精准的交通流量预测模型, 对于优化交通系统管理、缓解城市交通拥堵、提升路网运行效率具有至关重要的作用。然而, 实际交通流呈现出显著的非平稳特性与复杂的时空依赖关系, 尤其是由突发事件、早晚高峰、节假日等引起的流量分布偏移, 以及交通拥堵在路网中传播的延迟效应, 给传统预测方法带来了严峻挑战。现有模型大多基于平稳性假设或采用静态时空建模方式, 难以有效捕捉交通数据中的动态演化规律与异质性延迟依赖, 导致预测精度受限、工程适用性不足。针对上述问题, 本文提出一种基于延迟时空依赖的非平稳时间序列交通流量预测模型 (Non-Stationary time series Forecasting Model, NSFM), 旨在从频域和空域双重角度深入刻画交通流的动态演变机制。该模型首先利用傅里叶变换将非平稳时间序列分解为时变组分与时不变组分, 分别刻画局部动态波动与全局稳态趋势, 并通过正交性证明确保两类成分的独立性, 为后续差异化建模奠定理论基础。在此基础上, 模型进一步构建带有时延特征提取机制的特征融合模块, 结合逐点卷积与位置编码, 将交通流量、空间邻接关系、时间周期信息与延迟传播特征进行深度融合, 从而精准捕捉站点间交通状态的时空演化与滞后响应规律。为建模离散站点间的空间自相关结构, 本文引入 Moran 算子构建函数对函数回归预测框架, 通过基函数展开与正交化处理, 实现连续函数空间与离散观测站点之间的一致性映射, 有效量化区域间的空间依赖强度, 提升模型在复杂路网中的预测鲁棒性。为验证 NSFM 模型的有效性 & 泛化能力, 我们在四个真实世界交通流量数据集 (PEMS03, PEMS04, PEMS07, PEMS08) 上进行了系统实验, 实验表明, NSFM 在多个评估指标上均显著优于现有主流模型, 其中平均绝对百分比误差 (MAPE) 相较于 SOTA 模型, 分别降低了 7.48%、9.86%、3.20%、1.73%, 展现出在非平稳场景下更优的预测精度与稳定性。

关键词: 延迟时空依赖; 非平稳; 交通流量预测; 时空演变特征

基金项目: 湘江实验室重大项目 (No.23XJ01001, No.22XJ01001); 启元实验室创新基金 (No.2022-JCJQ-LA-001-088); 湖南省重点研发计划 (No.2024JK2007); 广西科技重大项目 (No.2024AA15007)

中图分类号: TP311.1

文献标识码: A

文章编号: 0372-2112(2025)11-4035-16

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250591

Non-Stationary Time Series Traffic Flow Forecasting Model Based on Delayed Spatio-Temporal Dependencies

LI Xiao-long^{1,2}, LI Xi¹, LIU Yang¹, LI Bing-ting¹, YI Chang-yan³, ZENG Ning-jun^{1*}

(1. Department of Computer Science, Hunan University of Technology and Business, Changsha, Hunan 410205, China;

2. Xiangjiang Laboratory, Changsha, Hunan 410205, China;

3. Department of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 211106, China)

Abstract: Building accurate traffic flow prediction models is crucial for optimizing traffic system management, alleviating urban congestion, and enhancing road network operational efficiency. However, real-world traffic flow exhibits significant non-stationary characteristics and complex spatio-temporal dependencies. In particular, the distribution shifts caused by unexpected events, rush hours, and holidays, coupled with the delayed propagation of traffic congestion across the network, pose severe challenges to traditional forecasting methods. Most existing models, relying on stationary assumptions or static spatio-temporal modeling, struggle to effectively capture the dynamic evolution patterns and heterogeneous

delayed dependencies within traffic data, leading to limited prediction accuracy and insufficient practical applicability. To address these limitations, this paper proposes a non-stationary time series traffic flow forecasting model based on delayed spatio-temporal dependencies (NSFM), designed to deeply characterize the dynamic evolution mechanisms of traffic flow from both frequency and spatial domains. The model first employs Fourier Transform to decompose the non-stationary time series into time-varying and time-invariant components, capturing local dynamic fluctuations and global steady-state trends respectively, with orthogonality proven to ensure the independence between the two components, laying a theoretical foundation for subsequent differentiated modeling. Furthermore, the model constructs a feature fusion module with a delay feature extraction mechanism, integrating traffic flow, spatial adjacency relationships, temporal periodic information, and delay propagation features through pointwise convolution and positional encoding, thereby accurately capturing the spatio-temporal evolution and lagged response patterns of traffic states between stations. To model the spatial autocorrelation structure among discrete stations, this paper introduces the Moran operator to build a function-on-function regression prediction framework. Through basis function expansion and orthogonalization processing, a consistent mapping between the continuous function space and discrete observation stations is achieved, effectively quantifying the spatial dependency strength between regions and enhancing the model's prediction robustness in complex road networks. To validate the effectiveness and generalization capability of the NSFM model, systematic experiments are conducted on four real-world traffic flow datasets (PEMS03, PEMS04, PEMS07, PEMS08). Experimental results demonstrate that NSFM significantly outperforms existing mainstream models across multiple evaluation metrics. Specifically, the mean absolute percentage error (MAPE) is reduced by 7.48%, 9.86%, 3.20%, and 1.73% respectively compared to SOTA models, demonstrating superior prediction accuracy and stability in non-stationary scenarios.

Key words: delayed spatio-temporal dependence; non-stationary; traffic flow prediction; spatio-temporal evolution characteristics

Foundation Item(s): Major Projects of Xiangjiang Laboratory (No.23XJ01001, No.22XJ01001); Innovation Fund of Qi-Yuan Laboratory (No.2022-JCJQ-LA-001-088); Key Research and Development Program of Hunan Province (No.2024JK2007); Major Science and Technology Project of Guangxi (No.2024AA15007)

1 引言

随着城市化进程加速和机动车保有量持续攀升,城市交通系统供需失衡引发的道路拥堵已成为制约经济社会高质量发展的瓶颈问题^[1]. 交通流量预测作为路网状态感知与动态调控的核心技术手段,其精度直接影响信号灯配时优化、诱导路径规划等管控策略的有效性,进而决定了交通系统运行效率的提升空间. 传统的交通流预测方法主要基于物理模型与统计学方法,如时间序列分析、卡尔曼滤波等. 该类方法结构简单、计算效率高,且对数据量的要求相对较低;然而,由于交通系统具有较强的非线性和随机性,传统方法往往难以充分捕捉复杂交通动态中的内在规律,预测精度存在瓶颈^[2]. 近年来,深度学习模型^[3-5]在交通流量预测领域展现出显著优势,其通过多层非线性变换自动挖掘时空特征,突破了传统方法依赖人工特征工程和受限参数空间的桎梏. 以时空图神经网络为代表的模型,如文献[6]构建的长期显隐时空网络,通过耦合图卷积与循环单元,在捕捉路网拓扑约束下的局部/非局部空间相关性方面取得进展;文献[7]提出的时空门控注意力转换器则利用改进的自注意力机制增强时间特征的细粒度提取能力. 特别地,基于Transformer架构的预测方法^[8,9]通过位置编码与多头注意力机制,理论

上具备建模长程时空依赖的潜力. 然而,这些方法仍存在两大本质局限:其一,对交通流量非平稳性的适应性不足. 现有模型普遍假设时空依赖具有稳态特性,未能有效刻画交通系统因突发事件(如事故、天气突变)或周期性场景(早晚高峰、节假日)引发的数据分布偏移. 例如,文献[10]虽引入概率建模量化预测不确定性,但仍忽略了外部事件驱动的时变耦合强度对分布规律的动态重构机制;其二,对延迟时空依赖的建模存在系统性缺失. 如图1所示,交通流量的异常波动往往存显著的传播延迟效应—A节点的拥堵状态可能经过 Δt 时间间隔才引发B节点的流量激增. 现有方法,如文献[11]结合胶囊网络的多节点预测框架,尝试通过多视角关联增强空间建模能力,但其静态权重分配机制难以捕获这种具备延时特性的动态关联模式. 更关键的是,基于固定时间窗口的特征聚合操作^[8],无法区分不同延迟路径下的传播动力学差异,导致模型对交通流突变信号的响应存在系统性滞后. 这种缺陷在包含复杂交叉口、潮汐车道等异质化设施的实际路网中尤为突出,严重制约了预测模型在动态交通场景下的工程适用.

总体而言,现有研究在从静态时空建模向动态非平稳建模演进的过程中,仍面临两大核心挑战:一是对由突发事件、周期性规律等引起的数据分布偏移(非平稳性),缺乏显式有效的建模机制;二是对交通流在路

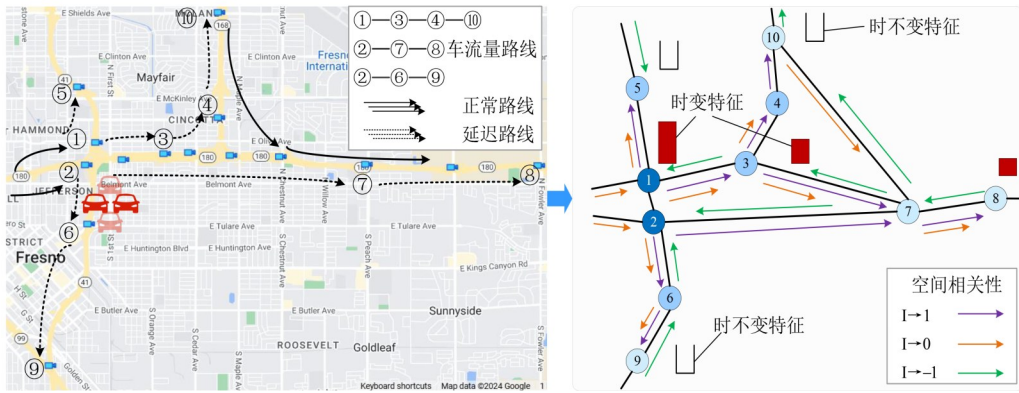


图1 交通流量延时传播

网中传播的延迟效应缺乏精细化的描述能力,多数模型仍依赖于同步或固定延迟的假设,难以适配真实的、异质化的交通动力学过程.

现有研究在应对交通流预测问题时,多基于平稳性假设构建时间序列模型,或采用图神经网络捕捉路段间的静态空间关联.然而,实际交通流具有显著的非平稳特性:一方面,早晚高峰、节假日等特殊时段产生的流量突增与周期性波动,导致时间维度呈现复杂多模态分布特征;另一方面,拥堵传播、车流扩散等微观交通动力学过程使得空间依赖性存在时延效应—A路段当前时刻的流量变化往往需经历数分钟至数十分钟的传递时间才会显著影响下游B路段.更为严峻的是,路网拓扑结构动态调整(如施工占道、临时交通管制)与突发交通事件(如交通事故、恶劣天气)的共同作用,加剧了流量变化的不平稳程度,如图2所示,使得时空依赖关系呈现非稳态演化规律.传统方法固化的时间窗口设定与静态邻接矩阵表示,难以适应此类延迟耦合、动态演化的时空交互机制,导致预测结果常出现时间滞后性误差与空间关联失准问题.

针对现有交通流量预测模型在非平稳性建模与延迟时空依赖特征捕获方面的局限性,本文提出一种基于延迟时空依赖的非平稳时间序列交通流量预测模型(Non-Stationary time series Forecasting Model, NSFM),主要贡献体现在以下3个方面.

(1)提出了一种基于空间自相关算子的非平稳时间序列交通流量预测模型 NSFM,该模型能够将交通流量数据分解为时变和时不变两类组分,从而能够在全局频率视角下描述交通流的稳态特性与局部动态趋势.

(2)在交通流量的时间特征和空间特征的基础上,特别引入了延时特征来捕捉交通流量的传播延迟特性,并通过逐点卷积和位置编码,将三者进行深度融合,从而更准确地捕捉交通流的动态传播规律.

(3)在PEMS03等四个不同城市公共交通流数据集上分别进行了实验验证,结果表明所提模型在四个数据集上的平均绝对百分比误差与SOTA模型相比,分别降低了7.48%、9.86%、3.20%、1.73%,证明了该模型在应对复杂非平稳场景下的交通流量预测具有一定的有效性和优越性.

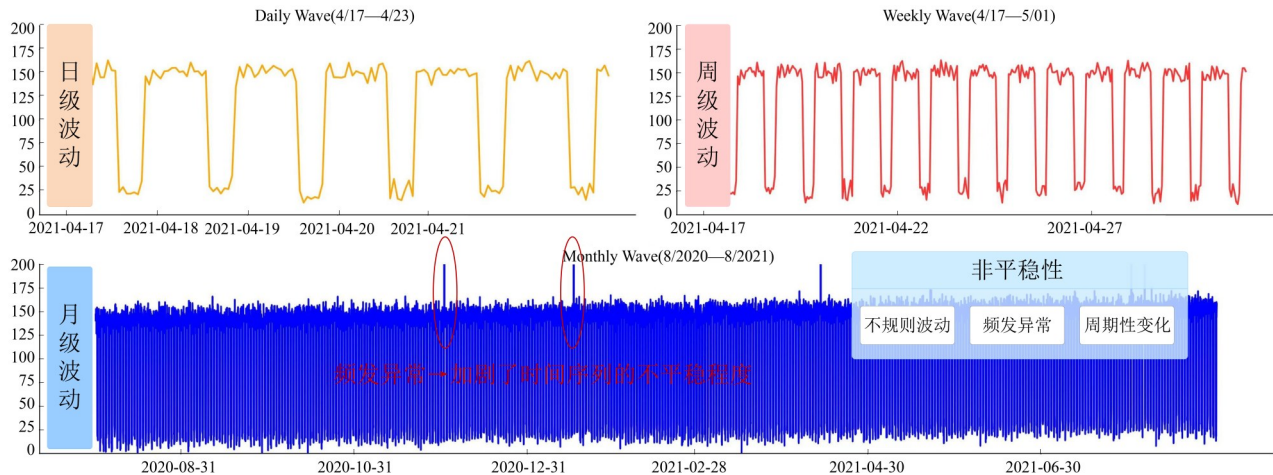


图2 交通流量时间序列非平稳性归因

2 相关工作

现有的交通流量预测主要可分为统计学模型和基于机器学习的模型,其中机器学习模型又可细分为循环神经网络模型、注意力机制增强模型和动态图神经网络模型等.本节系统梳理现有方法的行业背景、技术脉络,并分析其核心局限性.

2.1 基于统计学的交通流量预测模型

统计学模型,如向量自回归(Vector Auto Regression, VAR)^[12]、滑动平均模型(Moving Average, MA)^[13]及其组合模型自回归滑动平均模型(Auto Regressive Moving Average, ARMA)^[14]通过线性组合历史观测值建模时序依赖,但对非平稳交通流的非线性突变成分建模能力不足.学者提出统计非线性改进模型:非下采样Shearlet变换^[15]通过多尺度分解提取瞬态特征,双线性模型^[16]引入乘积项增强非线性表达能力.然而,这类方法仍依赖人工特征设计,难以应对高维时空交互场景.

2.2 基于机器学习的交通流量预测模型

机器学习模型通过核函数映射提升非线性建模能力,如支持向量回归(Support Vector Regression, SVR)^[17]利用结构风险最小化原则优化预测超平面.文献[18]进一步结合网格搜索优化SVR超参数,但其滑动时间窗口机制忽略长程依赖特征.文献[19]将主成分分析(Principal Component Analysis, PCA)与多元线性回归结合实现特征降维,但仅适用于单变量预测任务.

传统机器学习模型受限于浅层模型结构,无法有效捕捉动态时空耦合效应.近期针对交通流量预测的研究聚焦深度学习框架,重点突破时空依赖的动态建模能力,具体可分为以下三类:

(1) 循环神经网络模型

长短期记忆网络(Long Short-Term Memory, LSTM)是一种特殊的循环神经网络(Recurrent Neural Network, RNN),用于解决传统RNN在处理长序列数据时容易出现的梯度消失或梯度爆炸问题^[20].门控循环单元(Gated Recurrent Unit, GRU)^[21]是一种改进的循环神经网络,与LSTM类似,旨在解决传统RNN在处理长序列数据时的梯度消失或梯度爆炸问题,但结构更加简化.LSTM和GRU与图卷积网络(Graph Convolutional Network, GCN)的级联架构^[22],利用路网拓扑构建静态邻接矩阵.但GCN的固定聚合权重难以刻画交通状态的时变特性(如早晚高峰流量突变).文献[23]通过元学习生成动态邻接矩阵,但未显式建模延迟传播效应.

(2) 注意力机制增强模型

基于Transformer的模型利用自注意力捕捉全局时空依赖^[24,25].文献[9]集成多通道数据输入与时空变换器的先进架构,捕捉多维数据中的复杂时空特征与交

互关系,但其固定窗口的全局注意力机制导致事件响应滞后(平均延迟2.3个时间步).文献[26]提出因果稀疏注意力降低计算复杂度,但静态稀疏模式难以适应突发的流量波动.

(3) 动态图神经网络模型

文献[27,28]通过动态节点嵌入与微分方程更新图结构.文献[29]基于动态图框架实现大规模路网建模,但其空间权重分配未考虑交通拥堵的延时传播特性;文献[30]提出动态马尔可夫模型描述时序图演化,但马尔可夫假设限制了对长延迟链的建模能力.文献[31]开发自适应延迟图卷积层,可学习延时参数 Δt 调节消息传递速度,但采用均匀延迟分配策略,忽视异质节点对拥堵传播的差异化响应(如主干道与支路延时差异可达8~15 min)^[32].

2.3 现有交通流量预测模型局限性总结

当前研究存在三方面不足.第一,非平稳性建模缺失.现有方法假设交通流服从平稳或分段平稳分布^[33],对突发事件(如交通事故)引发的分布偏移缺乏鲁棒适应性.第二,延迟依赖建模粗粒度.时空依赖建模普遍采用同步聚合策略^[34],忽视交通拥堵传播的延时效应.第三,动态拓扑更新低效性.图结构更新依赖节点属性相似性度量^[35,36],未建立传播延迟与网络拓扑的显式耦合关系.上述问题导致现有模型在城市复杂路网中的预测误差显著增加,突发拥堵场景下,主流模型如Graph Wave Net^[37]、STG2Seq(Spatial-Temporal Graph to Sequence)^[38],最新模型CCHMM(Causal Conditional Hidden Markov Model)^[39]、Cycle LLH(Cycle Little Linear Head)^[4]的平均绝对误差(MAE)较平稳场景显著提升,进一步说明了现有方法在应对非平稳性和突发交通事件时的局限性.

3 非平稳时间序列交通流量预测模型

本节对交通路口流量预测问题进行了定义,并详细阐述了所提的基于延迟时空依赖的非平稳时间序列流量预测模型的整体框架和各功能模块.

3.1 问题定义

交通流量预测旨在通过历史交通数据,预测未来某一时段内特定路段的交通流量数据来分析交通状态.本文将整个城市沿着经纬度均匀等分成 N 个网格,每个网格代表一个区域.采集数据的时间段定义为时间间隔.以 t 为时间间隔将历史时间划分为 T 个时间段,则第 t 个时间间隔内通过第 n 个区域的车辆流量可以表示为 X_t^n .使用 $G=(V,I,X)$ 将所有区域合并成数据结构.其中 V 是网格节点集合,每个节点表示一个网格区域 $V=\{v_1, v_2, \dots, v_N\}$; I 为Moran算子,对应区域之间的空间相关性; X 为各节点输入数据的集合,

$\mathbf{X} \in \mathbb{R}^{T \times N \times C}$, \mathbb{R} 表示实数集, C 为模型输入数据的维数. 将交通流量预测问题形式化表述为寻找一个最优映射函数 J , 该函数基于区域信息 G , 利用交通流量时间序列的历史步长来预测交通流量的未来步长. 上述映射关系可表示为

$$[Y_{t+1}, Y_{t+2}, \dots, Y_{t+S}] = f([X_{t-T+1}, X_{t-T+2}, \dots, X_t]) \quad (1)$$

$$\mathbf{X}_T = [X_{t-T+1} \quad X_{t-T+2} \quad \dots \quad X_t] \in \mathbb{R}^{T \times N \times C} \quad (2)$$

$$\mathbf{Y}_S = [Y_{t+1} \quad Y_{t+2} \quad \dots \quad Y_{t+S}] \in \mathbb{R}^{S \times N \times C} \quad (3)$$

式(2)和式(3)分别为描述历史和未来流量数据的公式.

3.2 模型架构与实现

本文针对不平稳变化的交通流量数据集提出了基

于延迟时空依赖的非平稳时间序列预测模型, 该模型主要由 3 个部分组成, 如图 3 所示.

(1) 时序分解模块. 为捕捉非平稳序列中的时变和时不变分量, 将目标区域的历史观测数据进行时序分解, 构成不同的特征向量, 作为子网络层的输入.

(2) 特征提取模块. 基于时序分解模块的输出, 提取 T 个历史时刻的交通流量、空间特征和时间特征向量, 分别为 $\mathbf{E}_p \in \mathbb{R}^{N \times D}$, $\mathbf{E}_T \in \mathbb{R}^{N \times D}$ 和 $\mathbf{E}_S \in \mathbb{R}^{N \times D}$, 融合卷积将这些表示集成为统一的时空特征 $\mathbf{H} \in \mathbb{R}^{N \times 3D}$.

(3) 预测模块. 将全局特征输入全连接网络层获得预测结果. 最后变换函数对函数回归方程调整模型中的参数, 使得预测结果能够自适应数据特征.

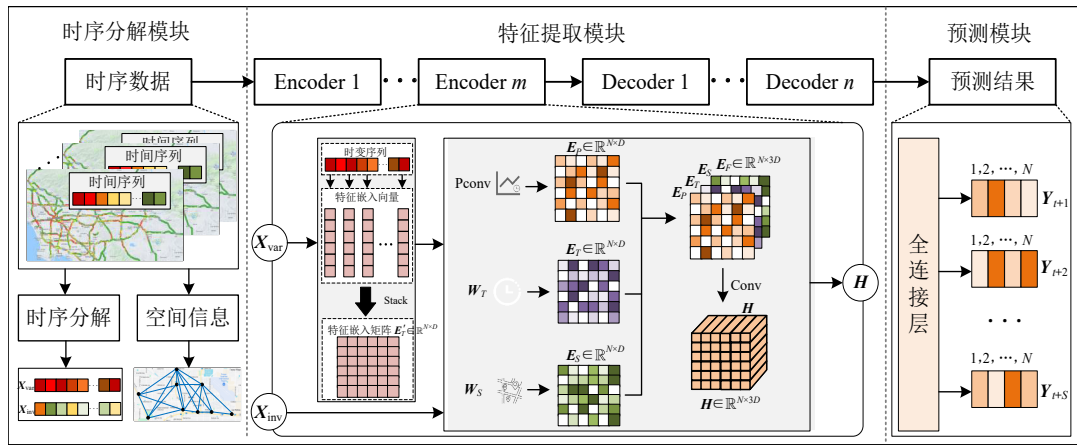


图 3 模型框架图

3.2.1 非平稳时间序列分解

对交通流量数据的非平稳特性, 区分不同的动力学因素并采用合适的算子. 基于傅里叶变换, 利用频域统计来分离确定性分量和随机分量 (即时不变和时变分量), 获得特征提取模块的输入. 在预测模块过程中, 将算子设置为预测窗口全局学习的模型参数.

傅里叶滤波器将非平稳时间序列集合分解:

$$\mathbf{X}_{\text{var}}, \mathbf{X}_{\text{inv}} = \mathcal{F}(\mathbf{X}) \quad (4)$$

其中, 下标 inv/var 分别代表时不变/时变组分. 时序分析的两个维度一般包含时域与频域, 其中, 频率所描述的正是一个全局概念, 利用傅里叶变换, 可以获得序列各种频率的强度 (频谱), 进而提取数据集的时不变特性. 具体而言, 对每个序列进行傅里叶变换 \mathcal{F} , 计算训练数据的频谱强度分布, 强度高的频率即为数据集 (整段时期) 的主要频率, 记强度排名前 α 的频率集合为 G_α . 每次输入待预测序列时, 利用傅里叶变换将频率在的组分 G_α 过滤为时不变组分, 而剩余的组分就是每个序列所处时期的独特成分, 即为时变组分. 计算过程如下:

$$\mathbf{X}_{\text{inv}} = \mathcal{F}^{-1}(\text{Filter}(G_\alpha, \mathcal{F}(\mathbf{X}))) \quad (5)$$

$$\mathbf{X}_{\text{var}} = \mathcal{F}^{-1}(\text{Filter}(\bar{G}_\alpha, \mathcal{F}(\mathbf{X}))) = \mathbf{X} - \mathbf{X}_{\text{inv}}$$

下面我们给出该分解为正交分解的证明:

定理 1 给定非平稳交通流量时间序列 \mathbf{X} , 其傅里叶变换为 $\mathcal{F}(\mathbf{X})$. 设 G_α 为根据训练集频谱强度选取的前 α 个频率构成的集合, \bar{G}_α 为其补集. 通过式(5)定义的滤波器, 将其分解为时不变组分和时变组分. 那么, 在 L^2 空间内积意义下, 其和是正交的.

$$\langle \mathbf{X}_{\text{inv}}, \mathbf{X}_{\text{var}} \rangle = 0 \quad (6)$$

证明 傅里叶变换的基函数系 $\{e^{i2\pi f t}\}$ 构成一组标准正交基:

$$\langle e^{i2\pi f_1 t}, e^{i2\pi f_2 t} \rangle = \begin{cases} 1, & f_1 = f_2 \\ 0, & f_1 \neq f_2 \end{cases} \quad (7)$$

序列 \mathbf{X} 可表示为所有频率组分的线性叠加: $\mathcal{F}(\mathbf{X}(t)) = \sum_f c_f e^{i2\pi f t}$ 根据定义, 时不变组分 $\mathbf{X}_{\text{inv}}(t) = \sum_{f \in G_\alpha} c_f e^{i2\pi f t}$, 时变组分 $\mathbf{X}_{\text{var}}(t) = \sum_{f \in \bar{G}_\alpha} c_f e^{i2\pi f t}$. 由于集合 G_α 与 \bar{G}_α 不相交, 故 \mathbf{X}_{inv} 和 \mathbf{X}_{var} 不包含任何共同的频率分

量. 因此, 它们的内积为

$$\begin{aligned} \langle \mathbf{X}_{\text{inv}}, \mathbf{X}_{\text{var}} \rangle &= \left\langle \sum_{f_p \in G_a} c_{f_p} e^{i2\pi f_p t}, \sum_{f_q \in G_a} c_{f_q} e^{i2\pi f_q t} \right\rangle \\ &= \sum_{f_p \in G_a} \sum_{f_q \in G_a} c_{f_p} \bar{c}_{f_q} \langle e^{i2\pi f_p t}, e^{i2\pi f_q t} \rangle = 0 \end{aligned} \quad (8)$$

证毕.

该分解的正交性保证了分解出的时不变组分 \mathbf{X}_{inv} 和时变组分 \mathbf{X}_{var} 在数学上相互独立, 这种独立性为后续差异化建模奠定了基础——时不变组分能纯净地刻画全局稳态趋势, 而时变组分能无混杂地捕捉局部动态与延迟传播特征.

3.2.2 带延时特性的时空特征提取

特征提取模块通过关注数据的时空邻近特性, 深入挖掘流量数据的动态演变规律, 提高预测准确度. 将每个位置的交通流量数据细化为特征向量, 包括空间相关性、一天中的小时、一周中的日期等特征表示. 通过逐点卷积来嵌入数据, 其中输入数据 \mathbf{X} 被转换为向量 $\mathbf{E}_p \in \mathbb{R}^{N \times D}$:

$$\mathbf{E}_p = \text{PConv}(\mathbf{X}; \theta_p) \quad (9)$$

其中, PConv 表示使用具有 1×1 核大小的滤波器进行逐点卷积操作; D 为隐藏层维度; θ_p 表示逐点卷积的可学习参数. 为保留数据中的时间信息, 本文为每个交通数据执行绝对位置编码, 分别在“日”和“周”分辨率上进行, 生成的位置编码为 $\mathbf{X}_{\text{day}} \in \mathbb{R}^{N \times T_d}$ 和 $\mathbf{X}_{\text{week}} \in \mathbb{R}^{N \times T_w}$. 一天中的小时数据 $\mathbf{E}_T^d \in \mathbb{R}^{N \times D}$ 和一周中的日期数据 $\mathbf{E}_T^w \in \mathbb{R}^{N \times D}$ 计算如下:

$$\mathbf{E}_T^d = \mathbf{W}_{\text{day}}(\mathbf{X}_{\text{day}}), \quad \mathbf{E}_T^w = \mathbf{W}_{\text{week}}(\mathbf{X}_{\text{week}}) \quad (10)$$

其中, $\mathbf{W}_{\text{day}} \in \mathbb{R}^{T_d \times D}$ 和 $\mathbf{W}_{\text{week}} \in \mathbb{R}^{T_w \times D}$ 分别是一天中的小时和一周中的日期的可学习参数, 通过将这两者相加, 得到时间表示 $\mathbf{E}_T^s \in \mathbb{R}^{N \times D}$.

交通流量在路网中的传播具有延时特性, 以交通事故为例, 其对相邻区域交通状态的影响通常存在数分钟的延迟. 为准确表征这一时变特征, 本文基于时变组分数据提取延时特征, 以捕捉站点交通流量的延时传播规律, 从时变组分数据集中提取单站点交通流数据, 使用集合 $\mathbf{X}_v = \{\mathbf{X}_{\text{var}, i} | i \in [1, 2, \dots, N]\}$ 表示. 相似的交通流量数据表明站点之间可能存在相似的交通状况, 因此将每个站点的历史时变组分序列与提取的流量模式集 \mathbf{X}_v 进行比较, 将相似模式的信息融合到每个站点的历史流量序列表示中.

首先将单个站点的 \mathbf{X}_{var} 通过嵌入矩阵 \mathbf{W}^e 得到高维形式 $\mathbf{e} = \mathbf{X}_{\text{var}} \mathbf{W}^e$, 同时通过另一个嵌入矩阵将每个站点的时变组分序列转换为记忆向量 $\mathbf{m} = \mathbf{X}_{\text{var}, i} \mathbf{W}^m$, 将高维形式的单个站点时变组分序列与记忆向量比较得到相

似性向量 $\mathbf{s} = \text{softmax}(\mathbf{e}^T \mathbf{m})$. 根据相似性向量 \mathbf{s} 对集合 \mathbf{X}_v 加权求和从而得到时间延迟的时变组分序列 $\mathbf{v} = \sum \mathbf{s}(\mathbf{X}_{\text{var}, i} \mathbf{W}^c)$, 其中 \mathbf{W}^c 为可学习的参数矩阵. 经过完整的计算从而得到单个站点的时变组分序列, 通过累计得到 N 个站点的综合时变组分序列 \mathbf{E}_v , 通过 \mathbf{E}_v 的补充得到最终的时间特征 \mathbf{E}_T :

$$\mathbf{E}_T = \mathbf{E}_T^s + \mathbf{E}_v \quad (11)$$

为表示区域内离散站点的空间相关性, 设计空间表示 $\mathbf{E}_S \in \mathbb{R}^{N \times D}$:

$$\mathbf{E}_S = \sigma(\mathbf{W}_S \cdot \mathbf{X} + \mathbf{b}_S) \quad (12)$$

其中, σ 表示激活函数; $\mathbf{W}_S \in \mathbb{R}^{D \times D}$ 和 $\mathbf{b}_S \in \mathbb{R}^D$ 是可学习的参数. 最后引入融合卷积 (FConv), 整合交通流量、空间和时间特征向量, 以表示 T 个历史时刻的延迟时空特征:

$$\mathbf{H} = \text{FConv}(\mathbf{E}_p \| \mathbf{E}_S \| \mathbf{E}_T; \theta_f) \quad (13)$$

其中, $\mathbf{H} \in \mathbb{R}^{N \times 3D}$; $\|$ 表示连接操作; θ_f 代表 FConv 可学习参数.

3.2.3 基于 Moran 算子的回归预测

预测模块使用全连接层 (FC) 来预测 S 个时间步长的交通流量:

$$\mathbf{Y}_S = \text{FC}(\mathbf{H}; \theta_s) \quad (14)$$

其中, $\mathbf{Y}_S \in \mathbb{R}^{S \times N \times C}$; θ_s 代表 FC 的可学习参数.

考虑数据之间潜在的空间相关性, 评估不同区域站点之间的空间特征依赖关系, 构造 Moran 算子 I , 描述地理空间中特征值的全局自相关性. 其中:

$$I = \frac{N}{\mathbf{W}} \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (15)$$

其中, x_i 和 x_j 是特征值; \bar{x} 是特征值的均值; w_{ij} 是 i 和 j 之间的空间权重; \mathbf{W} 是所有空间权重的总和; I 是 Moran's I 统计量, 它衡量了特征值的全局空间自相关性. 如果 I 的值接近于 1, 则表明特征值在空间上具有正的自相关性 (即聚集), 如果 I 的值接近于 -1, 则表明特征值在空间上具有负的自相关性 (即分散), 如果 I 的值接近于 0, 则表明特征值在空间上具有随机分布.

为将 Moran 算子融入函数对函数回归框架, 并解决连续空间 (基函数展开) 与离散空间 (站点观测) 之间的映射一致性问题, 我们引入基函数展开与正交化. 考虑每个站点的交通流量观测值为时间函数 $\mathbf{Y}_{s_i}(t), \mathbf{X}_{s_i}(r)$, 其中 $t, r \in \Omega$ 为连续时间域. 通过 QR 分解方法对基函数进行正交化, 将各函数变量展开为关于正交基函数 $\{\phi_k\}_{k \in N} \in \Gamma(\Omega)$ 和 $\{\phi_g\}_{g \in N} \in \Gamma(v)$ 的线性组合, 降低数据变量之间的共线性:

$$\begin{aligned}
 Y_{s_i}(t) &= \sum_{k=1}^{\infty} y_{ik} \phi_k(t) \\
 X_{s_i}(r) &= \sum_{g=1}^{\infty} x_{ig} \phi_g(r) \\
 W_{s_i}(t) &= \sum_{k=1}^{\infty} w_{ik} \phi_k(t) \\
 e_{s_i}(t) &= \sum_{k=1}^{\infty} e_{ik} \phi_k(t)
 \end{aligned} \tag{16}$$

由于 t, s_i 分别存在曲线内和曲线间依赖关系, 还需要对协方差函数进行建模, 增加计算复杂度的同时还会产生复杂的似然函数. 因此, 本文采用正交基函数的基变换方法, 将数据空间转换为基函数空间, 在基空间中使用较简单的协方差来拟合模型. 即对正交基函数进行内积运算, 得到

$$\begin{aligned}
 \int_{\Omega} \phi_k(t) Y_{s_i}(t) dt &= \int_{\Omega} \phi_k(t) \int_{\psi} \psi(r, t) X_{s_i}(r) dr dt \\
 &+ \int_{\Omega} \phi_k(t) W_{s_i}(t) dt + \int_{\Omega} \phi_k(t) e_{s_i}(t) dt
 \end{aligned} \tag{17}$$

其中, $\{\phi_k(t)\}_{k \in N}$ 是一项标准正交基函数, 根据支配收敛性, $\int_{\Omega} \phi_k(t) Y_{s_i}(t) dt$ 可以变换为

$$\begin{aligned}
 \int_{\Omega} \phi_k(t) Y_{s_i}(t) dt &= \int_{\Omega} \phi_k(t) \left(\sum_{k=1}^{\infty} y_{ik} \phi_k(t) \right) dt \\
 &= \sum_{k=1}^{\infty} y_{ik} \left(\int_{\Omega} \phi_k(t) \phi_k(t) dt \right) = \sum_{k=1}^{\infty} y_{ik}
 \end{aligned} \tag{18}$$

类似地, 得到等式右边每一项. 因此, 用有限数量的基函数 K 和 G 来近似上述的无限基函数展开式, 其中 k_n 和 g_n 随着 $n \rightarrow \infty$ 而增大. 设 $\tilde{Y} = (y_{ik})_{N \times K}$ 为目标函数矩阵, $\tilde{X} = (x_{ig})_{N \times G}$ 为设计矩阵, 其系数为 $\tilde{\psi} = (\psi_{gk})_{G \times K}$, $\tilde{W} = (w_{ik})_{N \times K}$ 为变换后的随机效应矩阵, $\tilde{e} = (e_{ik})_{N \times K}$ 为变换后的误差矩阵, 得到

$$\tilde{Y} = \tilde{X} \tilde{\psi} + \tilde{W} + \tilde{e} \tag{19}$$

由于观测数据来源于离散站点, 需将连续函数模型映射至离散空间. 设观测站点位置为 t_1, t_2, \dots, t_N , 定义投影算子 $P: L^2(\Omega) \rightarrow \mathbb{R}^N$, 将函数映射为站点观测值:

$$P: Y_{s_i}(t) \mapsto y_{s_i} = [Y_{s_i}(t_1), Y_{s_i}(t_2), \dots, Y_{s_i}(t_N)]^T \tag{20}$$

类似地, 对基函数也进行离散采样, 构建基函数矩阵 $\Phi = [\phi_1 \quad \phi_2 \quad \dots \quad \phi_K] \in \mathbb{R}^{N \times K}$, 其中, $\phi_k = [\phi_k(t_1), \phi_k(t_2), \dots, \phi_k(t_N)]^T$. 为实现空间自相关建模, 将 Moran 算子引入随机效应项矩阵 \tilde{W} . 设空间权重矩阵 $W = (w_{ik})_{N \times N}$, Moran's I 可写为

$$I = \frac{z^T W z}{z^T z} \tag{21}$$

其中, z 为去均值后的特征向量. 为将 Moran 算子融入函数回归框架, 假设随机效 \tilde{W} 具有空间结构, 即

$$\tilde{W} = P \delta + \eta \tag{22}$$

其中, $P \in \mathbb{R}^{N \times p}$ 为投影矩阵, 由 Moran 算子的前 p 个特征向量构成, $\delta \in \mathbb{R}^{p \times K}$ 为低维随机系数, η 为残差.

为证明连续基函数空间与离散 Moran 空间之间映射的一致性, 定义分段线性基函数 $A \in \mathbb{R}^{N \times M}$. 矩阵 A 的行对应观测位置, 列对应网格顶点; A 的第 i 行包含要插值到目标观测点位置的权重. 则观测函数 y_{s_i} 可表示为

$$y_{s_i} = A y_g \tag{23}$$

其中, y_g 为网格顶点上的函数值. 进一步, 将 Moran 算子的特征向量矩阵 $M \in \mathbb{R}^{M \times p}$ 作用于网格顶点, 可得投影矩阵:

$$P = A M \in \mathbb{R}^{N \times p} \tag{24}$$

该矩阵实现了从连续网格空间到离散站点空间的映射. 将式(24)代入式(19), 得到

$$\tilde{Y} = \tilde{X} \tilde{\psi} + P \delta + \tilde{e} \tag{25}$$

当基函数数量足够多、网格分辨率足够高的条件下, 通过插值矩阵 A 与 Moran 特征向量矩阵 M 构建的投影矩阵 P , 能够实现从连续函数空间到离散站点空间的一致性映射, 从而保证了式(25)的近似合理性.

非平稳时间序列流量预测方法流程图如图 4 所示, 直观展示了提出的 NSFM 模型主要结构: 频率解耦、延迟依赖建模、函数回归预测.

4 实验与分析

4.1 实验配置

4.1.1 数据集

本文基于真实交通路口的流量数据集对所提出模型的性能进行了验证. 所用数据为加利福尼亚州交通部门发布的公共交通流量数据集 PEMS^[39], 记录了四个城市区域每 5 min 一次的交通流量信息. 数据集基本情况如表 1 所示. 各子数据集代表了不同区域的交通特征, 具体如下. PEMS03(旧金山湾区): 地形复杂且人口密集, 交通流量呈现高频波动; PEMS04(洛杉矶地区): 高速公路网络发达但事故频发, 导致瞬时流量易出现异常; PEMS07(圣地亚哥及周边): 多中心城市结构使得流量波动可能在多个局部节点同时发生; PEMS08(圣何塞与硅谷地区): 以通勤为主, 流量周期性显著, 工作日与周末差异明显. 通过涵盖不同城市结构与交通行为的多样性, 这些数据集为模型的泛化能力评估提供了良好基础.

4.1.2 评价指标

本文采用过去的 12 个时间点 (60 min) 来预测未来的 12 个时间点. 使用三个统计指标来评估模型的性能: 平均绝对误差 (MAE)、均方根误差 (RMSE)、平均绝对百分比误差 (MAPE). 其中, RMSE、MAE、MAPE 越

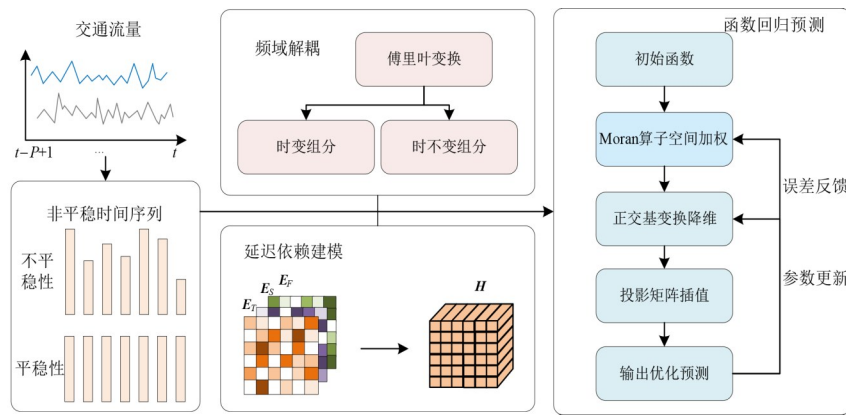


图4 方法流程图

表1 数据集描述

数据集	传感器数量	线路	时间步长	时间范围
PEMS03	358	866	26 208	2018.09.01—2018.11.30
PEMS04	307	680	16 992	2018.01.01—2018.02.28
PEMS07	883	340	28 224	2017.05.01—2017.08.31
PEMS08	170	548	17 856	2016.07.01—2016.08.31

小,说明预测误差越小,即预测效果越好. 设 $y_n (1 \leq n \leq N)$ 和 $\hat{y}_n (1 \leq n \leq N)$ 分别表示第 n 个真实值和预测值, n 为测试集中样本的个数. 公式定义如下:

$$\text{MAE} = \frac{1}{N} \sum_{n=1}^N |y_n - \hat{y}_n| \quad (26)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - \hat{y}_n)^2} \quad (27)$$

$$\text{MAPE} = \frac{1}{N} \sum_{n=1}^N \left| \frac{y_n - \hat{y}_n}{y_n} \right| \times 100\% \quad (28)$$

4.1.3 模型配置

本文非平稳时间序列预测模型的实现步骤与参数设置如下. 实现步骤:(1)在数据信息提取部分,使用一个卷积核为 1×1 的二维卷积层进行逐点卷积操作;(2)在时间特征提取部分,使用1个线性层提取数据“每日小时”的特征,使用1个线性层提取数据“每周日期”的特征,使用1个线性层提取数据“延时传播”的特征,将三者相加获得时间特征;(3)在空间特征提取部分,使用一个线性层、一个激活函数;(4)在特征整合部分,使用一个二维卷积层作为融合卷积将所有特征整合在一起;(5)在预测部分,使用全连接网络层实现预测. 参数设置:将PEMS数据集按6:2:2的比例划分为训练集、验证集和测试集. 基于深度学习库构建和训练模型,使用Ranger21优化器优化参数,学习率设置为0.001,批量大小设置为64,使用MAE损失训练模型. 具体的模型参数配置如表2所示.

NSFM模型的时间复杂度可以由下式计算: $O(B \times$

表2 模型参数

模型参数	值
学习率	1×10^{-3}
批次大小	64
每批次数据量	100
延迟时间	0.001
隐藏层数	3
隐藏层维数	512
输入维数	5(PEMS04,PEMS08),3(PEMS03,PEMS07)
延迟时间	12

$N \times (T \times F + N \times D + D \times \text{output_len})$ 其中, B 为批次大小, N 为节点数量, T 为时间序列长度, F 为输入特征维度, D 为隐藏层维度. 最终模型训练得到的平均单次训练时长约为150 s.

4.2 实验结果与分析

为证明模型的有效性和准确性,本文将NSFM模型与以下10种现有预测方法进行对比,包含3种传统时间序列预测方法以及7种深度学习预测方法:VAR^[12]、SVR^[17]、LSTM^[22]、Graph WaveNet^[37]、STG2Seq^[38]、STTN (Spatial-Temporal Transformer Networks)^[24]、DGCN (Dynamic Graph Convolutional Recurrent Network)^[29]、CCHMM^[32]、MC-STTM^[9]、Cycle LLH^[4].

为保证比较实验的公平性,上述所有模型使用相同的数据集来进行训练、验证和测试. 所有模型的输入数据相同,仅根据不同模型所需的数据格式进行了相应的调整. 图5展示了所提的NSFM模型分别在如表1所示的4个数据集上训练的收敛结果,其中蓝色实线为训练损失(Training Loss),红色虚线为验证损失(Validation Loss). 可以看出,所提模型在前几个epoch内迅速下降,且没有较大波动,表明模型能迅速适应数据的非平稳性质. 并且随着Epoch的增加,训练损失和验证损失都趋于稳定,且两者之间的差距较小,体现了模型具有良好的泛化能力.

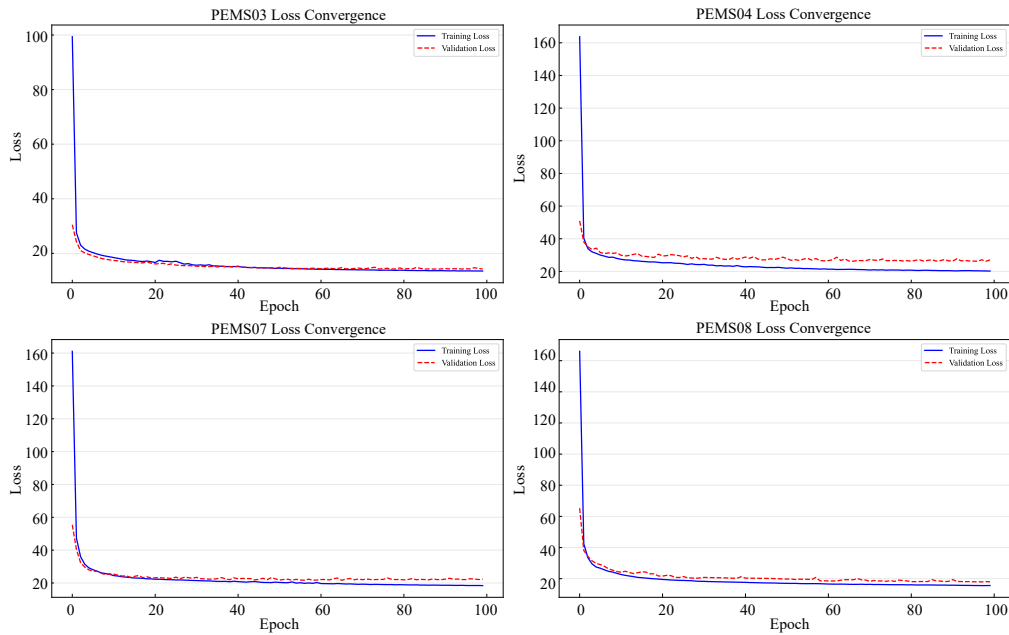


图5 NSFM模型在PEMS03、PEMS04、PEMS07、PEMS08数据集上的训练和验证损失下降曲线

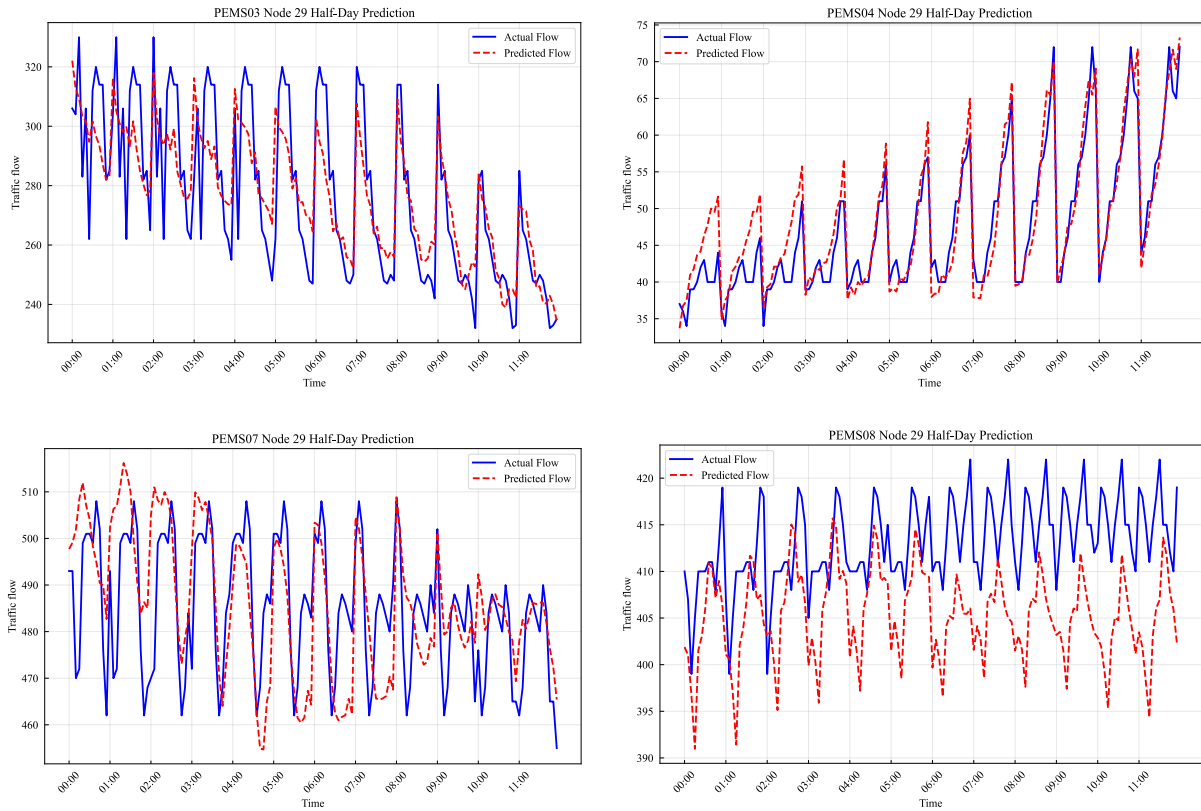


图6 NSFM在PEMS03、PEMS04、PEMS07、PEMS08数据集上的交通流预测结果

将训练好的NSFM模型在4个数据集上进行了预测结果测试,如图6所示,其中蓝色实线代表实际交通流量(Actual Flow),红色虚线代表预测交通流量(Pre-

dicted Flow)。可以看出,预测的交通流量曲线与实际流量曲线的趋势基本一致,不存在明显的时间滞后,表明模型能够较好地捕捉交通流量的非平稳性特征和延时

变化趋势. 表3和表4展示了所提的NSFM模型与前述10种模型的对比, 其中的数据是在同一测试集上进行10次实验的平均结果. 观察表3和表4可知3种评价指

标(MAE、RMSE、MAPE)在所有模型上基本呈现出一致性, 且本文提出的NSFM模型在3种测量条件下都取得了最佳的预测结果.

表3 不同模型在PEMS03和PEMS04数据集上的比较

模型	PEMS03			PEMS04		
	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
VAR(2015)	23.65	38.26	24.51	23.75	36.66	18.09
SVR(2020, Neurocomputing)	21.97	35.31	21.97	28.71	44.57	19.21
LSTM(2023, EAAI)	21.57	35.61	27.56	27.34	41.80	18.60
Graph WaveNet(2019, IJCAI)	19.88	33.12	19.90	25.48	39.74	17.53
STG2Seq(2019, IJCAI)	19.54	30.25	21.55	25.65	38.98	19.62
STTN(2023, Inf. Sci.)	24.39	35.94	23.01	26.80	39.98	20.39
DGCRN (2023, KDD)	19.54	30.97	18.79	24.25	36.86	17.20
CCHMM(2023, AAAI)	23.10	36.61	13.14	25.71	38.29	11.52
MC-STTM(2024, Inf. Sci.)	22.30	33.26	17.88	25.63	37.59	20.90
Cycle LLH(2024, 计算机学报)	19.71	31.39	11.21	24.72	36.66	11.08
NSFM	18.28	29.33	10.40	24.64	36.45	11.04

表4 不同模型在PEMS07和PEMS08数据集上的比较

模型	PEMS07			PEMS08		
	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
VAR(2015)	75.63	115.24	32.22	23.46	36.33	15.42
SVR(2020, Neurocomputing)	32.49	50.23	14.29	23.26	36.18	14.75
LSTM(2023, EAAI)	30.40	46.41	13.73	22.38	34.38	14.69
Graph WaveNet(2019, IJCAI)	26.90	42.85	12.53	21.21	31.12	12.68
STG2Seq(2019, IJCAI)	35.98	50.82	24.52	21.66	31.32	16.32
STTN(2023, Inf. Sci.)	32.24	46.14	14.65	22.60	33.47	12.92
DGCRN (2023, KDD)	28.73	45.73	13.27	19.70	30.13	11.70
CCHMM(2023, AAAI)	33.01	54.61	10.55	24.51	37.31	10.51
MC-STTM(2024, Inf. Sci.)	28.18	42.07	12.31	20.11	25.98	10.04
Cycle LLH(2024, 计算机学报)	32.96	50.29	10.54	21.70	31.24	9.31
NSFM	28.50	43.89	9.11	19.38	28.01	8.31

从结果中可以看出: 传统模型, 如VAR和SVR, 由于其算法原理的局限性, 无法有效捕获区域间的空间相关性, 在实际预测时难以准确把握流量在不同区域之间的关联变化, 导致预测效果并不理想, 尤其是在PEMS07数据集上, VAR的MAE和RMSE分别高达75.63和115.24, 表明传统模型在处理复杂时空依赖性时存在较大局限性. 传统神经网络LSTM网络虽在处理时间序列数据上有一定优势, 但其仅着重考虑时间依赖性的影响, 而忽视了空间维度的信息, 使得在面对复杂的交通流量数据时, 产生了较大的误差.

相比之下深度学习模型如Graph WaveNet和STG2Seq两者考虑时空相关性的网络模型, 打破了传统模型的局限, 在预测性能上有了显著提升, 相较于仅考虑时间因素的传统神经网络方法LSTM, 它们能更好地

捕捉时空依赖关系, 从而有更好的预测效果. 最新模型如MC-STTM和Cycle LLH在大多数数据集上表现更佳, 尤其是在PEMS07和PEMS08数据集上, 表明最新模型在处理非平稳时间序列和复杂时空依赖性方面都有了显著提升.

本文提出的NSFM模型, 在综合考虑时空相关性的基础上, 还充分考虑了异常事件的延时影响. 从表3和表4中可以看出, NSFM模型在4个数据集上都展现了显著的性能优势, 尤其是在PEMS03和PEMS04数据集上, 其MAE、RMSE和MAPE指标均优于其他对比模型. 如表2所示, 在PEMS03数据集上, NSFM的MAE、RMSE和MAPE分别为18.28、29.33和10.4%, 显著低于其他模型, 如利用图结构的Graph WaveNet和DGCRN模型, 显示出NSFM在捕捉局部和全局时空特征方面的优势.

在 PEMS04 数据集上, NSFM 三种评价指标均优于最新模型 MC-STTM 和 Cycle LLH, 体现了 NSFM 在解决复杂交通流量预测问题上的有效性和先进性. PEMS07 数据集路网节点数最多, 如表 3 所示, NSFM 的 MAPE 指标表现最优, 进一步验证了其在处理非平稳性和复杂时空依赖关系方面的有效性. 在 PEMS08 数据集上, NSFM 的 RMSE 指标相较于基于 Transformer 的架构, 如 STTN 和 Cycle LLH, 分别降低了 5.46、3.23, 体现了其在捕捉交

通流时空特征方面的显著优势.

4.3 步长敏感性分析

为验证 NSFM 模型对交通流延迟传播效应的捕捉能力, 本小节设计了延迟敏感性实验. 我们选取 PEMS04 数据集(该数据集包含频繁的交通事件, 延迟效应显著), 在保持其他参数不变的情况下, 系统调整模型中的延迟步长参数 T , 分别选择了 6、8、10、12 作为实验对比, 其预测结果如图 7 所示, 误差热力图如图 8 所示.

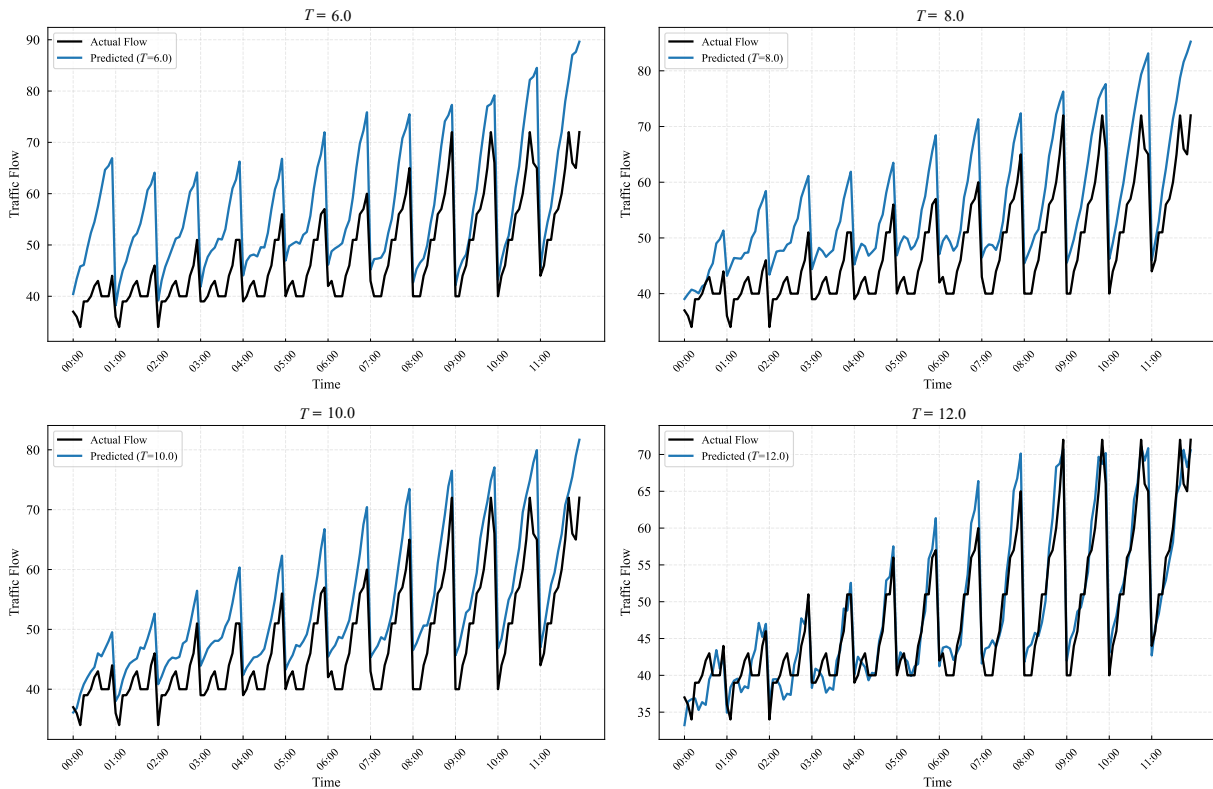


图 7 同一数据集不同时间步长结果预测对比图

由图 7 可知, 当 T 取较小时, 模型无法充分捕捉拥堵传播的完整过程, 产生较大的预测误差; 当 T 的取值较大, 如 12 时, NSFM 模型能够有效捕捉交通流的延迟传播特性, 取得较小的预测误差. 从图 8 也可以明显看出, 其误差热力图呈现明显的上浅下深的分布. 因此, 当 T 取 12 时, NSFM 模型获得较好的预测性能.

4.4 频带选择分析

为确定傅里叶分解中时不变组分的最优频带范围(即 K 值选择), 本小节设计了频带选择实验. 我们采用累积能量百分比方法确定 K 值: 对训练集所有序列的频谱强度进行归一化, 计算前 K 个频率的累积能量占比, 当累积能量达到预设阈值时确定 K 值. 我们分别测试了 10%、20%、30%、40% 四种 K 值下 NSFM 模型在 PEMS04 数据集上的预测结果, 如图 9 所示, 其误差热力图如图 10 所示.

从图 9、10 的结果可以看出, 若 K 值过小, 时不变组分包含的频带过窄, 将丢失部分重要的稳态交通模式, 导致模型对长期趋势的捕捉能力不足; 若 K 值过大, 时不变组分过于宽泛, 会压缩时变组分的表征空间, 削弱模型对动态波动特征的提取能力. 因此本文中 K 取 20% 是较为理想的最优值, 取得了相对较低的预测误差.

4.5 消融研究

为了验证本文提出的不同模块能否提高模型在交通流量数据预测上的表现, 本文设置了 NSFM 模型的三个变体. 在四个不同的 PEMS 数据集上进行了消融实验. 表 5 为不同变体模型在三个评价指标下的平均结果. w/o TDM 表示移除时序分解模块, 不考虑延时时间关系, 同时保持其他部分不变; w/o Moran 表示不插值 Moran 算子评估站点间的空间相关性来优化预测; w/o DSTFE 表示移除延时时空特征提取部分, 仅用全连接

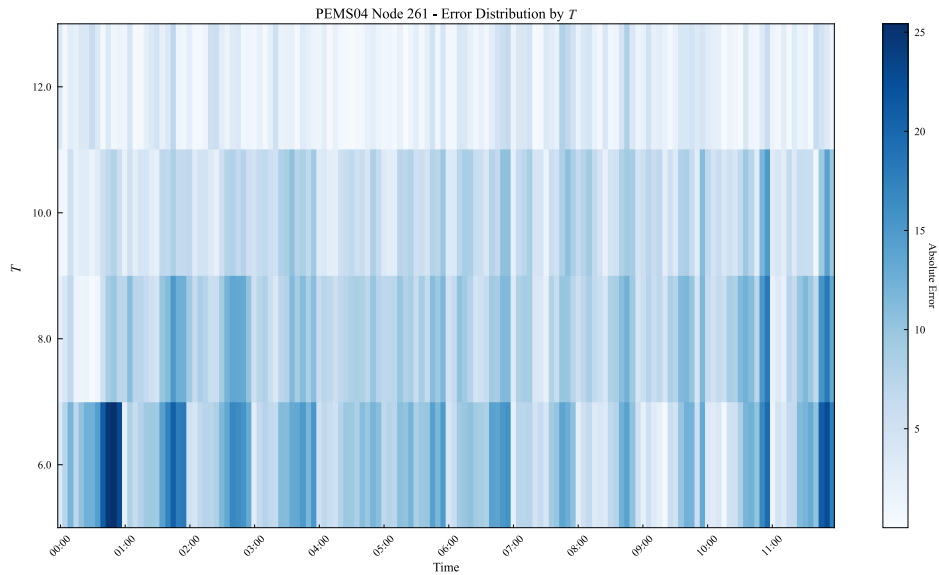


图8 不同步长参数下的误差热力图对比

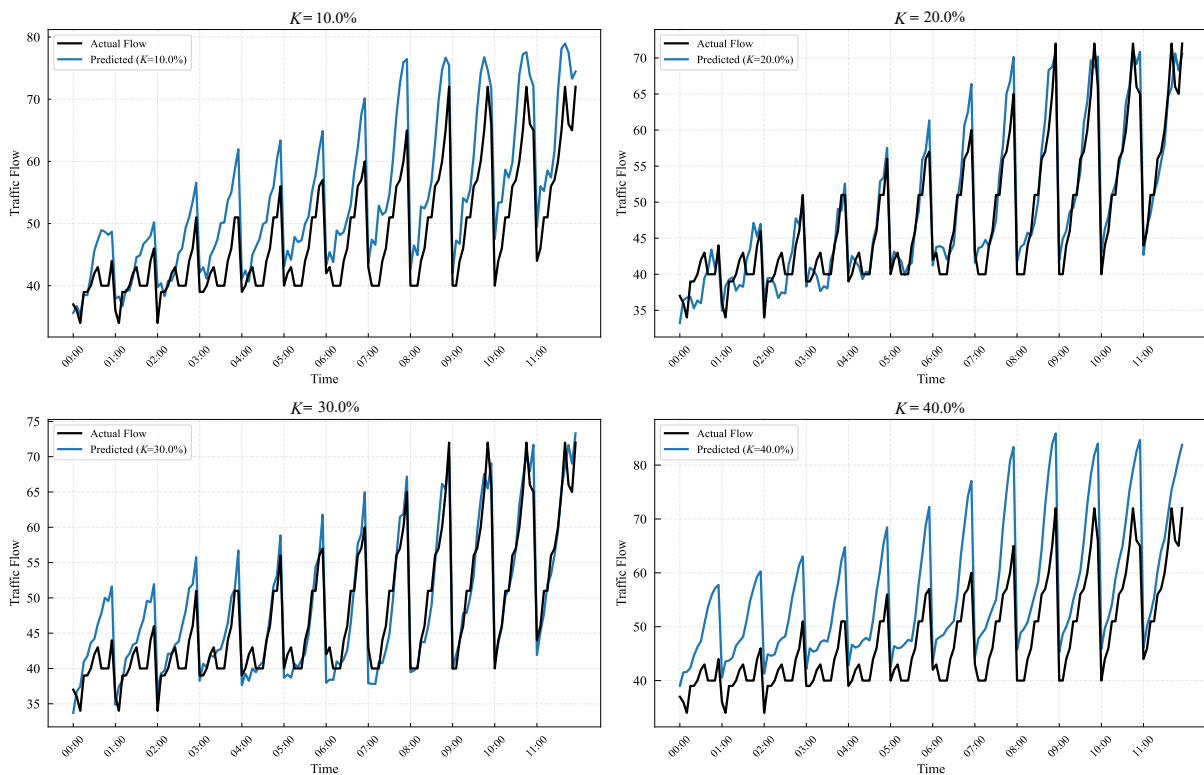


图9 不同K值下的同一数据集的预测结果对比图

网络进行预测。

从表5中可以看出,在PEMS03、PEMS04、PEMS07和PEMS08四个数据集中,移除不同模块模型的预测性能均有所下降.以PEMS07数据集为例,移除时序分解模块后,模型的MAE、RMSE和MAPE指标分别上升了4.8724、6.2768、1.56%,表明时序分解模块能够有效分解时序数据,捕捉交通流量数据中的延时

时间关系,提升模型的预测精度.以PEMS03数据集为例,移除Moran算子评估空间相关性后,评价指标MAE从18.2792上升至20.6073, RMSE从29.3309上升至31.8895, MAPE从10.40%上升至11.72%.这说明通过Moran算子评估站点间的空间相关性,在优化模型的预测结果方面起到了重要作用.移除延时时空特征提取模块后,模型预测性能最差,无法有效捕

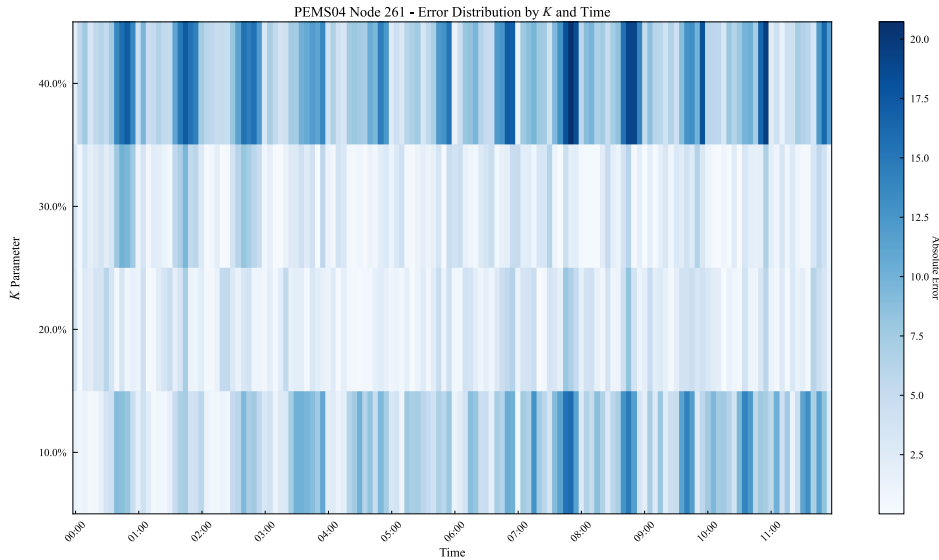


图 10 不同频带选择下的误差热力图对比

捉时序数据长期依赖和局部波动,说明仅使用全连接网络层进行预测无法充分捕捉交通流量数据中的复杂时空依赖关系.通过考虑交通场景中的延时时空关系,NSFM 能够更全面地建模交通流量的动态变化规律.

表 5 消融实验结果

数据集	模型	评价指标		
		MAE	RMSE	MAPE/%
PEMS03	w/o TDM	19.150 5	30.953 6	10.90
	w/o Moran	20.607 3	31.889 5	11.72
	w/o DSTFE	21.501 4	33.546 9	12.23
	NSFM	18.279 2	29.330 9	10.40
PEMS04	w/o TDM	25.003 6	37.067 0	11.21
	w/o Moran	25.886 5	38.415 2	11.60
	w/o DSTFE	26.511 6	39.154 9	11.88
	NSFM	24.639 8	36.453 8	11.04
PEMS07	w/o TDM	33.367 8	50.167 8	10.67
	w/o Moran	34.091 3	51.707 2	10.90
	w/o DSTFE	36.045 9	54.173 2	11.53
	NSFM	28.495 4	43.891 0	9.11
PEMS08	w/o TDM	20.981 9	30.743 1	9.00
	w/o Moran	21.069 8	31.141 8	9.04
	w/o DSTFE	22.840 7	35.787 7	13.00
	NSFM	19.382 9	28.009 6	8.31

5 结束语

本文研究交通场景下非平稳时间序列预测问题.针对已有交通流量预测模型不能有效建模交通流量的

非平稳性和延迟时空依赖关系的问题,本文提出了一种基于延迟时空依赖的非平稳时间序列流量预测模型.该模型将时间序列分解为时变和时不变组分,并结合频域分析方法提取局部化特征,有效捕捉了交通流量的动态波动特性.引入延迟时空依赖特征建模,模型能够准确刻画交通流量传播的延时效应,进一步提升了预测精度.最后利用空间函数对函数回归方程和 Moran 算子,量化站点间的空间相关性,增强了预测结果的鲁棒性.实验结果表明,本文提出的模型在多个真实数据集上均表现出显著的性能优势,相较于现有的多个基准模型,能够提供更加精确的预测结果.尽管本文在利用深度学习技术进行时间序列预测的研究上取得了一定进展,未来还可以在可解释性方面进行更深入的研究.同时,可以进一步研究跨序列全局模式的特征提取方法,以及如何结合领域知识发现和利用全局和局部特征的方法等.

参考文献

[1] ZHANG J P, WANG F Y, WANG K F, et al. Data-driven intelligent transportation systems: A survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(4): 1624-1639.

[2] LIU R P, SHIN S Y. A review of traffic flow prediction methods in intelligent transportation system construction[J]. Applied Sciences, 2025, 15(7): 3866.

[3] 邓攀, 刘俊廷, 王晓, 等. STCTN: 一种基于时域偏倚校正与空域因果传递的时空因果表示学习方法[J]. 计算机学报, 2023, 46(12): 2535-2550.

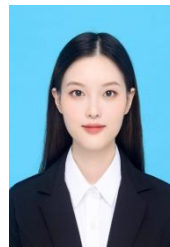
- DENG P, LIU J T, WANG X, et al. STCTN: A spatio-temporal causal representation learning method based on temporal bias adjustment and spatial causal transition[J]. Chinese Journal of Computers, 2023, 46(12): 2535-2550. (in Chinese)
- [4] 唐文杰, 肖一磊, 孔祥宇, 等. CycleLLH: 一种基于周期性整合的新型网络流量预测模型[J]. 计算机学报, 2024, 47(12): 2867-2888.
- TANG W J, XIAO Y L, KONG X Y, et al. CycleLLH: A new network traffic prediction model based on cycle integration[J]. Chinese Journal of Computers, 2024, 47(12): 2867-2888. (in Chinese)
- [5] 黄昕, 毛政元. 基于时空多图卷积网络的网约车乘客需求预测[J]. 地球信息科学学报, 2023, 25(2): 311-323.
- HUANG X, MAO Z Y. Prediction of passenger demand for online car-hailing based on spatio-temporal multi-graph convolution network[J]. Journal of Geo-Information Science, 2023, 25(2): 311-323. (in Chinese)
- [6] LAI Q, CHEN P. LEISN: A long explicit-implicit spatio-temporal network for traffic flow forecasting[J]. Expert Systems with Applications, 2024, 245: 123139.
- [7] GENG Z L, XU J, WU R S, et al. STGAFormer: Spatial-temporal gated attention transformer based graph neural network for traffic flow forecasting[J]. Information Fusion, 2024, 105: 102228.
- [8] PU B, LIU J S, KANG Y, et al. MVSTT: A multiview spatial-temporal transformer network for traffic-flow forecasting[J]. IEEE Transactions on Cybernetics, 2024, 54(3): 1582-1595.
- [9] XIAO J L, LONG B C. A multi-channel spatial-temporal transformer model for traffic flow forecasting[J]. Information Sciences, 2024, 671: 120648.
- [10] QIAN W Z, ZHAO Y, ZHANG D L, et al. Towards a unified understanding of uncertainty quantification in traffic flow forecasting[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(5): 2239-2256.
- [11] LIU X Y, LI X, FIUMARA G, et al. Link prediction approach combined graph neural network with capsule network[J]. Expert Systems with Applications, 2023, 212: 118737.
- [12] BOX G E P, JENKINS G M, REINSEL G C, et al. Time Series Analysis: Forecasting and Control[M]. 5th Ed. Hoboken: John Wiley & Sons, 2015.
- [13] QIAN Y R, CAI Q, PAN Y W, et al. Boosting diffusion models with moving average sampling in frequency domain[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 8911-8920.
- [14] YANG H M, PAN Z S, TAO Q, et al. Online learning for vector autoregressive moving-average time series prediction[J]. Neurocomputing, 2018, 315: 9-17.
- [15] LONG L F, LIU Q, PENG H, et al. Multivariate time series forecasting method based on nonlinear spiking neural P systems and non-sampled shearlet transform[J]. Neural Networks, 2022, 152: 300-310.
- [16] TRAN D T, IOSIFIDIS A, KANNIAINEN J, et al. Temporal attention-augmented bilinear network for financial time-series data analysis[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(5): 1407-1418.
- [17] XUE Z W, ZHANG Y, CHENG C, et al. Remaining useful life prediction of lithium-ion batteries with adaptive unscented Kalman filter and optimized support vector regression[J]. Neurocomputing, 2020, 376: 95-102.
- [18] BILLINGS S A, HONG X. Dual-orthogonal radial basis function networks for nonlinear time series prediction[J]. Neural Networks, 1998, 11(3): 479-493.
- [19] 乔少杰, 吴凌淳, 韩楠, 等. 情景感知驱动的移动对象多模式轨迹预测技术综述[J]. 软件学报, 2023, 34(1): 312-333.
- QIAO S J, WU L C, HAN N, et al. Multiple-motion-pattern trajectory prediction of moving objects with context awareness: A survey[J]. Journal of Software, 2023, 34(1): 312-333. (in Chinese)
- [20] ZHENG H F, LIN F, FENG X X, et al. A hybrid deep learning model with attention-based conv-LSTM networks for short-term traffic flow prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(11): 6910-6920.
- [21] SHI X M, QI H, SHEN Y M, et al. A spatial-temporal attention approach for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(8): 4909-4918.
- [22] MÉNDEZ M, MERAYO M G, NÚÑEZ M. Long-term traffic flow forecasting using a hybrid CNN-BiLSTM model[J]. Engineering Applications of Artificial Intelligence, 2023, 121: 106041.
- [23] GUO L B, WANG W Q, CHEN Z, et al. Newton-Cotes graph neural networks: On the time evolution of dynamic systems[EB/OL]. (2023-10-20)[2025-10-10]. <https://arxiv.org/abs/2305.14642>.
- [24] XU M X, DAI W R, LIU C M, et al. Spatial-temporal

- transformer networks for traffic flow forecasting[EB/OL]. (2023-03-29)[2025-10-10]. <https://arXiv.org/abs/2001.02908>.
- [25] WANG X, ZHOU T, WEN Q S, et al. CARD: Channel aligned robust blend transformer for time series forecasting[EB/OL]. (2024-02-16)[2025-10-10]. <https://arxiv.org/abs/2305.12095>.
- [26] ZHAO K, ZHANG L. Causality-inspired spatial-temporal explanations for dynamic graph neural networks[C]//12th International Conference on Learning Representations (ICLR 2024). Appleton: ICLR, 2024: 1-13.
- [27] YADAV H, SCHAEFER M, ZHAO K, et al. CASPFormer: Trajectory prediction from BEV images with deformable attention[C]//Pattern Recognition. Cham: Springer, 2025: 420-434.
- [28] BAI L, YAO L N, LI C, et al. Adaptive graph convolutional recurrent network for traffic forecasting[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: ACM, 2020: 17804-17815.
- [29] LI F X, FENG J, YAN H, et al. Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution[J]. ACM Transactions on Knowledge Discovery from Data, 2023, 17(1): 1-21.
- [30] BAI J D, ZHU J W, SONG Y J, et al. A3T-GCN: Attention temporal graph convolutional network for traffic forecasting[J]. ISPRS International Journal of Geo-Information, 2021, 10(7): 485.
- [31] LONG B C, ZHU W, XIAO J L. ST-RetNet: A long-term spatial-temporal traffic flow prediction method[EB/OL]. (2024-07-13)[2025-10-10]. <https://arxiv.org/abs/2407.11074>.
- [32] ZHAO Y, DENG P, LIU J T, et al. Causal conditional hidden Markov model for multimodal traffic prediction[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(4): 4929-4936.
- [33] ZHOU Z, SHOJAFAR M, ABAWAJY J, et al. ECMS: An edge intelligent energy efficient model in mobile edge computing[J]. IEEE Transactions on Green Communications and Networking, 2022, 6(1): 238-247.
- [34] DAS S, DAS S K. A probabilistic link prediction model in time-varying social networks[C]//2017 IEEE International Conference on Communications. Piscataway: IEEE, 2017: 1-6.
- [35] BHATKAR S, GOSAVI P, SHELKE V, et al. Link prediction using GraphSAGE[C]//2023 International Conference on Advanced Computing Technologies and Applications. Piscataway: IEEE, 2024: 1-5.
- [36] SONG C, LIN Y F, GUO S N, et al. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 914-921.
- [37] WU Z H, PAN S R, LONG G D, et al. Graph WaveNet for deep spatial-temporal graph modeling[EB/OL]. (2019-05-31)[2025-10-01]. <https://arxiv.org/abs/1906.00121>.
- [38] BAI L, YAO L N, KANHERE S S, et al. STG2Seq: Spatial-temporal graph to sequence model for multi-step passenger demand forecasting[EB/OL]. (2019-05-24)[2025-10-10]. <https://arxiv.org/abs/1905.10069>.
- [39] GUO S N, LIN Y F, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[C]//Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence. New York: ACM, 2019: 922-929.

作者简介



李小龙 男,1981年6月出生于湖南省常德市.现为湖南工商大学计算机学院教授、博士生导师.主要研究方向为智慧交通、大模型、物联网等.
E-mail: lxl@hutb.edu.cn



李曦 女,2000年11月出生于湖南省沅江市.湖南工商大学计算机学院硕士.主要研究方向为深度学习、数据分析等.
E-mail: lmz1191652939@163.com



刘 洋 女,1981年7月出生于湖南省常德市. 现为湖南工商大学计算机学院助理研究员、博士. 主要研究方向为深度学习、时空网络、知识图谱等.

E-mail: 4889329@qq.com



易畅言 男,1989年10月出生于浙江省绍兴市. 现为南京航空航天大学计算机科学与技术学院教授. 主要研究方向为随机优化、博弈论、激励机制、排队调度和人工智能等. 中国电子学会会员编号:E190186644M.

E-mail: changyan.yi@nuaa.edu.cn



李柄廷 男,1999年8月出生于湖南省长沙市. 现为湖南工商大学计算机学院硕士生. 主要研究方向为图神经网络、并行图着色等.

E-mail: 437138609@qq.com



曾宁俊 男,1994年12月出生于湖南省益阳市. 现为湖南工商大学计算机学院讲师. 主要研究方向为复杂非线性系统建模、控制与优化等.

E-mail: 3149@hutb.edu.cn