

DOI:10.13232/j.cnki.jnju.2026.01.008

面向大规模路网的车辆行程时间不确定性量化方法研究

吴庭锐¹, 毛潇苇¹, 郭晟楠^{1,3*}, 林友芳^{1,2}, 万怀宇^{1,2}

(1. 北京交通大学计算机科学与技术学院, 北京, 100044; 2. 北京交通大学交通数据挖掘与具身智能北京市重点实验室, 北京, 100044; 3. 交通大数据与人工智能教育部重点实验室, 北京, 100044)

摘要: 随着智能交通系统和共享出行服务的迅猛发展, 人们对行程时间预测的需求日益增长, 准确的行程时间预测已成为提升交通效率和优化用户体验的重要任务. 传统的行程时间估计方法大多侧重于预测均值, 提供点估计结果, 忽略了复杂且动态变化的交通状况带来的不确定性. 量化行程时间的不确定性以提供包含置信区间的结果, 可以为用户和共享出行平台提供更全面可信的预测信息. 但是, 由于动态变化的路段通行时间分布以及多个路段通行时间不确定性的累积问题, 难以准确量化行程时间的不确定性. 为解决上述问题, 提出了一种基于动态交通路况的行程时间预测与不确定性量化方法, 设计了一个分布感知行程时间不确定性估计模型 (Distribution Aware Travel Time Estimation, DATE), 该模型包括路网模块化模块、全局分布感知器模块以及分布融合不确定性估计模块. 该模型能在准确预测行程时间的同时, 提供可靠的置信区间, 全面量化不确定性. 实验结果表明, DATE 在两个真实数据集上的表现优于现有方法, 且能有效提高行程时间预测的精度和可靠性, 为智能交通系统提供更为可靠的决策支持.

关键词: 行程时间预测, 不确定性量化, 分布感知, 时空数据挖掘, 机器学习

中图分类号: TP183

文献标志码: A

Research on vehicle travel time uncertainty quantification method for large-scale road networks

Wu Tingrui¹, Mao Xiaowei¹, Guo Shengnan^{1,3*}, Lin Youfang^{1,2}, Wan Huaiyu^{1,2}

(1. School of Computer Science and Technology, Beijing Jiaotong University, Beijing, 100044, China;

2. Beijing Key Laboratory of Traffic Data Mining and Embodied Intelligence, Beijing, 100044, China;

3. Key Laboratory of Big Data & Artificial Intelligence in Transportation, Ministry of Education, Beijing, 100044, China;)

Abstract: With the rapid development of intelligent transportation systems and shared mobility services, the demand for accurate travel time prediction has been increasing. As a result, accurate travel time estimation has become a crucial task for improving traffic efficiency and optimizing user experience. Traditional travel time estimation methods mostly focus on predicting the mean value and provide point estimates, while ignoring the uncertainty caused by complex and dynamically changing traffic conditions. Quantifying the uncertainty of travel time and providing results with confidence intervals can offer more comprehensive and trustworthy predictions for users and mobility platforms. However, due to the dynamically varying travel time distributions of road segments and the accumulated uncertainty across multiple segments, it remains challenging to quantify travel time uncertainty accurately. To address this issue, this paper proposes a travel time prediction and uncertainty quantification method based on dynamic traffic conditions. A novel model, Distribution Aware Travel Time Estimation (DATE), is designed, which consists of a road network partitioning module, a global distribution-aware module, and a distribution fusion-based uncertainty estimation module. This model not only improves the accuracy of travel time prediction

基金项目: 国家自然科学基金(62202043), 北京市自然科学基金(4242029)

收稿日期: 2025-09-09

* 通信联系人, E-mail: guoshn@bjtu.edu.cn

but also provides reliable confidence intervals for comprehensive uncertainty quantification. Experimental results on two real-world datasets demonstrate that DATE outperforms existing methods in terms of both prediction accuracy and reliability, offering robust decision support for intelligent transportation systems.

Keywords: travel time estimation, uncertainty quantification, distribution awareness, spatial-temporal data mining, machine learning

行程时间估计(Travel Time Estimation, TTE)指的是预测车辆沿查询路径从一个地点到另一个地点所需的时间,对导航和行程规划至关重要。行程时间估计可以视为一个时空领域内的回归任务^[1],其特点是观测数据稀疏、不规则且片段化^[2]。与传统的时空数据预测(如时间序列预测)相比,行程时间估计更复杂,主要体现在其时序的连续性和行程序列中固有的复杂空间依赖性。因此,行程时间估计近年来受到了广泛关注,为自动驾驶决策和导航等技术提供了重要的数据支持。

然而,在许多实际应用中,仅提供行程时间的平均估计值是远远不够的。一种更为可靠、信息量更丰富的方法是量化行程时间不确定性。具体而言,我们的目的是估计具有指定置信水平的行程时间的置信区间,以量化不确定性,这在许多场景中都有重要意义。例如,网约车服务可以从为顾客提供出行时间的置信下界和置信上界中获益,以便更好地规划自己的行程。此外,了解行程时间不确定性可以帮助网约车和物流平台提高决策有效性,例如,订单调度和车辆调度。

量化行程时间不确定性的关键在于预测每个路段通行时间分布,此外,还需要捕获多个路段通行时间不确定性的累积。要完成这一任务将会面临许多挑战。

由于交通流量在相邻路段之间具有较强的时空依赖性,前一个路段的交通状况直接影响到后续路段的行程时间。例如,在城市道路网络中,前一个路段的拥堵情况或交通事件(如交通信号灯、事故等)直接影响车辆在下一个路段的行驶速度,导致后续路段的行程时间发生变化。忽视这种相邻路段间的依赖关系可能导致行程时间预测不准确。因此,如何准确捕捉和建模这种时空依赖性,成为了行程时间不确定性量化中的一大挑战。

此外,每个路段上的行程时间状况都是随时间不断变化的,对总体路径不确定性的贡献也在随之变化,如何捕捉这种变化、有效建模不同条件下各路段行程时间对整体行程时间不确定性的影响十分具有挑战性。如图 1 所示,行程 1 中车辆到达路段 e_6 的时间为 9:00,由于 8:50—9:00 的统计行程时间分布集中在 30 s 左右,因此,行程 1 中车辆在 e_6 的行程时间是相对确定的。然而,在 9:10—9:20,车辆在 e_6 的行程时间分布范围更广,表现出更高的方差和波动性。增加了行程 2 在 9:20 开始的不确定性。受交通状况、出发时刻等复杂因素的影响,路段层面的行程时间不确定性不断变化,难以准确量化整体路径层面的行程时间不确定性。

在过去几十年中,许多算法已被用于精准预测未来的交通状况,这些方法从统计模型^[3]到数据驱动模型^[4]都有涉及。最初,时序建模方法例如循环神经网络(RNN)和自回归综合滑动平均(ARIMA)模型用于学习交通数据中单一时序的时间演化^[5-6],但这些方法忽略了多个交通时序之间的空间传递关系。随着技术的发展,时空图神经网络(Spatio-Temporal Graph Neural Networks, ST-GNN)成为捕捉时空依赖关系的前沿方法,能协同捕捉空间和时间依赖性^[7-9]。然而,传统时空图神经网络中的固定静态图无法反映不同时间阶段的各种空间相关性,这是因为交通点的空间关联性随时间的变化而变化。因此,动态空间建模技术逐渐成为交通预测的主流研究方向^[10],旨在揭示每个时间片的空间相关性并动态地传播空间信息。然而,大多数动态空间建模方法的计算复杂度相对行程长度呈二次幂,因此交通预测通常只能在较小的区域尺度进行,难以满足现实城市中包含数千个交通点的交通预测需求^[11]。

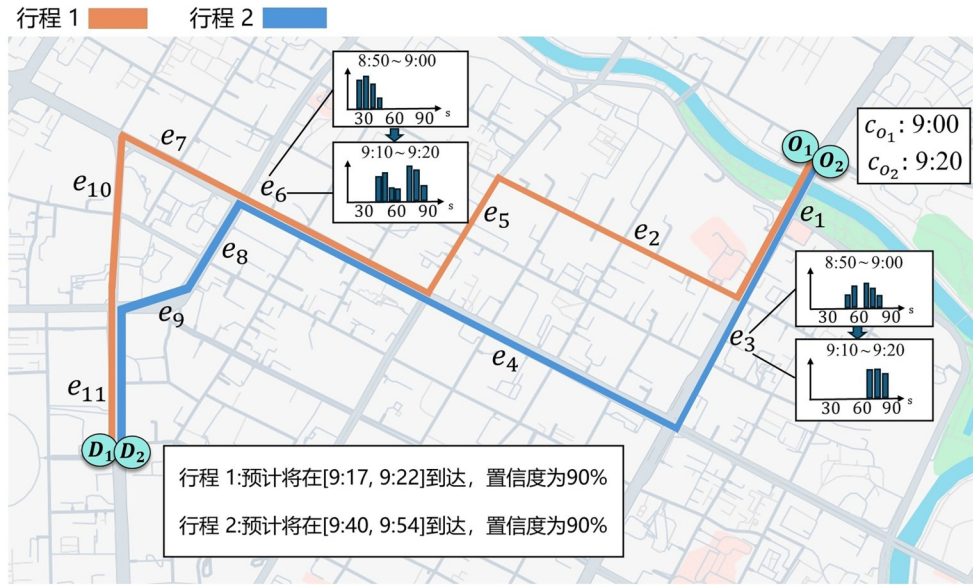


图1 动态变化的路段行程时间分布

Fig. 1 Dynamic changing travel time distribution of road segments

为了解决上述问题,本文提出了一种基于分布感知的行程时间预测模型——DATE (Distribution Aware Travel Time Estimation). 模型有一个时间分布感知模块,首先将路网中的路段划分成一个个板块(Patch)并构建子图,然后将路段信息编码后分别对板块内和板块间交替使用深度和广度注意力机制来得到各个路段之间复杂的空间和时序依赖关系,从而捕捉各个路段行程时间分布的变化情况,同时给出行程轨迹中未来每个路段的通行时间分布预估. 这样的设计降低了动态空间建模的复杂度,保证了可解释性和保真性. 接着,模型通过分布融合不确定性估计模块,结合各路段的历史信息和分布预估进行端到端训练,最终输出通行时间预测值和置信区间的估计. 这一创新性方法不仅提高了行程时间预测的准确性,还提供了对预测不确定性的全面量化,为智能交通系统中的决策支持提供了更为可靠的数据依据.

1 相关工作

1.1 行程时间估计 行程时间估计的研究可以追溯到20世纪90年代初. 早期方法是将路段历史行程时间相加来估计到达时间,这些方法能为出行者在一定程度上提供行程时间估算. 然而,

随着大数据和实时交通信息的不断发展,基于动态数据的行程时间预测逐渐成为研究主流. 当前的研究方法主要包括统计回归、机器学习以及深度学习等模型. 基于统计回归的估计方法主要通过回归分析技术,例如,加权回归和多元线性回归,为行程时间提供初步的估计,但由于缺乏对动态交通流和时空变动的考虑,常常无法适应复杂的交通状况.

近年来,随着机器学习方法的兴起,许多研究者开始采用支持向量机(SVM)^[12]、决策树^[13]和随机森林(RF)^[14]等算法进行行程时间预测. 这些方法能在非线性和高维数据处理中展现出较强的优势,能更好地应对复杂交通环境. 例如,支持向量回归(SVR)^[15]模型通过引入非线性关系,显著提高了行程时间的预测精度. 随着深度学习技术的突破,卷积神经网络(CNN)^[16]和循环神经网络(RNN)^[17]逐渐成为行程时间估计的主流方法. 基于长短期记忆网络(LSTM)^[18]的预测方法,通过建模历史交通数据的时序特性,能更好地捕捉动态交通流的波动,为行程时间提供准确的预测.

为了进一步提升预测的准确性,研究者们提出了更复杂的模型来捕捉动态复杂的时空依赖关系. 许森^[19]构建了基于动态时空图卷积网络和多组件融合机制的预测框架,在综合考虑交通流动

态性与异常事件影响的基础上,实现了在复杂城市路况下的高精度行程时间预测. Fang et al^[20]提出的 ConSTGAT 设计了一个上下文感知的时空图注意力网络,在统一框架内融合了交通预测与路径的上下文信息. Chen et al^[21]提出的 HierETA 则采用分层自注意力网络,创新地从路段、路链和交叉口等多个不同粒度的视角来综合建模轨迹. 这些先进的模型通过更强大的网络结构,得到更准确的行程时间预测结果.

这些方法虽然在行程时间预测方面取得了显著进展,但它们主要关注点估计,未考虑行程时间的不确定性,难以提供可信的置信区间;同时对动态交通变化的适应性有限,在面对大规模路网中复杂且动态变化的交通状况时,未充分考虑路段间的时空依赖关系与交通流波动性,影响了在实际复杂场景中的稳健性与可靠性.

1.2 行程时间不确定性量化 尽管现有的行程时间估计方法在准确性上取得了显著进展,但通常仅提供均值估计,缺乏对行程时间不确定性的量化. 为了解决这一问题,许多研究开始关注如何在行程时间估计中引入不确定性量化,尤其是置信区间的设计. 贝叶斯神经网络(BNN)^[22]是一种常见的用于量化不确定性的方法,通过在神经网络中引入概率分布,贝叶斯方法能提供行程时间估计的置信区间. 例如,一些研究采用贝叶斯神经网络对交通流进行建模,不仅估计了行程时间的均值,还提供了相应的不确定性范围,显著提升了预测结果的可靠性. 蒙特卡罗方法通过多次随机采样和概率统计来计算行程时间的不确定性,能为每个预测提供置信区间,尽管该方法能提供较为全面的结果,但其计算复杂度较高,且在实时预测中存在一定的限制. 近年来,研究者们探索了多种结合深度学习的概率估计框架. 例如,基于图神经网络(GNN)^[23]的方法有效捕捉了路段间复杂的空间相关性,并结合时间序列模型(如 RNN 或 LSTM)建模路况随时间的动态变化. 这些方法通过融合概率推断与深度神经网络的强大特征表示能力,显著提高了行程时间预测的可靠性与不确定性量化的精度.

此外, Mao et al^[24]提出的 GMDNet 创新性地使用了混合密度网络来预测行程时间的多峰形态

分布,能有效捕捉由不同交通状况导致的多种可能结果,但其基于全图的编码方式在应用于大规模城市路网时面临扩展性挑战. Xu et al^[25]提出的 ProbTTE 模型,通过学习路段的低秩表示来构建多趟行程间的联合高斯分布,并基于实测的相邻行程进行条件推断,但其单峰高斯分布的假设可能难以完全捕捉城市环境中复杂的行程时间分布. 最近,一些新兴的基于高斯过程的框架(如 SPTTE 模型^[26])采用时空联合分布建模,能在处理数据稀疏性和空间异质性时提供更加稳健的不确定性量化结果. 这些方法通过建模路径上的各路段之间的相关性,将动态交通流与空间拓扑关系结合起来,实现了对复杂交通环境的不确定性描述. 通过这些研究,行程时间不确定性量化的方法在理论和应用方面都得到了显著提升.

尽管上述方法在行程时间不确定性量化方面各具优势并取得了显著成果,但由于建模复杂、计算开销大、对特定场景的依赖或模型假设的限制等问题,普遍难以在大规模路网中高效应用.

2 问题定义

2.1 道路网络定义 道路网络(以下简称路网)定义为一个研究区域或城市内所有路段的集合,被建模为一个有向图 $\mathcal{G}=(\mathcal{V}, \mathcal{E})$, 其中, \mathcal{V} 是节点 v_i 的集合, v_i 用以表示路口或路段的起止点; \mathcal{E} 是边 e_i 的集合, e_i 用以表示路段,即节点之间的连接道路. 路网中的每个节点和路段都有一个唯一的索引.

2.2 轨迹、行程和路径定义 轨迹 \mathcal{T} 是一组带有时间戳的 GPS 点的序列,如式(1)所示:

$$\mathcal{T}=\langle\langle g_1, c_1 \rangle; \dots; \langle g_{|\mathcal{T}|}, c_{|\mathcal{T}|} \rangle\rangle \quad (1)$$

其中, $g_i=(lng_i, lat_i)$ 表示第 i 个 GPS 点, lng 表示经度, lat 表示纬度, $|\mathcal{T}|$ 表示轨迹中的总 GPS 点数.

行程 $x_{\mathcal{T}}$ 是一组节点的序列,表示为 $x_{\mathcal{T}}=\{(v_1, c_1), \dots, (v_k, c_k)\}$, 其中,时间戳 c 是单调递增的. 行程的总时间 $y=c_k-c_1$. 值得注意的是,多个 GPS 数据点可以位于同一条路段上.

路径被定义为一系列节点的序列,表示为

$x = (v_1, \dots, v_{|x|})$, 其中, 每一对相邻节点是相邻的, 即 $\forall i = 0, \dots, |x| - 1, (v_i, v_{i+1}) \in \mathcal{E}$, 其中, \mathcal{E} 是路网中的边集合, 表示相邻节点之间的连接关系.

2.3 问题陈述 我们旨在学习一个映射函数 f_θ , 该函数基于给定路径和出发时间的输入, 预测总行程时间 y 及其上下置信区间, 分别记为 \hat{l} 和 \hat{u} , 以量化行程时间的不确定性. 这样, $[\hat{l}, \hat{u}]$ 表示一个能覆盖实际行程总时间的置信区间, 且置信度为 $1 - \rho$, 将路径和出发时间的输入对记作 $req = (r, s)$, 其中, r 表示路径, s 表示行程出发时间.

3 本文方法

DATE 模型的整体架构如图 2 所示, 划分为三个模块: 路网板块化模块 (包含图中板块划分和板块填充)、全局分布感知器模块以及分布融合不确定性估计模块.

为了降低处理大规模路网的复杂度, 首先将整个路网上的路段按照一阶邻接关系不重不漏地切分为若干板块 (*Patch*) 来构建子图, 这样可以尽量保留路段间的空间相关性.

接着, 通过对每个 *Patch* 中已有的路段信息进行简单的平均处理, 用以填充当前 *Patch*, 使每个 *Patch* 里的元素个数统一, 方便后续模型处理.

随后, 处理好的所有 *Patch* 被送入双注意力编码器, 深度注意力作用于每个 *Patch* 内部, 使每个路段都能融合邻居路段的信息. 广度注意力作用于每个 *Patch* 之间, 使每个路段都能融合更远的全局的路段的信息. 两者交替使用且均采用多头注意力机制, 以多视角并行建模并融合时空信息, 以便每个路段更好地捕捉到整个路网的信息. 编码结果经投影解码器映射为离散的行程时间概率分布, *Softmax* 保证各时间区间概率和为 1, 可以直接得到各路段的完整通行时间分布, 而不是单点预测. 将预测分布嵌入向量与额外的历史信息一同输入 LSTM, 进一步建模序列动态变化.

最终, 输出行程时间的点估计及上下置信区间, 既准确预估了行程时间, 又实现不确定性量化. 整套流程借助 *Patch* 级局部拓扑、双重注意力以及分布式输出与序列建模的结合, 达到对行程时间及其不确定性的高精度联合预测.

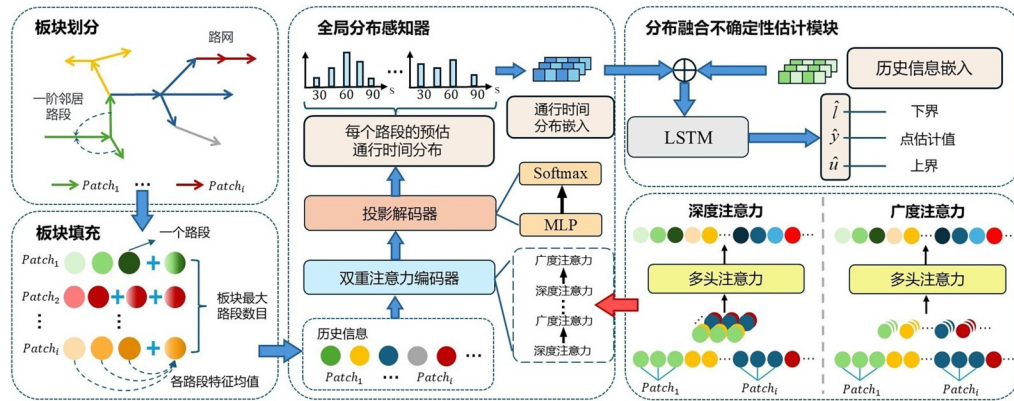


图 2 DATE 总体架构

Fig. 2 The DATE framework

3.1 路网板块化模块 为了提高模型在复杂交通网络中的适应性和计算效率, 路网板块化模块通过合理划分路网并构建板块子图, 能优化时空特征的处理和融合, 降低动态空间建模的计算复杂度.

在构建路网板块化输入之前, 首先对时间

域进行离散化处理, 将一天 24 h 等分为长度为 10 min 的时间片, 共有 144 个时间片, 每个时间片都是一个离散区间. 给定某个具体时刻 τ , 其所在的时间片索引 s 表示如式 (2) 所示:

$$s = \left\lfloor \frac{hour(\tau) \times 60 + minute(\tau)}{10} \right\rfloor, s \in \{0, \dots, 143\} \quad (2)$$

然后,对每个路段 u ,构建其在目标时刻所在时间片 s 之前 T 个时间片 $([s-T, \dots, s-1])$ 上的历史通行时间信息序列. 令在时间片 t 上路段 u 的通行时间分布如式(3)所示:

$$\phi'_u = [\phi'_u(1), \phi'_u(2), \dots, \phi'_u(K)]^T \in \mathbb{R}^K \quad (3)$$

其中, K 为离散区间数,其平均通行时间为标量 $\mu'_u \in \mathbb{R}$,将这两部分特征在时间维度上拼接,得到路段 u 的历史特征张量:

$$X_u = [(\phi_u^{s-T}, \mu_u^{s-T}), (\phi_u^{s-T+1}, \mu_u^{s-T+1}), \dots, (\phi_u^{s-1}, \mu_u^{s-1})] \in \mathbb{R}^{T \times (K+1)} \quad (4)$$

随后,利用路网的邻接矩阵 $A \in \{0, 1\}^{N \times N}$ (N 为路段总数),将每个路段 u 与其一阶邻居集合共同组成局部子图 (*Patch*),如式(5)所示:

$$N(u) = \{v | A_{uv} = 1 \text{ or } A_{vu} = 1\} \cup \{u\} \quad (5)$$

为了不重不漏地划分路网,每划分一个 *Patch* 后,都要把 *Patch* 里所有的路段从路网中删除,再对剩余路段进行新一轮划分,直到所有路段都被划分到各自的 *Patch* 中. 由于不同路段的一阶邻居数可能不一致,为了后续模型批量处理,需要将每个 *Patch* 统一为大小为 M 的定长集合 (M 为最大 *Patch* 的长度). 记 $|\mathcal{P}_u|$ 为每个 *Patch* 中的实际路段数目,若 $|\mathcal{P}_u| < M$,则用该子图内所有节点在每个时间片 t 上的特征均值进行填充:

$$\bar{X}' = \frac{1}{|\mathcal{P}_u|} \sum_{v \in \mathcal{P}_u} X'_v, X'_v = (P'_v, \mu'_v) \quad (6)$$

将 \bar{X}' 复制至补齐位置,直至节点数达到 M . 经过上述处理后,每个路段 u 对应的 *Patch* 表示为一个形状为 $M \times T \times (K+1)$ 的张量,如式(7)所示:

$$P_u = [X_1, X_2, \dots, X_M] \in \mathbb{R}^{M \times T \times (K+1)} \quad (7)$$

所有 *Patch* 一并输入后续的全局分布感知器模块,为空间-时序特征学习打下统一且可并行的输入基础.

3.2 全局分布感知器模块 在全局分布感知器模块中,我们的目标是捕捉并建模路网中各路段间复杂的时空依赖关系,更好地估计路径的通行时间及其不确定性.

如图 2 和图 3 所示,本模块接收路网板块化模块输出的路网划分后的 *Patch* 张量:

$$P = \{P_i\}, P \in \mathbb{R}^{P \times M \times T \times (K+1)} \quad (8)$$

其中, P 表示 *Patch* 的数量,首先通过线性映射将原始通行时间分布与平均通行时间特征投射到隐藏维度 d :

$$\mathcal{H}^{(0)} = PW^0 + b^0, W^0 \in \mathbb{R}^{(K+1) \times d}, b^0 \in \mathbb{R}^d, \mathcal{H}^{(0)} \in \mathbb{R}^{P \times M \times T \times d} \quad (9)$$

随后对投射后的特征进行 L 层交替编码,每层统一采用如下表示:

$$\mathcal{H}^{(1)} = \text{LayerNorm}(\mathcal{H}^{(t-1)} + \text{DAttn}(\mathcal{H}^{(t-1)}) + \text{BAttn}(\mathcal{H}^{(t-1)})) \quad (10)$$

其中, *DAttn* 和 *BAttn* 分别表示深度注意力和广度注意力.

深度注意力 (Depth Attention) 部分作用于每个 *Patch* 内的 M 个路段节点,用于捕获局部空间依赖. 举例来说,对某 *Patch* \mathcal{P}_i ,时间片 t 上的特征矩阵,记为 $X \in \mathbb{R}^{M \times d}$,深度注意力表示如式(11)所示:

$$\text{DAttn}(X) = \text{Softmax}\left(\frac{Q_d K_d^T}{\sqrt{d_k}}\right) V_d \quad (11)$$

其中, Q_d, K_d, V_d 均由输入 X 经线性映射得到.

广度注意力 (Breadth Attention) 部分作用于跨 *Patch* 的同一“相对位置”节点,用以捕捉全局空间依赖. 对跨所有 *Patch*,在某个相同位置上的节点特征合并起来,记为 $Y \in \mathbb{R}^{P \times d}$,广度注意力表示如式(12)所示:

$$\text{BAttn}(Y) = \text{Softmax}\left(\frac{Q_b K_b^T}{\sqrt{d_k}}\right) V_b \quad (12)$$

其中, Q_b, K_b, V_b 均由输入 Y 经线性映射得到.

在完成 L 层深度和广度注意力机制的交替编码过后,将得到向量 $\mathcal{H}^{(L)} \in \mathbb{R}^{P \times M \times T \times d}$,这时候将所有的 *Patch* 合并起来,丢掉一开始使用路段历史通行时间分布均值填充得到的节点,提取出原始的 N 条有效路段,得到 $X_{\text{eff}} \in \mathbb{R}^{T \times N \times d}$,最后将每个节点的最后表示通过一个线性层映射到 K 维通行时间分布的 logits:

$$l = X_{\text{eff}} W^0 + b^0 \in \mathbb{R}^{T \times N \times K} \quad (13)$$

将 l 在时间维度上取最后一片 (代表往后预测一个时间片):

$$\hat{D} = \text{Softmax}(l_{T-1, \dots}) \in \mathbb{R}^{1 \times N \times K} \quad (14)$$

其中, *Softmax* 在最后一个维度 (K 个区间) 上进行,输出即为每条路段在其所在轨迹出发时间片

的后一个时间上的通行时间分布预测,用于后续的不确定性量化。

本模块的核心是深度和广度注意力机制,这样的组合能让模型在两个尺度上同时学习交通网络的空间结构。深度注意力专注于同一Patch内相邻路段之间的细粒度依赖,捕捉局部道路网节点在微观层面的相互影响,例如交叉口附近的拥堵传播和信号灯周期带来的短时序上下文;广度注意力则跨越不同Patch,将城市不同区域之间的长距离关联纳入考量,如主干道与支路之间在高峰期的联动效应或事故影响的远程扩散。通过这样的“先局部再全局”的分层建模,模型不仅可以

获得更丰富、更具辨识度的多尺度表征,还可以提高对不同交通模式(如早晚高峰、节假日出行)的适应性和鲁棒性。

此外,这样的机制大幅降低了计算复杂度。传统点积方式对所有 N 个路段直接做一次全局注意力的复杂度为 $O(N^2)$,而本文改进之后的方法对 P 个Patch乘 M 个节点做一次深度和广度注意力的复杂度为 $O(PM^2 + MP^2)$ 。由实验可知, P 比 N 小一至两个量级, M 为个位数,因此 $O(PM^2 + MP^2) \ll O(N^2)$,极大地提升了模型的计算效率和可扩展性。

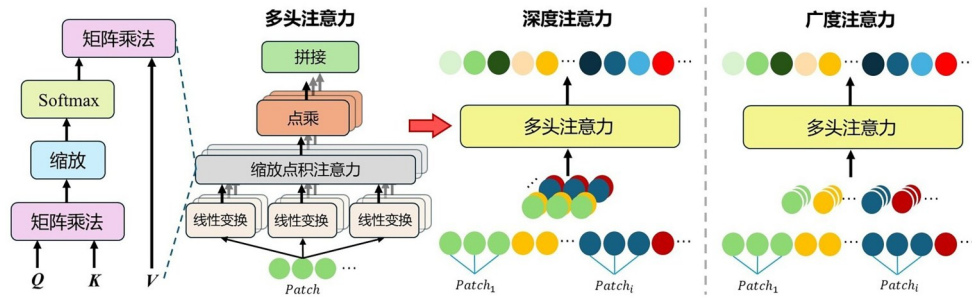


图3 深度和广度注意力机制

Fig. 3 The depth and breadth attention mechanisms

3.3 分布融合不确定性估计模块 本模块旨在在全局分布感知器输出的基础上,进一步融合每条轨迹中各路段的通行时间分布预测结果与其历史观测信息,实现对整体路径通行时间的精准估计及其不确定性区间的量化。本模块以轨迹为单位开展建模,综合考虑了多路段间的动态关联与误差累积特性,并通过端到端的方式输出行程时间的点估计及上下置信区间。设某条轨迹由 Q 条路段组成,各路段在轨迹出发时刻所在时间片的下一个时间片上的分布预测为:

$$\{\hat{d}_i\}_{i=1}^Q, \hat{d}_i = [\hat{d}_{i,1}, \dots, \hat{d}_{i,K}]^T, \sum_{k=1}^K \hat{d}_{i,k} = 1 \quad (15)$$

其中, K 为离散时间区间数目,各区间中心为 $\{c_k\}_{k=1}^K$,对于第 i 路段,基于该预测分布,计算加权平均的通行时间如式(16)所示:

$$\bar{t}_i = \sum_{k=1}^K \hat{d}_{i,k} c_k \quad (16)$$

记该路段在轨迹出发时间片的历史观测平均通行时间为 r_i ,为了充分融合历史信息与分布预

测信息,将两者及预估的完整分布拼接,构成第 i 路段的特征向量:

$$x_i = [r_i, \bar{t}_i, \hat{d}_{i,1}, \dots, \hat{d}_{i,K}]^T \in \mathbb{R}^{K+2} \quad (17)$$

将序列 $\{x_i\}_{i=1}^Q$ 输入LSTM模型,递归更新式如式(18)所示:

$$(h_i, c_i) = LSTM(x_i, h_{i-1}, c_{i-1}), i = 1, \dots, Q \quad (18)$$

取最后一步隐藏态 h_Q 作为整条轨迹的全局表征,最终,模型分别通过三组由两层全连接网络组成的输出头,预测轨迹总行程时间的点估计值、下界偏差 δ^l 与上界偏差 δ^u ,生成完整的置信区间如式(19)所示:

$$[\underline{y}, \bar{y}] = [\hat{y} - \delta^l, \hat{y} + \delta^u] \quad (19)$$

由于本文模型是为了得到更准确的点估计、更窄的上下界,同时上下界能将真实通行时间覆盖在内,为了达到这一目标,做到更好的平衡,我们设计了独立的预测头来估计通行时间以及上下界区间,监督目标为均值区间得分(Mean Interval

Score, MIS) 对于置信度为 $(1 - \rho)$ 的置信区间, 第 i 个样本的通行时间的预测上界和下界定义如式(20)所示:

$$\hat{u}_i = \hat{y}_i + \delta_i^u, \hat{l}_i = \hat{y}_i - \delta_i^l \tag{20}$$

MIS 损失函数表示如式(21)所示:

$$\mathcal{L}_{MIS} = \frac{1}{N} \sum_{i=1}^N \left[(\hat{u}_i - \hat{l}_i) + \frac{2}{\rho} \mathbb{I}\{y_i > \hat{u}_i\} (y_i - \hat{u}_i) + \frac{2}{\rho} \mathbb{I}\{y_i < \hat{l}_i\} (\hat{l}_i - y_i) + |y_i - \hat{y}_i| \right] \tag{21}$$

其中, \mathbb{I} 表示指示函数. 当 y_i 大于 \hat{u}_i 时, 惩罚项将增加; 当 y_i 小于 \hat{l}_i 时, 惩罚项也将增加, 从而鼓励预测区间包含实际的到达时间, 确保置信区间尽可能包含真实的到达时间. MIS 损失函数奖励较窄的置信区间, 同时鼓励它们包含实际的到达时间.

4 实验分析

为了评估本文提出的方法在行程时间预测和不确定性量化问题上的有效性, 我们在两个真实数据集上进行了实验, 并将所提出的算法与其他的几种先进方法进行了比较. 此外, 进行了消融实验, 旨在验证模型各个组件的有效性.

4.1 实验数据集 本文使用两个真实世界的数据集 XA 和 CD. 在对数据集进行预处理操作中, 将数据组织成路段通行时间分布以及路段通行时间均值的记录. XA 和 CD 数据集都是由滴滴公司发布的包含了 2018 年 10 月 31 日至 2018 年 11 月 20 日期间在成都和 2018 年 10 月 14 日至 2018 年

10 月 31 日期间在西安运营的出租车的 GPS 轨迹. 首先从 OpenStreetMap 中获取成都和西安两个城市的道路网络, 并将轨迹映射到该网络上. 接着, 将每天的 24 h 按照 10 min 一个时间片的方式均匀划分成 144 个时间片. 数据集统计情况如表 1 所示.

对 XA, CD 两个数据集进行初步处理和路网匹配后, 又对轨迹路段数量进行了统计. 统计结果如图 4 所示.

两个数据集大致按照 6:2:2 的比例来划分训练集, 验证集和测试集. 模型在训练集上进行训练, 并在测试集上进行评估, 验证集用于调整超参数和实现早停.

4.2 评价指标 为了全面地评估模型在行程时间预测和不确定性量化上的综合性能, 采用一系列评价指标. MIS 作为衡量模型不确定性量化综合质量的主要指标, 置信区间宽度 ($Width$) 和区间覆盖率 (Prediction Interval Coverage Probability, $PICP$) 两个指标作为参考. 此外, 为了专门评估模型点估计的准确性, 引入平均绝对误差 (Mean Absolute Error, MAE)、均方根误差 (Root Mean Squared Error, $RMSE$) 和平均绝对百分比误差 (Mean Absolute Percentage Error, $MAPE$) 三个标准回归指标.

在效率评估方面, 记录模型在整个测试集上的推理时间. XA 和 CD 数据集的测试样本数量分别为 159028 和 295785, 据此可以衡量模型在大规模路网下的实际运行效率.

(1) 不确定性量化评价指标

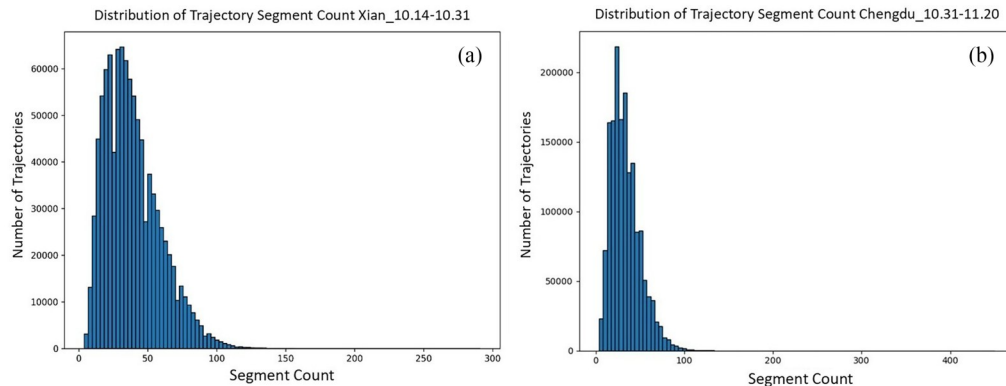


图 4 数据集轨迹路段数量统计

Fig. 4 Statistics on the number of segments in trajectories

表1 数据集统计情况

Table 1 Dataset statistics

数据集	XA	CD
时间跨度	2018.10.14— 2018.10.31	2018.10.31— 2018.11.20
全局轨迹数	991022	1620963
全局平均路段数	38.49	33.33
全局平均通行时间(s)	1006.91	907.85
全局平均轨迹长度(m)	4021.33	3782.47
路网路段数	9823	9112

$$MIS = \frac{1}{N} \left[\left(\hat{u}_i - \hat{l}_i \right) + \frac{2}{\rho} \mathbb{I}\{y_i > \hat{u}_i\} (y_i - \hat{u}_i) + \frac{2}{\rho} \mathbb{I}\{y_i < \hat{l}_i\} (\hat{l}_i - y_i) + |y_i - \hat{y}_i| \right] \quad (22)$$

MIS 同时衡量点估计的准确性、置信区间的宽度、置信区间对真实行程时间的覆盖率,是模型综合性能的直接度量,也是主要指标。 MIS 越小,模型越好。

$$Width = \frac{1}{N} \sum_{i=1}^N (\hat{u}_i - \hat{l}_i) \quad (23)$$

$Width$ 是置信区间的宽度,即预测的通行时间上界与下界之间的差值,衡量了模型对预测不确定性的表达程度。较宽的置信区间意味着模型对预测的不确定性较高,而较窄的置信区间表明模型对预测结果的信心较强。在一定程度上 $Width$ 越小,模型越好,但过小的 $Width$ 可能导致 $PICP$ 极低,将 $Width$ 作为参考指标。

$$PICP = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(\hat{l}_i \leq y_i \leq \hat{u}_i) \quad (24)$$

$PICP$ 表示预测置信区间能够覆盖真实通行时间的比例,反映了模型预测区间的可靠性。较高的 $PICP$ 表明模型在大多数情况下能够准确地预测通行时间,并且置信区间有效地覆盖了真实值。但 $PICP$ 过高也可能意味着 $Width$ 极大,将 $PICP$ 也作为参考指标。

通过 MIS , $Width$, $PICP$ 这三个指标的结合,可以更全面地评估模型在处理不确定性和动态环境下的表现。

(2)点估计评价指标

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (25)$$

MAE 计算点估计值与真实值之间绝对误差

的平均值。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (26)$$

$RMSE$ 计算预测误差平方的平均值的平方根,其对较大的预测误差给予了更高的权重,更能反映模型的偏差情况。

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (27)$$

$MAPE$ 计算预测误差占真实值的百分比的平均值。

MAE , $RMSE$, $MAPE$ 这三个指标越小,说明模型的点估计结果越准确。

4.3 基准模型 将提出的方法与其他几种行程时间估计的方法进行对比来衡量提出的方法的综合性能,对点估计模型进行修改,如ConSTGAT和HierETA,让其额外输出置信区间的上下界,以便参与 MIS 等指标的评估。

(1)历史分布卷积(Historical Distribution Convolution, HDC)通过结合每条轨迹每个路段的历史通行时间分布数据,利用卷积计算近似得到整条路径的通行时间分布。

(2)高斯分布先验(Gaussian Prior Distribution, GPD)假设路径通行时间服从高斯分布,通过LSTM预测其均值和标准差,以构建置信区间和点估计值。

(3)ConSTGAT^[20]模型采用3D图注意力网络(3DGAT),在图结构上融合空间与时间特征进行交通状态预测,同时使用局部卷积结构捕捉路径中相邻路段的上下文信息,最终得到路径通行时间的预测值。

(4)MC Dropout^[27]在DATE模型的基础上,在训练、验证和测试过程中都启用Dropout层,进行多次前向传播,生成多个预测结果作为采样,最终计算这些采样的均值和上下分位数,得到置信区间和点估计结果。

(5)ProbTTE^[25]模型通过学习路段表示来构建多趟行程间的联合多元高斯分布,并基于实测的相邻行程数据进行条件推断,最后得到查询路径的通行时间预测值。

(6)GMDNet^[24]模型首先利用图神经网络(GNN)对路径上的各个路段进行编码,以捕捉复

杂的空间依赖关系,然后通过一个混合密度网络来估计行程时间分布的多个组成部分(如高斯混合模型的均值、方差和权重),从而对多峰不确定性进行建模,最终得到路径行程时间预测值。

(7) HierETA^[21]模型采用一种分层的自注意力网络(Hierarchical Self-Attention Network),通过并行编码器分别对路段内部的精细局部交通动态和跨路链(包含几个路段)与交叉口的宏观结构信息进行建模,最后通过一个层次感知的注意力解码器,动态融合这些多粒度特征,最终得到路径通行时间的预测值。

4.4 实验结果 本文模型和基准模型的结果比较如表2所示,其中,黑体表示结果最优。

由表可见,DATE模型框架在不确定性量化和点估计精度上均展现了最先进的性能。具体而言,在不确定性量化这一核心任务上,DATE(main)在综合指标MIS上表现最优,在两个数据集上均达到最低。这表明与ProbTTE和GMD-Net这样先进的概率模型相比,DATE(main)能更准确地捕获通行时间的不确定性,给出更加稳定可靠的置信区间。

在点估计方面,DATE-MC变体在两个数据集上均取得了MAE, RMSE和MAPE指标的全面最优,这证明了DATE模型的基础架构在特征提取和关系建模上具有卓越的能力,超越了所有基准模型。与DATE(main)相比,DATE-MC方法旨在捕获模型固有的不确定性,该方法将蒙特卡罗Dropout视为一种贝叶斯近似手段,通过对多次前向传播的结果求平均,虽然可以获得更精确的点估计,但是没有显式建模行程时间的上界和下界,导致区间过窄, PICP 过低,使MIS的结果较差。

在推理效率方面,DATE(main)在两个数据集上推理时间都最短,远优于HierETA等结构复杂的模型。值得注意的是,DATE-MC由于需要进行多次前向传播采样,其推理时间显著增加,这进一步凸显了DATE(main)在提供可靠不确定性量化的同时,兼具高效率的巨大优势。

总体而言,DATE在准确性、不确定性量化和计算效率上均具优势,展现了更强的实用价值。

4.5 消融实验 为了评估DATE中成分的有效性,我们比较了完整方法和两种变体的性能。

(1) 全局分布感知:去掉全局分布感知器模块,不去捕捉并建模路网中各路段间复杂的时空依赖关系,在进行行程时间估计和不确定性量化时,直接使用每个路段的历史信息作为输入。

(2) -LSTM:在分布融合不确定性估计模块中,进行行程时间估计和不确定性量化时,不使用LSTM来对轨迹中的路段序列进行处理,而是直接使用一个线性层进行处理。

这些变体在XA和CD数据集上的性能表现如表3所示,其中,表中黑体表示结果最优。由实验结果可见,去除全局分布感知器模块和LSTM模块都会导致模型性能的下降。具体而言,去除全局分布感知器模块后,模型无法有效捕捉路网中各路段间复杂的时空依赖关系,影响了行程时间估计和不确定性量化的精度;去除LSTM模块后,模型对轨迹的时序特征处理能力不足,导致预测准确性和不确定性量化效果下降。通过对比DATE的完整模型和去除模块后的变体,可以看出全局分布感知器模块在捕捉时空依赖关系方面的不可替代性,LSTM在处理轨迹的时序特征中起到了重要作用。

为了更深入地验证全局分布感知器模块的有效性和先进性,我们将其单独作为一个路段通行时间分布的预测器,与一个经典的时空图神经网络ASTGNN^[28]进行比较。ASTGNN通过在时间和空间维度上分别设计动态注意力机制来捕捉交通数据动态性,可以有效建模复杂的时空依赖关系。除了MAE, RMSE, MAPE等指标,我们还使用KL散度(Kullback-Leibler Divergence, KLD)作为评估指标,如式(28)所示:

$$D_{\text{KL}}(P|Q) = \sum_{k=1}^K P(k) \lg \left(\frac{P(k)}{Q(k)} \right) \quad (28)$$

其中, K 代表设定的离散区间的总数量。KLD用于衡量一个概率分布 Q (模型预测分布) 与另一个参考概率分布 P (真实分布) 之间的差异。这是一种非对称度量,值越小,表示两个分布越接近。在我们的任务中,其直接量化了模型预测的路段通行时间分布与真实统计分布的拟合程度,是评估分布预测质量的关键指标,

表2 各方法综合性能表现情况

Table 2 Overall performance comparison of all methods

数据集	方法	MIS (主)	Width	PICP	MAE	RMSE	MAPE	推理时间(s)
XA	HDC	7351.92	215.65	25.67%	265.77	354.99	26.35%	—
	GPD	1426.36	516.16	89.97%	131.69	188.96	14.62%	14.01
	ConSTGAT	1574.11	680.59	93.53%	152.42	219.76	15.46%	14.88
	ProbTTE	1366.15	475.43	87.16%	129.56	188.01	14.48%	12.51
	GMDNet	1468.74	608.95	90.47%	143.86	201.28	16.17%	12.31
	HierETA	1394.12	527.73	87.69%	129.95	189.26	14.60%	36.87
	DATE-MC	2307.32	235.89	59.45%	125.08	184.62	13.09%	150.17
	DATE (main)	1343.67	559.45	91.22%	126.34	187.16	13.67%	7.58
CD	HDC	8802.63	169.59	19.66%	292.76	373.69	31.80%	—
	GPD	1572.37	540.64	89.49%	140.11	204.50	15.48%	20.93
	ConSTGAT	1772.32	573.08	88.78%	161.62	239.45	16.30%	30.69
	ProbTTE	1429.42	573.55	92.15%	134.06	201.15	14.87%	27.83
	GMDNet	1509.27	526.90	89.05%	135.81	199.39	14.79%	27.78
	HierETA	1416.05	505.58	87.93%	134.07	195.57	14.58%	84.91
	DATE-MC	2667.14	223.77	55.31%	130.35	193.10	13.47%	280.21
	DATE (main)	1399.45	529.80	89.76%	131.52	204.51	13.98%	18.16

DATE中全局分布感知器模块与ASTGNN模型对路段通行时间分布预测的比较结果如表4所示,其中,表中黑体表示结果最优.由表可见,DATE全局分布感知器模块在所有评估指标上均显著优于ASTGNN.在分布预测质量上,其

KLD损失在两个数据集上分别降低了约42%和45%,表明DATE模块预测的分布与真实分布更接近.在基于分布均值的点估计精度上,其MAPE在两个数据集上也分别相对提升了约30%和23%.此外,得益于我们设计的板块化

表3 DATE的变体和完整模型的性能情况

Table 3 Performance comparison of DATE variants and the full model

数据集	方法	MIS (主)	Width	PICP	MAE	RMSE	MAPE
XA	—全局分布感知	1368.96	541.82	90.41%	130.66	194.32	13.98%
	—LSTM	1410.97	554.11	90.07%	132.75	196.64	14.17%
	DATE (full)	1343.67	559.45	91.22%	126.34	187.16	13.67%
CD	—全局分布感知	1422.84	550.24	91.27%	133.94	207.27	14.20%
	—LSTM	1479.23	544.21	88.59%	137.34	209.84	14.56%
	DATE (full)	1399.45	529.80	89.76%	131.52	204.51	13.98%

表4 DATE中全局分布感知器模块与ASTGNN模型对路段通行时间分布预测的比较

Table 4 Travel time distribution prediction: A comparison of DATE's global distribution-aware module and ASTGNN

数据集	方法	KLD	MAE	RMSE	MAPE	推理时间(s)
XA	ASTGNN	0.19	7.45	15.33	33.14%	7.53
	全局分布感知器	0.11	5.72	13.57	23.29%	2.51
CD	ASTGNN	0.22	7.12	13.46	31.70%	8.09
	全局分布感知器	0.12	6.08	12.61	24.29%	4.36

(Patching)和双重注意力机制,该模块的推理时间也远低于ASTGNN,效率提升了近2~3倍,表明了模型在大规模路网上高效应用的潜力.

图5进一步展示了全局分布感知器模块的预估效果,由图可见,该模块对所有路段所有时间片的通行时间分布的估计都是十分准确的.例如在11月5日第55个时间片(表示上午9:10—9:20)上,此时编号4465的路段交通情况比较复杂,有时还发生拥堵,但是全局分布感知器模块对这一

情况做出了相当准确的预测,说明其很好地捕捉到了交通的动态变化情况.通过合理的路网板块化和时空依赖建模,模型不仅有效地捕捉了局部交通模式,而且反映出整个路网的动态变化.与去除该模块后的模型相比,使用全局分布感知器模块的模型在准确性和不确定性量化上均有显著提升.结果表明,深度时空建模和全局分布感知的结合对提升行程时间预测的精度和可靠性具有重要作用.

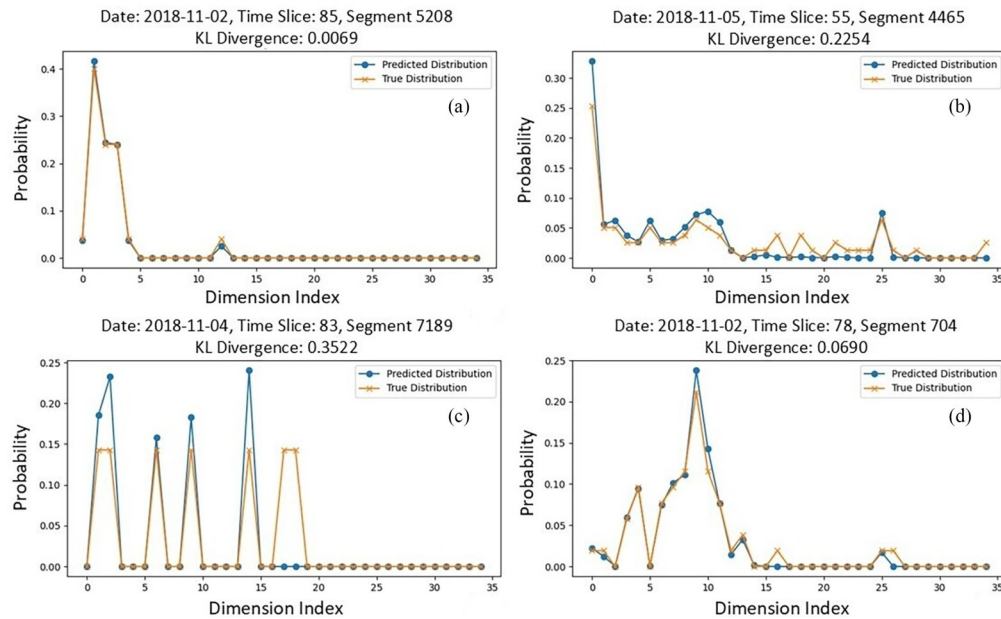


图5 全局分布感知器模块的预估结果: (a) 2018. 11. 02, 路段 5208 在时间片 85 的分布; (b) 2018. 11. 05, 路段 4465 在时间片 55 的分布; (c) 2018. 11. 04, 路段 7189 在时间片 83 的分布; (d) 2018. 11. 02, 路段 704 在时间片 78 的分布

Fig. 5 Estimation results of the global distribution-aware module: (a) 2018. 11. 02, distribution of segment 5208 at time slice 85; (b) 2018. 11. 05, distribution of segment 4465 at time slice 55; (c) 2018. 11. 04, distribution of segment 7189 at time slice 83; (d) 2018. 11. 02, distribution of segment 704 at time slice 78

5 结论

本文提出了一种面向大规模路网的行程时间预测和不确定性量化方法,旨在解决传统行程时间估计方法在面对动态复杂交通状况时的局限性.通过结合路网板块化、全局分布感知以及分布融合不确定性估计模块,本文提出的DATE模型能有效提高行程时间预测的精度,并提供可靠的预测置信区间,以实现不确定性的全面量化.在两个真实数据集上的大量实验结果表明了

DATE模型的有效性,充分验证了该方法在行程时间估计和不确定性量化中的可靠性和优势.

参考文献

[1] Yan C W, Johndrow J, Woodard D, et al. Efficiency of eta prediction. *SIAM Journal on Mathematics of Data Science*, 2024, 6(2):227—253.

[2] Chen X X, Cheng Z H, Sun L J. Bayesian inference for link travel time correlation of a bus route. *Transportmetrica B: Transport Dynamics*, 2024, 12(1):2416181.

- [3] Chandra S R, Al-Deek H. Predictions of freeway traffic speeds and volumes using vector autoregressive models. *Journal of Intelligent Transportation Systems*, 2009, 13(2):53–72.
- [4] Jin G Y, Liang Y X, Fang Y C, et al. Spatio-temporal graph neural networks for predictive learning in urban computing: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 2024, 36(10):5388–5408.
- [5] Kumar S V, Vanajakshi L. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review*, 2015, 7(3):21.
- [6] Lv Z J, Xu J J, Zheng K, et al. Lc-rnn: A deep learning model for traffic speed prediction// *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*. Stockholm: IJCAI, 2018:3470–3476.
- [7] Jiang X K, Zhuang D Y, Zhang X H, et al. Uncertainty quantification via spatial-temporal Tweedie model for zero-inflated and long-tail travel demand prediction// *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*. New York: Association for Computing Machinery, 2023:3983–3987.
- [8] Luo J, Zhang W, Fang Y, et al. Time Series supplier allocation via deep black-litterman model. *arXiv preprint 2024*, arXiv:2401.17350, .
- [9] Wang T F, Deng L W, Wang C, et al. COMET: NFT price prediction with wallet profiling// *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, 2024:5893–5904.
- [10] Guo S N, Lin Y F, Feng N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*. Honolulu: AAAI Press, 2019, 33(1):922–929.
- [11] Liu X, Xia Y, Liang Y, et al. Largest: A benchmark dataset for large-scale traffic forecasting. *Advances in Neural Information Processing Systems*, 2023, 36:75354–75371.
- [12] Cortes C, Vapnik V. Support-vector networks. *Machine Learning*, 1995, 20(3):273–297.
- [13] Quinlan J R. C4.5: Programs for machine learning. Amsterdam: Elsevier, 2014.
- [14] Breiman L. Random forests. *Machine Learning*, 2001, 45(1):5–32.
- [15] Awad M, Khanna R. Support vector regression// Awad M, Khanna R. CA: Efficient learning machines: theories, concepts, and applications for engineers and system designers. Berkeley: Apress, 2015:67–80.
- [16] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, 86(11):2278–2324.
- [17] Wang Z, Fu K, Ye J P. Learning to estimate the travel time// *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: Association for Computing Machinery, 2018:858–866.
- [18] Gers F A, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM. *Neural Computation*, 2000, 12(10):2451–2471.
- [19] 许淼. 基于深度学习的城市路网行程时间预测方法研究. 博士学位论文. 长春: 吉林大学, 2022.
- [20] Fang X M, Huang J Z, Wang F, et al. Constgat: contextual spatial-temporal graph attention network for travel time estimation at Baidu maps// *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: Association for Computing Machinery, 2020:2697–2705.
- [21] Chen Z B, Xiao X L, Gong Y J, et al. Interpreting trajectories from multiple views: A hierarchical self-attention network for estimating the time of arrival// *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, 2022:2771–2779.
- [22] Blundell C, Cornebise J, Kavukcuoglu K, et al. Weight uncertainty in neural networks// *Proceedings of the 32nd International Conference on International Conference on Machine Learning*. Lille: JMLR, 2015:1613–1622.
- [23] Derrow-Pinion A, She J, Wong D, et al. Eta prediction with graph neural networks in google maps// *The 30th ACM International Conference on Information and Knowledge Management*. New

- York: Association for Computing Machinery, 2021: 3767–3776.
- [24] Mao X W, Wan H Y, Wen H M, et al. GMDNet: A graph-based mixture density network for estimating packages' multimodal travel time distribution// Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(4):4561–4568.
- [25] Xu C, Wang Q, Sun L J. Link representation learning for probabilistic travel time estimation. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(11):21149–21161.
- [26] Xu C, Wang Q, Sun L G. SPTTEA: Spatiotemporal probabilistic framework for travel time estimation. (2024–11–27) [2025–05–28]. <https://doi.org/10.48550/arXiv.2411.18484>.
- [27] Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: representing model uncertainty in deep learning//Proceedings of the 33rd International Conference on International Conference on Machine Learning. New York: JMLR, 2016, 48:1050–1059.
- [28] Guo S N, Lin Y F, Wan H Y, et al. Learning dynamics and heterogeneity of spatial-temporal graph data for traffic forecasting. IEEE Transactions on Knowledge and Data Engineering, 2021, 34(11): 5415–5428.

(责任编辑 高善露)