

Document Version

Final published version

Licence

CC BY-NC

Citation (APA)

Li, A., Pan, Y., Xu, Z., Gao, B., Li, Y., Yu, H., & Chen, Y. (2026). 基于 Transformer 的智能网联车辆预测性运动规划. *Huanan Ligong Daxue Xuebao/Journal of South China University of Technology (Natural Science)*, 54(3), 52-64. <https://doi.org/10.12141/j.issn.1000-565X.250056>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership. Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

基于Transformer的智能网联车辆预测性运动规划

李安然^{1,2} 潘芋燕³ 徐震林⁴ 高博麟² 李永行¹ 于洪晟⁵ 陈艳艳¹

(1. 北京工业大学 城市交通学院, 北京 100124; 2. 清华大学 车辆与运载学院, 北京 100084; 3. 宾夕法尼亚州立大学 土木环境工程系, 宾夕法尼亚州 大学公园 16802; 4. 代尔夫特理工大学 土木工程和地球科学学院, 代尔夫特 2826 CN; 5. 中铁科学研究院集团有限公司 电子计算技术研究所, 北京 100081)

摘要: 智能网联车辆在复杂交通场景中的高效、安全运动规划是自动驾驶领域的关键挑战。该研究基于Transformer提出了ST-Trans交通预测模型, 并通过ST-Trans开发了用于智能网联车辆的预测性运动规划器。ST-Trans利用Transformer从动态高精度地图提供的实时车辆数据和车道段结构信息中挖掘交通时空演化规律来预测车道段的未来交通状态, 并利用车道段连通性和交叉口信号相位信息进一步提升预测准确性。模型采用编码器-解码器架构, 通过车道编码器融合车辆与车道特征, 道路编码器建模动态拓扑关系, 解码器迭代生成未来交通状态序列。预测结果显示ST-Trans在平均绝对误差、均方根误差和准确率上分别比最优基准线模型高出12.2%、12.1%和3.55个百分点。基于ST-Trans的预测结果, 预测性运动规划器采用双层结构, 底层路径规划器动态选择目标点并融合动态规划与二次规划生成平滑路径, 上层速度规划器构建时空走廊以压缩解空间, 并同样结合动态规划与二次规划生成安全高效的速度曲线, 从而显著降低运动规划任务的求解复杂度。该研究结合SUMO和CARLA对预测运动规划器进行了仿真实验, 结果表明, 基于ST-Trans的预测运动规划器能够实现预测性路径和速度规划, 并在安全性、效率、舒适性和计算速度方面优于传统运动规划器。实验验证了所提方法能有效缩短高风险状态持续时间, 提高通行效率, 并保持较低的计算延迟。

关键词: 动态高精度地图; 运动规划; 交通预测模型; 深度学习

中图分类号: U4; TP308

文章编号: 1000-565X(2026)03-0052-13

近年来, 通信、计算与传感技术的不断进步, 共同推动了动态高精度地图在智能网联车辆(ICV)运动规划中的广泛应用^[1]。与OpenDRIVE、Lanelet2及Apollo Map等现有地图格式类似, 动态高精度地图以专业测绘车辆采集的道路采样点为基础, 通过将其连接为封闭多边形, 构建具有三维坐标与语义特征的基本车道段单元^[2], 用于表征道路空

间。该表达方式实现了连续道路空间的离散化, 形成兼具精确几何特征与丰富语义信息的车道段, 从而为ICV的定位与感知提供底层支持。进一步地, 动态高精度地图融合来自网联车辆和智能路侧单元的动态交通数据, 并结合其内置的静态道路环境信息^[3], 借助高更新频率实现对时空离散化道路系统的动态映射, 最终构建出实时交通场景的数字孪生

收稿日期: 2025-03-05

基金项目: 国家自然科学基金青年项目(52402375); 交通部交通运输行业重点科技项目(2021-ZD2-047, 2022-ZD6-116); 北京市教委科技一般项目(KM202410005002)

Foundation items: Supported by the Young Scientists Fund of the National Natural Science Foundation of China (52402375) and the Key Science and Technology Program of the Transportation Industry, Ministry of Transport (2021-ZD2-047, 2022-ZD6-116)

作者简介: 李安然(1997—), 男, 博士生, 主要从事智能交通、自动驾驶研究。E-mail: lianran@emails.bjut.edu.cn

模型^[4]。基于车辆、车道、道路与交通控制等多源实时信息，动态高精度地图能够预测未来交通状态，从而为ICV提供合理的运动规划方案。

现有车辆运动规划方法通常可分为路径规划与轨迹规划两个阶段。路径规划旨在从道路采样点所构成的搜索空间中，寻得一条满足无碰撞、安全性、高效性与舒适性要求的几何路径^[5]。部分研究将路径规划建模为图优化问题，并采用Dijkstra、A*或D* Lite等图搜索算法生成路径^[6-7]。尽管这类方法能够在路径规划空间中获取最优路径，但其较高的计算复杂度限制了其在大范围规划任务中的应用。另一类研究采用概率采样方法进行路径规划，该方法无需显式构建路径规划空间，而是在道路采样点构成的近似空间中搜索无碰撞路径，通过压缩搜索空间以提高规划效率^[8-9]。然而，这类方法的采样过程随机性较强，导致所得解的一致性较差。还有研究将路径规划视为不确定性优化问题，并借助人工智能算法求解可行路径^[10-11]。这类方法虽然能够高效完成规划任务，但易陷入局部最优解，难以保证路径的全局最优性。在路径规划的基础上，轨迹规划进一步引入时间约束，生成描述位置与速度关系的时间序列，作为车辆可跟踪的时空参考轨迹^[12]。早期研究通常基于预设轨迹类型实现聚类规划策略，例如在结构化道路上采用直线-弧线曲线、螺旋曲线、多项式曲线、样条曲线或贝塞尔曲线等进行轨迹生成^[13-16]。然而，这些方法均存在一定局限性：直线-弧线曲线的曲率变化不连续；螺旋曲线的计算开销较大；多项式曲线中的关键参数不易确定；样条曲线难以保证最优性；而贝塞尔曲线的计算复杂度随阶数增加显著上升。另一类研究将轨迹规划建模为满足约束与性能指标的优化问题^[17-18]。尽管这类数值优化方法能够综合考虑多种约束条件，但仍存在计算耗时长、复杂度高的不足。近年来，随着人工智能技术的发展，基于深度学习的轨迹规划方法逐渐得到应用^[19-20]，但是其性能严重依赖训练数据，导致泛化能力较弱且可解释性有限。

针对上述问题，本研究提出一种基于动态高精度地图的车辆预测性运动规划方法。该方法首先基于Transformer架构构建时空交通预测模型(ST-Trans)，并利用其预测结果提升车辆运动规划在安全性、效率与舒适性方面的性能^[21]。在此基础上，本研究设计了一种双层运动规划架构，包括底层的路径规划器与上层的速度规划器。在运动规划过程

中，路径规划器依据高精度地图构建路径规划空间，根据ST-Trans的交通预测结果动态选择规划目标点，并结合动态规划与二次规划方法生成可行路径。随后，速度规划器在已生成路径曲线中引入时间约束，构建时空图模型，并依据高精度地图中的车道段结构及ST-Trans的预测结果对速度规划空间施加约束，以降低求解复杂度。速度规划器采用动态规划与二次规划方法，在时空图中搜索无碰撞且能够最小化目标函数的时空轨迹，作为最终可被车辆跟踪的速度规划结果。为验证所提交通预测模型与预测性运动规划方法的有效性，本研究基于SUMO与CARLA构建联合仿真平台，结合pNEUMA数据集所提供的真实轨迹数据与开放街道地图(OSM)，在虚拟交通场景中开展实验验证^[22]。

1 交通预测模型

1.1 问题描述

ST-Trans将动态高精度地图所提供的实时交通数据与道路环境信息编码为两个输入序列：交通状态序列 $X = \{X^t | t = 1, 2, \dots, T_h\}$ 和动态道路拓扑图 $G = \{G^t | t = 1, 2, \dots, T_h + T_p\}$ 。其中： T_h 表示输入序列的时间长度， T_p 表示预测序列的时间长度； $X^t = \{X_{\zeta}^t | \zeta = 1, 2, \dots, N_s\}$ ，表示场景中所有车道段在历史时段内各时刻的交通状态集合； N_s 为车道段总数。任意车道段 ζ 在时刻 t 的交通状态定义为 $X_{\zeta}^t = (V_{\zeta}^t, \theta_{\zeta}^t, s_{\zeta}^t)$ ，由该时刻位于车道段 ζ 内部的所有车辆状态序列 $V_{\zeta}^t = \{V_{\zeta,i}^t | i = 1, 2, \dots, N_{\zeta}^t\}$ 、车道中线方向 $\theta_{\zeta}^t \in [0, 2\pi]$ 及语义类型标签 $s_{\zeta}^t \in \mathbb{R}^d$ 共同构成。其中， $V_{\zeta,i}^t$ 表示 t 时刻车道段 ζ 中索引为 i 车辆的状态， N_{ζ}^t 表示该时刻车道段 ζ 内的车辆总数， s_{ζ}^t 为用于区分车道段类型的语义信息。车辆 i 在时刻 t 的状态定义为 $V_{\zeta,i}^t = (v_i^t, \theta_i^t, s_i^t)$ ，包括速度 $v_i^t \in \mathbb{R}$ 、航向角 $\theta_i^t \in [0, 2\pi]$ 和车辆语义类型信息 $s_i^t \in \mathbb{R}^d$ 。

ST-Trans将 t 时刻的道路拓扑结构表示为图 $G^t = (L, E^t, A^t)$ 。其中： L 为所有车道段的集合； E^t 表示由车道段间连通关系及该时刻信号交叉口相位信息共同确定的有向边集合； $A^t \in \mathbb{R}^{N_s \times N_s}$ ，表示拓扑图 G^t 的邻接矩阵； $D^t \in \mathbb{R}^{N_s \times N_s}$ ，为其度矩阵。邻接矩阵 A^t 和度矩阵 D^t 的计算公式如下：

$$A_{\zeta, \xi}^t = \begin{cases} 1, & \overrightarrow{\zeta\xi} \in E^t, \\ 0, & \text{其他,} \end{cases} \quad \zeta, \xi \in L \quad (1)$$

$$D_{\zeta, \zeta}^t = \sum_{\xi=1}^{N_b} A_{\zeta, \xi}^t \quad (2)$$

ST-Trans 的预测目标为所有车道段在预测时域内的交通状态序列 $Y = \{Y^t | t = T_h + 1, T_h + 2, \dots, T_h + T_p\}$ 。其中, $Y^t = \{N_{\zeta}^t | \zeta = 1, 2, \dots, N_b\}$, 表示 t 时刻所有车道段的交通状态。预测过程如下: 模型首先基于历史交通状态序列 X 与动态道路拓扑图 G 生成 $T_h + 1$ 时刻的交通状态预测 Y^{T_h+1} , 将该预测结果作为解码器输入的一部分, 进一步推断 $T_h + 2$ 时刻的交通状态 Y^{T_h+2} ; 依此类推, 迭代执行直至获得整个预测时段内所有时刻的交通状态预测, 最终输出完整的序列 $Y' = \{Y'^t | t = T_h + 1, T_h + 2, \dots, T_h + T_p\}$ 。

1.2 交通预测模型

1.2.1 模型架构

如图1所示, ST-Trans 由车道编码器、道路编码器和解码器3个核心模块构成。该模型首先利用车道编码器融合车辆特征与车道段特征, 提取每个车道段的嵌入表示。随后, 道路编码器将各车道段视为动态有向图中的节点, 依据交叉口信号相位信息构建节点之间的动态连接关系, 并通过注意力机制聚合历史时段中各车道段的编码信息, 生成交通状态的编码序列。最后, 解码器借助注意力机制对该编码序列进行解码, 提取隐含交通特征的序列, 并进一步映射为对未来交通状态的预测。ST-Trans 充分利用动态高精度地图所提供的实时交通与道路

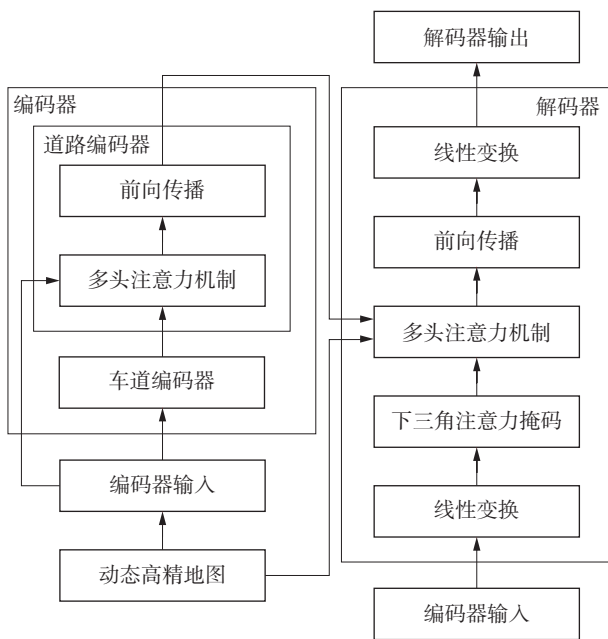


图1 ST-Trans 架构

Fig. 1 ST-Trans architecture

环境信息, 以高精度地图中的车道段为基本建模单元, 实现了对未来交通态势的精确预测。该模型的输出无需额外后处理, 即可直接用于构建车辆运动规划所需的时空走廊, 通过约束解空间提升运动规划的求解效率, 从而辅助目标车辆更安全、高效地通过目标路段。

1.2.2 车道编码器

在输入交通状态序列 X 和动态道路拓扑图 G 后, 车道编码器首先对各时刻的车辆速度特征进行正交分解, 并通过前馈神经网络(FFN)将其与车辆语义类型信息融合为车辆编码向量 $\epsilon_{\zeta, i}^t$ 。该过程的计算公式如下:

$$\epsilon_{\zeta, i}^t = \sigma \left(\mathbf{W}_{\text{veh}} \left[v_i^t \sin(\theta_i^t - \theta_{\zeta}^t), v_i^t \cos(\theta_i^t - \theta_{\zeta}^t), s_i^t \right] + b_{\text{veh}} \right) \quad (3)$$

其中, $\sigma(\cdot)$ 表示以 ReLU 为激活函数的前馈神经网络(FFN), \mathbf{W}_{veh} 表示线性变换的权重矩阵, b_{veh} 表示偏置项。该处理有助于强化横向速度突变特征, 捕捉潜在变道意图, 并区分不同类型车辆的运动特性, 从而提高交通预测的准确性。同时, 车道编码器使用另一个 FFN 将车道段特征转换为车道段特征向量 ϵ_{ζ}^t :

$$\epsilon_{\zeta}^t = \sigma \left(\mathbf{W}_{\text{lane}} \left[\theta_{\zeta}^t, s_{\zeta}^t \right] + b_{\text{lane}} \right) \quad (4)$$

其中, \mathbf{W}_{lane} 表示线性变换的权重矩阵, b_{lane} 表示对应的偏置项。随后, 车道编码器将 ϵ_{ζ}^t 线性变换为查询向量 q_{ζ}^t , 将 $\epsilon_{\zeta, i}^t$ 线性变换为键向量 $k_{\zeta, i}^t$ 和值向量 $v_{\zeta, i}^t$, 并通过注意力机制进行特征融合:

$$\text{head}_h = \text{Softmax} \left(\frac{q_{\zeta}^t k_{\zeta, i}^{tT}}{\sqrt{d_k}} \right) v_{\zeta, i}^t \quad (5)$$

$$z_{\zeta, i}^t = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h, \dots, \text{head}_{N_h}) \mathbf{W}^0 \quad (6)$$

其中, $z_{\zeta, i}^t$ 表示时刻 t 车道段 ζ 内车辆 i 的特征编码分量, head_h 表示多头注意力的 h 头索引, N_h 表示注意力头的总数, d_k 表示注意力向量维度, \mathbf{W}^0 是可学习的权重矩阵。最后, 车道编码器对该车道段内所有车辆的编码进行融合, 并再次通过 FFN 与 ϵ_{ζ}^t 进一步结合, 得到 t 时刻车道段 ζ 的车道编码向量 z_{ζ}^t 。该过程的计算公式如下:

$$z_{\zeta}^t = \sum_{i=1}^{N_{\zeta}^t} \mathbf{W}_z z_{\zeta, i}^t \quad (7)$$

$$z_{\zeta}^t = \sigma \left(\mathbf{W}_{\text{seg}} \left[\epsilon_{\zeta}^t, z_{\zeta}^t \right] + b_{\text{seg}} \right) \quad (8)$$

其中, N_{ζ}^t 表示 t 时刻位于车道段 ζ 内的车辆数目,

\mathbf{W}_z 和 \mathbf{W}_{seg} 表示线性变换的权重矩阵, b_{seg} 表示对应的偏置项。该机制使模型能够区分不同车道段的功能与特性差异, 理解交通负载与状态变化, 并学习车辆在相邻车道段之间的转移规律。

1.2.3 道路编码器

道路编码器首先对每一时刻的道路拓扑图 G^t 进行特征图分解, 其计算公式如下所示:

$$\mathbf{L}^t = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{A}^t \mathbf{D}^{-\frac{1}{2}} \quad (9)$$

$$\mathbf{L}^t = \mathbf{U}^t \mathbf{A}^t \mathbf{U}^{tT} \quad (10)$$

其中, \mathbf{I} 表示单位矩阵, \mathbf{L}^t 表示邻接矩阵 \mathbf{A}^t 的归一化拉普拉斯矩阵, \mathbf{U}^t 和 \mathbf{A}^t 分别表示特征向量矩阵和特征值矩阵。随后, 道路编码器通过3个独立的FFN将 \mathbf{U}^t 转化为时空嵌入矩阵 \mathbf{W}_q^t 、 \mathbf{W}_k^t 和 \mathbf{W}_v^t , 其计算公式如下所示:

$$\mathbf{W}_q^t, \mathbf{W}_k^t, \mathbf{W}_v^t = \sigma(\mathbf{W}_{road} \mathbf{U}^t + b_{road}) \quad (11)$$

其中, \mathbf{W}_{road} 表示线性变换的权重矩阵, b_{road} 表示对应的偏置项。接下来, 道路编码器将场景中所有车道段在 t 时刻的车道编码向量集合 $\mathbf{z}^t = \{\mathbf{z}_\zeta^t | \zeta = 1, 2, \dots, N_b\}$ 线性变换为查询向量 \mathbf{q}^t 、键向量 \mathbf{k}^t 和值向量 \mathbf{v}^t , 并在时空嵌入矩阵的指导下进行注意力计算:

$$\text{head}_h = \text{Softmax} \left(\frac{\mathbf{W}_q^t \mathbf{q}^t (\mathbf{W}_k^t \mathbf{k}^t)^T}{\sqrt{d_k}} \right) \mathbf{W}_v^t \mathbf{v}^t \quad (12)$$

$$\boldsymbol{\gamma}^t = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h, \dots, \text{head}_{N_b}) \mathbf{W}^0 \quad (13)$$

其中, $\boldsymbol{\gamma}^t$ 表示时刻 t 的交通状态编码向量。该模块将车道段视为图节点, 提取空间连接关系以表征道路结构, 并将信号相位信息映射为动态通行权关系, 从而学习动态拓扑对交通状态的影响。最终, 所有时刻的交通状态编码向量 $\boldsymbol{\gamma}^t$ 被合并为交通状态编码序列 $\boldsymbol{\gamma} = \{\boldsymbol{\gamma}^t | t = 1, 2, \dots, T_h\}$ 。

1.2.4 解码器

解码器根据初始输入 Y^{T_h} 和道路编码器生成的交通状态编码序列 $\boldsymbol{\gamma}$ 生成对预测时域内交通状态序列 Y 的预测序列 Y' 。与道路编码器类似, 解码器也会引入未来时刻的道路拓扑信息, 并采用掩码机制防止信息泄露。该过程的计算公式如下:

$$\text{head}_h = \text{Softmax} \left(\frac{\mathbf{W}_q^t \mathbf{q}^t (\mathbf{W}_k^t \mathbf{k}^t)^T}{\sqrt{d_k}} + \mathbf{M} \right) \mathbf{W}_v^t \mathbf{v}^t \quad (14)$$

$$\boldsymbol{\Psi} = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h, \dots, \text{head}_{N_b}) \mathbf{W}^0 \quad (15)$$

其中, $\boldsymbol{\Psi}$ 表示解码器初始输入的编码, \mathbf{M} 表示作为掩码的下三角矩阵。在此基础上, 解码器将 $\boldsymbol{\Psi}$ 线性变换为查询向量, 将 $\boldsymbol{\gamma}$ 线性变换为键向量和值向量, 采用多头注意力生成解码序列。最后, 解码器会对解码序列进行归一化生成对未来交通状态的预测 Y' 。

1.2.5 损失函数

为提升预测精度, 本研究采用平均绝对误差 (MAE) 作为损失函数对 ST-Trans 进行训练, 其计算公式如下:

$$\text{MAE}(Y, Y') = \frac{1}{T_p} \sum_{i=1}^{T_p} |Y^i - Y'^i| \quad (16)$$

在训练过程中, 模型通过梯度下降法不断减小 MAE, 以逐步提高预测准确性。

2 预测性运动规划模型

本研究提出了一种基于 ST-Trans 预测结果的运动规划模型, 该模型采用双层结构, 包括底层的路径规划器和上层的速度规划器。路径规划器以目标车辆当前位置为起点, 并依据 ST-Trans 输出的交通状态预测, 从车辆前方车道段中动态选取可最快通行的位置作为终点, 构建参考路径。在此基础上, 路径规划器以起点与终点构建 Frenet 坐标系, 并结合动态规划与二次规划方法, 生成连续平滑的最优路径。速度规划器则沿该路径构建时空图作为初始解空间, 并将 ST-Trans 的预测信息映射至该空间中, 标识出被占用区域和信号交叉口通行状态。随后, 在时空图中搜索一条能够避开所有障碍、并可快速通过交叉口的时空走廊。基于时空走廊, 速度规划器采用动态规划与二次规划算法, 生成一条平滑且安全的速度曲线。该预测性运动规划模型充分利用了 ST-Trans 的高精度交通预测能力, 能够快速确定路径终点, 并通过时空走廊有效约束解空间, 从而显著降低运动规划问题的计算复杂度, 提升整体求解效率。

2.1 底层路径规划器

如图2所示, 路径规划器以目标车辆的质心位置作为路径起点, 并依据 ST-Trans 的预测结果, 从前方第 N 列车道段中动态选择可最大化通行效率的车道段末端点作为终点, 据此生成连接起点与终点的参考线。该方法可高效确定路径规划的终点位置, 缩小解空间规模, 从而显著提高计算效率。在此基础上, 路径规划器以自车位置为原点建立

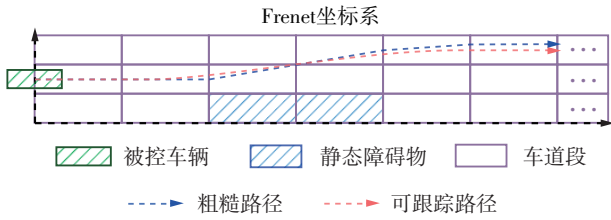


图2 路径规划过程

Fig. 2 Path planning process

Frenet坐标系^[23], 将高精度地图中的采样点映射至该坐标系中作为路径参考点, 同时将静态障碍物映射至同一坐标系, 并将其所占用的车道段从解空间中排除, 进一步压缩可行解范围。随后, 路径规划器采用动态规划算法在解空间中搜索满足安全性、平滑性与道路边界约束的无碰撞路径曲线^[24-25]。路径规划器的总目标 J_{DP} 的函数表达式定义如下:

$$J_{DP} = w_{smo} J_{smo} + w_{safe} J_{safe} + w_{ref} J_{ref} \quad (17)$$

其中, J_{smo} 、 J_{safe} 、 J_{ref} 分别表示路径的平滑性目标、安全性目标和车道保持目标, w_{smo} 、 w_{safe} 、 w_{ref} 为对应的权重系数。

为提升路径的平滑性, 路径规划器将平滑目标 J_{smo} 的函数表达式定义如下:

$$J_{smo} = w_c \sum \kappa(s)^2 + w_{cc} \sum \kappa'(s)^2 \quad (18)$$

其中, s 表示路径规划点序列, $\kappa(s)$ 和 $\kappa'(s)$ 分别表示路径规划的曲率及其变化率, w_c 和 w_{cc} 分别表示对应的权重系数。

为规避静态障碍物, 路径规划器将安全目标 J_{safe} 的函数表达式定义如下:

$$J_{safe} = \sum g_{safe}(\tau) \quad (19)$$

$$g_{safe}(\tau) = \begin{cases} 0, & \tau > \tau_{safe} \\ k(\tau_{safe} - \tau), & \tau_{risk} \leq \tau \leq \tau_{safe} \\ C_{max}, & \tau < \tau_{risk} \end{cases} \quad (20)$$

其中, τ 表示目标车辆与包含静态障碍物的车道段的间距, τ_{safe} 表示安全间距阈值, τ_{risk} 表示高风险间距阈值, k 表示比例系数, C_{max} 表示一个极大的常数惩罚项。

为将车辆行驶限制在道路边界内, 路径规划器将车道保持目标 J_{ref} 的函数表达式定义如下:

$$J_{ref} = \sum w_{lane} (l - l_{ref})^2 \quad (21)$$

其中, l 表示目标车辆的横向位置, l_{ref} 表示参考线的横向位置, w_{lane} 表示权重参数。

路径规划器首先通过动态规划算法初步求解最小化总目标函数的路径曲线, 再采用二次规划方法对该结果进行优化, 生成可供车辆跟踪的最终路

径^[26]。二次规划过程需同时优化路径平滑性与对参考线的跟随性能, 其目标函数 J_{QP} 的计算公式如下所示:

$$J_{QP} = w_1 \sum (l - l_{ref})^2 + w_2 \sum l'^2 + w_3 \sum l''^2 + w_4 \sum l'''^2 \quad (22)$$

其中: $\sum (l - l_{ref})^2$ 的最小化用于减小目标车辆对路径曲线进行跟踪的横向偏移; $\sum l'^2$ 、 $\sum l''^2$ 、 $\sum l'''^2$ 的最小化则分别用于抑制横向速度、加速度和加加速度, 以优化路径曲线的平顺性; w_1 、 w_2 、 w_3 、 w_4 分别表示对应的权重系数。此外, 路径曲线在优化过程中还需要满足起始状态约束、路径连续性约束及各阶导数约束。

2.2 上层速度规划器

速度规划器基于路径规划器所生成的路径曲线, 构建以 Δs 为空间间隔、 Δt 为时间间隔的等距时空图, 作为速度规划的初始解空间。为压缩解空间并提高计算效率, 速度规划器将高精度地图中的车道段信息与ST-Trans的交通预测结果投影至该解空间中, 进而构造出连续的时空走廊, 作为速度规划的可行边界。如图3(a)所示, 时空走廊有效避开各时刻所有被占据的车道段, 显著缩小了速度规划的解空间, 提高了速度规划的求解效率。在此基础上, 速度规划器对目标车辆周围的环境车辆进行短时轨迹预测^[27], 并对与目标车辆路径存在冲突的环境车辆轨迹进行膨胀处理。如图3(b)所示, 膨胀后的环境车辆轨迹被映射至时空走廊中, 冲突区域被排除出解空间, 从而进一步缩减求解范围并增强速度曲线的安全性。随后, 速度规划器采用动态规划与二次规划算法, 在可行解空间中搜索最优速度曲线^[28]。速度规划模块的总目标函数 J_{dp} 的表达式如下所示:

$$J_{dp} = c_{smo} J_{smo} + c_{safe} J_{safe} + c_{eff} J_{eff} \quad (23)$$

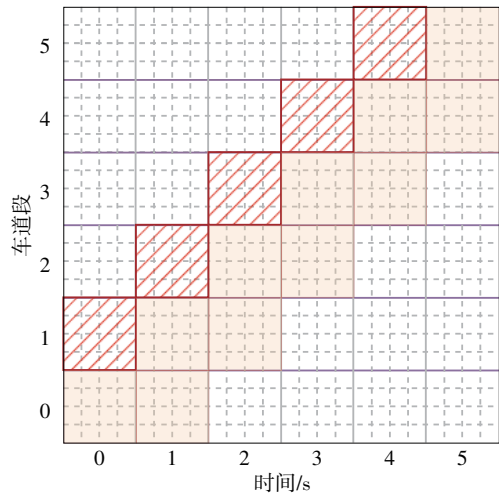
其中, J_{smo} 、 J_{safe} 、 J_{eff} 分别表示速度规划曲线的平滑目标、安全目标和通行效率目标, c_{smo} 、 c_{safe} 、 c_{eff} 表示对应的权重系数。

为保证速度曲线的平滑性, 速度规划器将平滑目标 J_{smo} 的函数表达式定义如下:

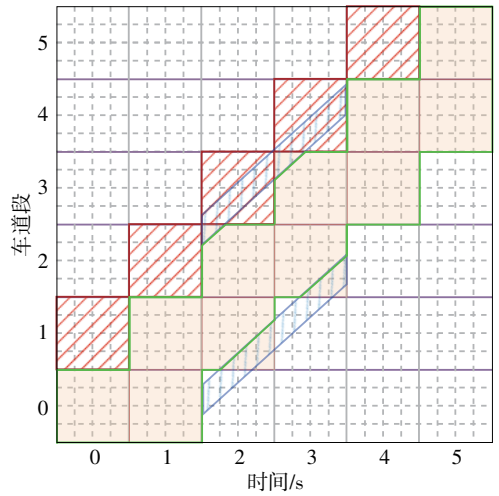
$$J_{smo} = c_{vel} \sum \dot{s}_i^2 + c_{acc} \sum \ddot{s}_i^2 + c_{jerk} \sum \dddot{s}_i^2 \quad (24)$$

其中, \dot{s}_i 、 \ddot{s}_i 、 \dddot{s}_i 代表离散化后的速度、加速度和跃度, c_{vel} 、 c_{acc} 、 c_{jerk} 分别表示速度项、加速度项和跃度项的权重系数。

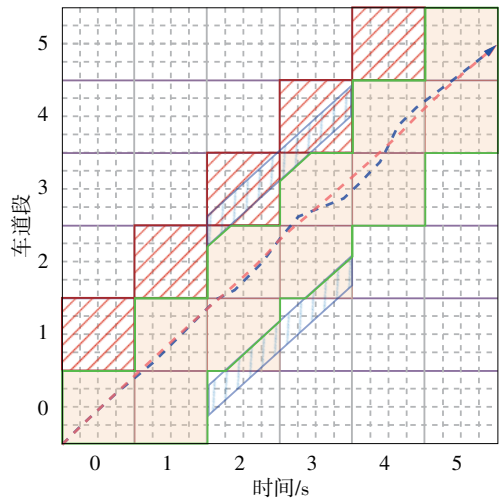
为避免速度规划曲线与动态障碍物发生冲突, 速度规划器继续采用式(19)和(20), 将安全目标 J_{safe}



紫色直线表示车道段边缘线，红色阴影表示被占据车道段，棕色阴影表示时空走廊
(a) 构建速度规划时空图



蓝色阴影表示环境车辆轨迹预测，绿色线条所围空间表示速度规划求解边界
(b) 构建速度规划解空间



蓝色曲线表示动态规划的粗糙速度曲线，红色曲线表示二次规划的可跟踪速度曲线
(c) 求解速度规划

图3 速度规划过程

Fig. 3 Speed planning process

定义为分段惩罚函数 $g_{\text{safe}}(\cdot)$ 。

为使目标车辆尽快通过规划路段，速度规划器将通行效率目标 J_{eff} 的函数表达式定义如下：

$$J_{\text{eff}} = \sum (\dot{s}_i - \dot{s}_{\text{ref},i})^2 \quad (25)$$

其中， $\dot{s}_{\text{ref},i}$ 表示由驾驶任务需求或道路限速确定的参考速度。通过最小化通行效率目标 J_{eff} ，速度规划器能够促使目标车辆尽快达到并维持参考速度 $\dot{s}_{\text{ref},i}$ ，从而提高目标车辆的通行效率。

如图3(c)所示，速度规划器采用二次规划算法对动态规划输出的初始速度曲线进行进一步优化。优化过程中需综合考虑平滑性约束、安全性约束、车辆运动学约束以及速度曲线连续性约束。其目标函数 J_{qp} 的表达式如下所示：

$$J_{\text{qp}} = \sum (c_s (s_i^* - s_i)^2 + c_v (\dot{s}_i^* - \dot{s}_i)^2 + c_a \ddot{s}_i^{*2} + c_j \ddot{s}_i^{*2}) \quad (26)$$

其中， s_i^* 、 \dot{s}_i^* 、 \ddot{s}_i^* 、 \ddot{s}_i^* 代表二次规划化后的纵向位置、速度、加速度和跃度， c_s 、 c_v 、 c_a 、 c_j 分别表示对应的权重系数。此外，速度曲线在优化过程中还需满足起始状态约束、车辆动力学约束及乘客舒适性约束。

3 实验与分析

3.1 交通预测模型性能分析

3.1.1 实验数据集

本研究基于 pNEUMA 开源数据集的车辆轨迹数据与 OSM 道路数据构建虚拟交通场景^[29]。pNEUMA 数据集通过无人机群于雅典中央商务区采集，覆盖约 1.4 km²、包含数 10 个交叉口的区域，在 10 月 24 日至 11 月 1 日期间 5 个工作日的早高峰时段(当地时间 08:30 至 10:30)记录了以 25 Hz 采样频率获取的近 500 000 辆独立车辆的高精度 GPS 轨迹。本研究首先从 OSM 获取对应区域的路网信息，并借助地图转换工具将其转换为 OpenDRIVE 格式。随后对原始轨迹进行预处理：采用无迹卡尔曼滤波对含噪声轨迹进行平滑处理，并将处理后的轨迹精确映射至 OpenDRIVE 路网坐标系中，从而生成连续可靠的车道级轨迹序列。为平衡数据规模与计算效率，将所有轨迹数据统一降采样至 1 Hz，作为最终虚拟场景的输入。基于上述虚拟场景，本研究选取了 3 个信号控制交叉口区域，构建了一个专门用于交通行为建模的数据集。该数据集共包含 323 557 个样本，按 7:1:2 的比例划分为训练集、验证集和测试集，对应样本数量分别为 205 942、

39472和78143。

3.1.2 实验参数设置

ST-Trans 部署于配备有 Intel Xeon Gold 5218CPU 及 NVIDIA Tesla V100S GPU 的刀片服务器上, 运行于 Ubuntu 22.04.2 LTS 操作系统。该模型在 Python 3.8 环境下基于 PyTorch 1.5 框架实现, 训练阶段采用 Adam 优化器, 其学习率设为 1×10^{-3} , 批量大小为 32, 一阶矩估计衰减率为 0.9, 二阶矩估计衰减率为 0.99, 数值稳定性常数为 1×10^{-8} 。

现有研究表明, 引入车道类型与车型属性信息有助于提高交通预测精度^[22]。因此, 本研究参考既有分类体系, 将车道段划分为上游、直行、左转和右转 4 种类型; 同时, 基于 pNEUMA 数据集提供的车辆类型信息, 将交通参与者分为公交车、小客车、重型车、中型车、摩托车和出租车共 6 类。根据超参数分析结果, 模型关键参数设置如下: 输入序列时间维度 T_h 为 7, 预测序列时间维度 T_p 为 20, 隐藏层维度 d_h 为 64, 多头注意力头数 N_h 为 8, 并指定了相应的多头注意力维度 d_h/N_h 。

3.1.3 实验评价指标

本研究采用 MAE、均方根误差(RMSE)和准确率(ACC)作为模型性能的评估指标。其中, RMSE 的计算公式如下:

$$\text{RMSE}(Y, Y') = \sqrt{\frac{1}{T_p} \sum_{t=1}^{T_p} (Y^t - Y'^t)^2} \quad (27)$$

ACC 的判定准则为: 若某车道段的预测值低于 0.5, 则判定其为自由状态; 若高于 0.5, 则视为占用状态。本研究最终根据被正确分类的车道段数量占总数的比例计算 ACC。

3.1.4 预测模型性能分析

为评估 ST-Trans 的性能, 本研究将其与 6 种基线模型(自回归积分滑动平均模型(ARIMA)、随机森林(RF)、支持向量回归(SVR)、FNN、自适应图卷积循环网络(AGCRN)和图波网络(GWN))在预测准确性方面进行了比较^[30-35]。如表 1 所示, ST-Trans 在 MAE、RMSE 和 ACC 这 3 项指标上显著优于所有基线模型。与表现最佳的基线模型 GWN 相比, ST-Trans 的 MAE 和 RMSE 分别降低了 12.2% 和 12.1%, 同时 ACC 提高了 3.55 个百分点。从模型演进的角度来看, 传统时间序列模型(ARIMA、RF 和 SVR)由于难以有效捕捉时空依赖性, 其 MAE、RMSE 和 ACC 指标均显著落后; 浅层神经网络 FNN 通过特征学习将 ACC 提升至 84.77%; 而时空图模型

表 1 ST-Trans 与基准线模型的性能比较

Table 1 Performance comparison between ST-Trans and baseline models

模型	MAE	RMSE	ACC/%
ARIMA	0.509	1.007	69.91
RF	0.417	0.815	78.95
SVR	0.434	0.849	77.81
FNN	0.337	0.664	84.77
AGCRN	0.316	0.624	86.83
GWN	0.303	0.593	88.43
ST-Trans	0.266	0.521	91.98

AGCRN 和 GWN 进一步利用路网拓扑信息, 将 ACC 提高至 86% 以上。相比之下, ST-Trans 通过将车辆类型与单元类型信息嵌入车道编码, 来区分不同车辆与车道的属性, 从而提升预测准确性。此外, 该模型能够基于动态道路结构识别交通信号状态变化, 并学习不同信号阶段中交通流的演化规律, 进一步增强预测精度。

3.1.5 预测模型功能模块分析

为评估不同功能模块对 ST-Trans 性能的贡献, 本研究采用消融实验依次从完整模型中移除各功能模块, 并分析其对预测准确性的影响。所有变体模型均在与完整 ST-Trans 相同的实验设置和测试集下进行评估, 具体包括以下变体模型: ST-Trans_{noVT}, 移除车道编码器中的车型嵌入模块; ST-Trans_{noVA}, 移除车辆航向角嵌入模块; ST-Trans_{noLT}, 移除车道段类型嵌入模块; ST-Trans_{noLD}, 移除车道段方向嵌入模块; ST-Trans_{noST}, 移除道路编码器中的时空嵌入矩阵模块。

消融实验结果如表 2 所示, 移除任意模块均导致模型性能显著下降, 表明各模块均对预测准确性具有重要贡献。其中, 时空嵌入矩阵模块的缺失对性能影响最为显著, ACC 从 91.98% 降至 88.34%, 说明该模块在动态交通信号状态建模中起核心作用。车型嵌入模块和车道段类型嵌入模块的移除对性能产生较大影响: 缺少车型嵌入信息时 ACC 降至 90.02%, 缺少车道段类型信息时 ACC 降至 90.19%, 表明区分交通实体语义特征对模型性能至关重要。车辆航向角嵌入模块和车道段方向嵌入模块的影响相对较小, ACC 分别降至 90.66% 和 91.00%, 仍较完整模型下降 1.32 个百分点和 0.98 个百分点, 说明微观运动方向信息对精度提升具有辅助作用。在全部变体模型中, 仅移除车道方向模块时性能损失最小, 而时空嵌入矩阵缺失所造成的影响最为严

表2 不同功能模块对ST-Trans性能的影响

Table 2 Impacts of diverse function modules to the performance of ST-Trans

模型	MAE	RMSE	ACC/%
ST-Trans _{noVT}	0.284	0.561	90.02
ST-Trans _{noVA}	0.274	0.539	90.66
ST-Trans _{noLT}	0.280	0.563	90.19
ST-Trans _{noLD}	0.271	0.536	91.00
ST-Trans _{noST}	0.314	0.627	88.34
ST-Trans	0.266	0.521	91.98

重，凸显了建模道路动态拓扑结构在交通预测中的关键作用。

3.1.6 预测模型超参数分析

为优化ST-Trans在交通预测中的性能，本研究系统评估了多个关键超参数对其预测准确性的影响，包括输入序列时间长度 T_h 、预测序列时间长度 T_p 、隐藏层维度 d_h 以及多头注意力头数 N_h 。如表3所示，增大 T_h 可显著提升模型的预测准确性，但其改善效果呈现明显的边际递减趋势。当 T_h 增加至7时，性能提升幅度显著减缓；继续增大至8时，性能增益已趋于忽略不计。因此，本研究将 T_h 设定为7，以兼顾预测精度与计算效率。在预测时域 T_p 方面，ST-Trans在 $T_p=10$ 时达到最佳预测性能。随着 T_p 进一步延长，模型误差迅速增大，表明时间累积效应加剧，因此将 T_p 确定为10。隐藏层维度 d_h 的增加有助于增强特征表征能力，从而提升预测性能，但同样伴随边际效益递减现象。鉴于 $d_h=64$ 与 $d_h=128$ 之间的性能差异有限，本研究选择 $d_h=64$ ，以在准确性与模型效率之间取得更好的平衡。多头注意力头数 N_h 的实验显示，当 N_h 为8时模型性能最优，减少或增加头数均会导致性能下降，因此将 N_h 最终设置为8。

3.1.7 预测模型训练分析

如图4所示，学习率对ST-Trans的训练收敛过程具有显著影响。当学习率设置过高(例如0.1000或0.0100)时，训练损失曲线出现明显振荡，表明参数更新步长过大，导致优化过程反复越过最优解区域，难以稳定收敛，最终影响模型性能。相反，学习率设置过低(如0.0001)会导致损失值下降缓慢，收敛速度显著降低，不仅导致训练效率下降和计算资源消耗增加，也更容易陷入局部最优解。相比之下，当学习率为0.0010时，ST-Trans表现出最佳的收敛行为：损失值在训练初期迅速下降，随后平稳收敛至较低水平。该学习率

表3 不同超参数对ST-Trans性能的影响

Table 3 Impacts of diverse hyperparameters on the performance of ST-Trans

超参数	MAE	RMSE	ACC/%
$T_h=2$	0.404	0.791	80.61
$T_h=3$	0.353	0.698	85.79
$T_h=4$	0.309	0.605	88.84
$T_h=5$	0.286	0.566	90.29
$T_h=6$	0.272	0.538	91.16
$T_h=7$	0.266	0.521	91.98
$T_h=8$	0.264	0.517	92.25
$T_p=10$	0.248	0.497	92.96
$T_p=15$	0.254	0.508	92.50
$T_p=20$	0.266	0.521	91.98
$T_p=25$	0.289	0.568	89.98
$d_h=16$	0.274	0.537	90.81
$d_h=32$	0.269	0.529	91.62
$d_h=64$	0.266	0.521	91.98
$d_h=128$	0.264	0.518	92.11
$N_h=2$	0.279	0.560	90.34
$N_h=4$	0.268	0.528	91.69
$N_h=8$	0.266	0.521	91.98
$N_h=16$	0.268	0.526	91.77

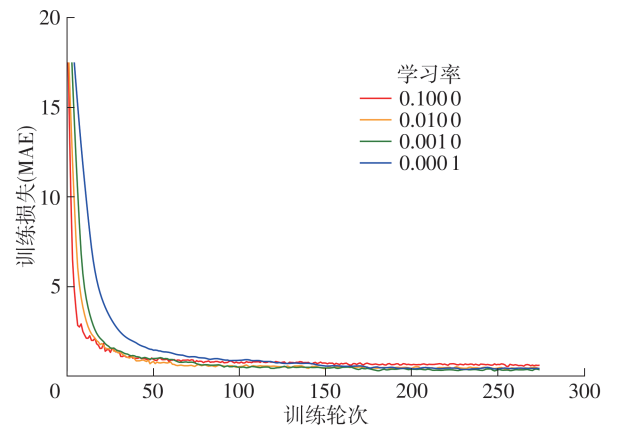


图4 不同学习率下ST-Trans的收敛曲线

Fig. 4 Convergence curves of ST-Trans at different learning rates

在收敛速度与稳定性之间取得了良好平衡，既有效避免了训练过程中的振荡现象，也保证了优化效率，因此被确定为ST-Trans在交通预测任务中的较优选择。

3.2 预测性运动规划器性能分析

3.2.1 联合仿真平台

本研究基于SUMO与CARLA构建了联合仿真平台，并结合pNEUMA车辆轨迹数据及其对应的OSM地图数据，建立了信号交叉口虚拟交通环境，

以评估基于ST-Trans的预测性运动规划器的性能^[36-37]。平台构建过程包括以下几个步骤:首先,基于pNEUMA数据集关联的OSM地图,采用SUMO的netconvert插件生成基础路网拓扑,并依据Shapefile数据与交通规则对其进行语义校正;随后,将修正后的路网通过netconvert命令转换为OpenDRIVE格式,并导入CARLA中完成交通场景的几何构建。在背景交通流生成方面,研究采用无迹卡尔曼滤波对pNEUMA原始轨迹数据进行平滑处理,将处理后的轨迹通过路网匹配算法映射至SUMO坐标系,并借助TraCI接口实时注入SUMO仿真环境,作为背景车辆流。

如图5所示,所构建的虚拟仿真场景为单向双车道主干道与单车道支路相交的信号控制交叉

口,其中主干道与支路的信号周期分别为25s与20s。在仿真初始化阶段,平台基于pNEUMA轨迹数据加载仿真开始时位于交叉口上游100m至下游50m范围内的所有车辆,总车辆数不超过30辆。仿真开始2s后停止生成环境车辆,并在主干道上游100m处的右侧车道生成代表目标车辆的小汽车模型。该平台通过桥接器实现SUMO与CARLA之间的双向同步:SUMO负责背景交通流的微观仿真,并实时传输背景车辆位姿与信号相位至CARLA;CARLA则负责执行控制指令,驱动目标车辆并将其状态反馈回SUMO,从而构成闭环验证系统。同时,利用CARLA传感器系统采集目标车辆的运动数据,用于后续对ST-Trans规划器的性能评估。

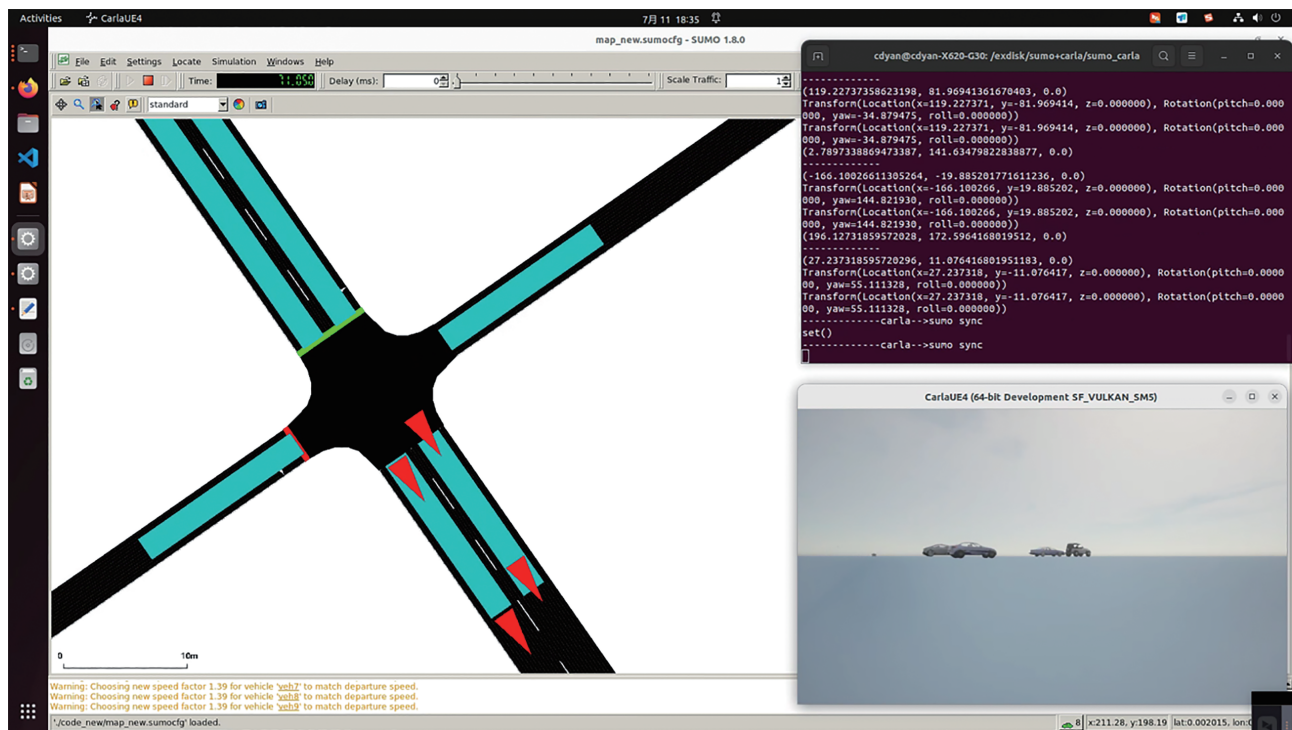


图5 SUMO-CARLA联合仿真平台

Fig. 5 SUMO-CARLA joint simulation platform

3.2.2 预测性运动规划器定性分析

在本研究构建的虚拟交通场景中,对ST-Trans规划器与百度Apollo提出的期望最大化(EM)规划器进行了对比实验^[38]。如图6所示,目标车辆进入实验场景后,EM规划器和ST-Trans规划器均控制车辆减速驶向信号交叉口,但EM规划器所产生的运动规划呈现出更频繁的加减速振荡。相比之下,ST-Trans规划器生成的速度曲线更为平缓,且穿越整个场景所需时间更短。实验结果表明,ST-Trans规划器在通行效率与舒适性方面均优于EM规划器。

3.2.3 预测性运动规划器定量分析

为评估ST-Trans规划器的性能,本研究基于pNEUMA数据集在联合仿真平台中重构了约40个信号交叉口交通场景。以40辆/(km·ln)作为密度阈值,将交通状态划分为低密度与高密度两种交叉口场景,并依据车辆行驶行为将车辆分为直行、左转和右转3类分别进行仿真分析。本研究对EM规划器和ST-Trans规划器分别开展测试,并对两者的运动规划效果进行了综合比较。参考现有研究中采用的碰撞时间阈值,本研究将目标车辆的驾驶状态划

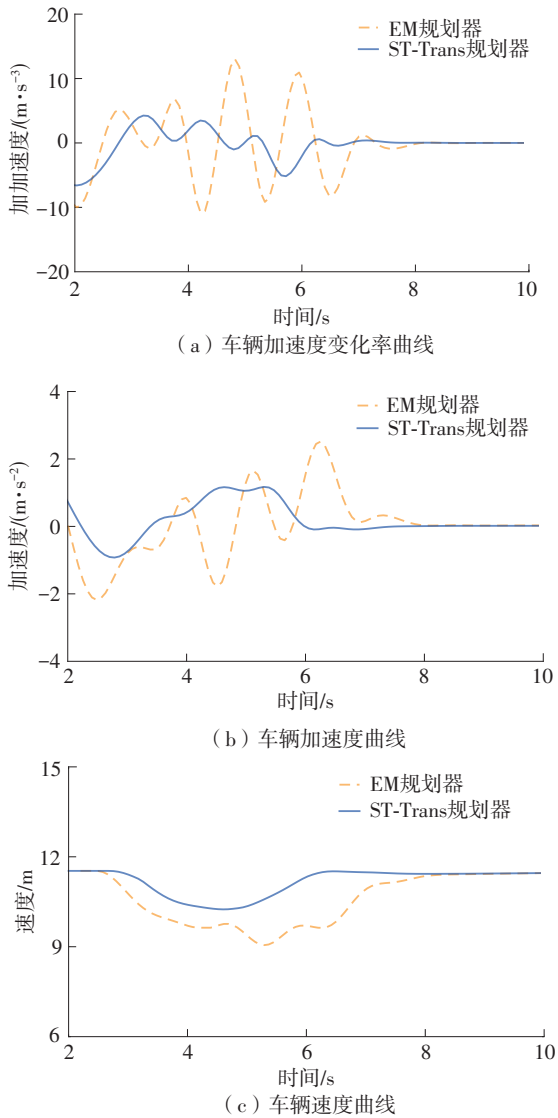


图6 车辆运动参数曲线

Fig. 6 Curves of vehicle motion parameters

分为3类^[39]：高风险状态(碰撞时间 $\leq 1.6\text{s}$)、风险预警状态(碰撞时间介于 1.6s 至 3.0s 之间)和安全状态(碰撞时间 $\geq 3.0\text{s}$)。通过统计目标车辆处于各状态的持续时间，以量化不同规划器在安全性方面的差异。同时，采用平均通行时间和单轮规划的平均计算延误分别评估通行效率与计算效率。

如表4所示，在低密度场景中，ST-Trans规划器在维持较低计算延迟的同时，显著缩短了高风险驾驶状态的持续时间；在高密度场景中，该规划器仍能有效控制冲突风险，整体性能优于EM规划器。实验结果表明，ST-Trans规划器在不同交通密度条件下均表现出良好的适应性与稳定性，在跟驰场景中能够保持更高安全性，其在通行效率与实时性方面的表现也总体优于对比方法。

4 结语

本研究提出了一种基于Transformer架构的交通预测模型ST-Trans，并以此为基础构建了一种预测性双层运动规划器。ST-Trans利用动态高精度地图所提供的车辆与车道段信息，对未来交通状态进行预测；通过将车道段建模为有向图中的节点，并将信号交叉口的信号相位变化表征为特定节点间的动态连接关系，有效提升了预测准确性。

基于ST-Trans的预测输出，本研究进一步设计了一种分层运动规划器，其结构包括底层路径规划层与上层速度规划层。路径规划器依据ST-Trans的预测结果动态选定路径终点，结合高精度地图提供的定位信息与采样点构建路径规划空间，并采用动态规划算法和二次规划算法生成可行路径。速度规

表4 EM规划器与ST-Trans规划器的性能比较

Table 4 Performance comparison between EM planner and ST-Trans planner

场景	行为	模型	碰撞时间			平均通行时长	平均计算延误
			高风险状态	风险预警状态	安全状态		
低密度	直行	EM规划器	1.76	3.27	6.12	11.15	0.087
		ST-Trans规划器	0.65	6.31	3.07	10.03	0.069
	左转	EM规划器	2.55	3.13	10.95	16.63	0.102
		ST-Trans规划器	0.98	9.65	4.53	15.16	0.075
	右转	EM规划器	2.29	3.57	7.60	13.46	0.095
		ST-Trans规划器	1.14	7.34	3.59	12.07	0.073
高密度	直行	EM规划器	2.52	4.33	8.42	15.27	0.094
		ST-Trans规划器	1.04	8.79	4.23	14.06	0.070
	左转	EM规划器	3.95	4.53	14.86	23.34	0.106
		ST-Trans规划器	1.98	13.14	6.23	21.35	0.077
	右转	EM规划器	3.75	4.91	10.46	19.12	0.102
		ST-Trans规划器	1.92	10.09	5.53	17.54	0.076

划器则基于轨迹规划结果建立时空图,借助高精度地图中的车道段结构及ST-Trans的交通预测输出构建时空走廊,作为速度规划的求解空间,从而显著降低优化问题的复杂度。同时,该层还将周围车辆的轨迹预测结果进行膨胀处理,并映射至时空走廊中,以增强运动规划的安全性。速度规划采用动态规划算法求解最优速度曲线,并通过二次规划算法进行平滑处理,最终生成可被目标车辆稳定跟踪的时空曲线。

为验证ST-Trans及所提出预测性运动规划器的性能,本研究搭建了SUMO与CARLA联合仿真实验平台,并基于pNEUMA数据集中的真实车辆轨迹与OSM地图数据构建了虚拟交通场景。测试结果表明,ST-Trans能够准确预测未来交通状态,所提出的规划器在运动规划的安全性、通行效率、舒适性及计算效率等方面均表现出优良性能。未来研究将聚焦于更高效地利用高精度地图车道段信息与交通预测结果以构建时空走廊,并进一步提升ST-Trans规划器在复杂交通场景中的综合性能。

参考文献:

- [1] GAO B, WAN K, CHEN Q, et al. A review and outlook on predictive cruise control of vehicles and typical applications under cloud control system [J]. *Machine Intelligence Research*, 2023, 20(5): 614-639.
- [2] BAO Z, HOSSAIN S, LANG H, et al. High-definition map generation technologies for autonomous driving [EB/OL]. (2022-06-24) [2025-11-10]. arXiv preprint, 2022: arXiv: 2206.05400.
- [3] DONG J, XU Q, WANG H, et al. Mixed cloud control testbed: validating vehicle-road-cloud integration via mixed digital twin [J]. *IEEE Transactions on Intelligent Vehicles*, 2023, 8(4): 2723-2736.
- [4] 冯昶, 杜清运, 范晓宇, 等. 高精动态地图基础平台众源更新技术路线研究 [J]. *测绘地理信息*, 2023, 48(1): 10-15.
FENG Chang, DU Qingyun, FAN Xiaoyu, et al. A crowdsourcing update technology route of HD dynamic map basic platform [J]. *Journal of Geomatics*, 2023, 48(1): 10-15.
- [5] SUNG I, CHOI B, NIELSEN P. On the training of a neural network for online path planning with offline path planning algorithms [J]. *International Journal of Information Management*, 2021, 57: 102142/1-9.
- [6] NOTO M, SATO H. A method for the shortest path search by extended Dijkstra algorithm [C]//*Proceedings of the 2000 IEEE International Conference on Systems, Man and Cybernetics*. Nashville: IEEE, 2000: 8-11.
- [7] ZHANG Z, ZHAO Z. A multiple mobile robots path planning algorithm based on A-star and Dijkstra algorithm [J]. *International Journal of Smart Home*, 2014, 8(3): 75-86.
- [8] SCHMERLING E, JANSON L, PAVONE M. Optimal sampling-based motion planning under differential constraints: the drift case with linear affine dynamics [C]//*Proceedings of the 54th IEEE Conference on Decision and Control*. Osaka: IEEE, 2015: 5-18.
- [9] KUFFNER J, LAVALLE S. RRT-connect: an efficient approach to single-query path planning [C]//*Proceedings of the 2000 IEEE International Conference on Robotics and Automation*. San Francisco: IEEE, 2000: 995-1001.
- [10] REDDY T, SUPREETHI K. Optimization of K-means algorithm: ant colony optimization [C]//*Proceedings of the 2017 International Conference on Computing Methodologies and Communication*. Erode: IEEE, 2017: 530-535.
- [11] XIN J, ZHONG J, YANG F, et al. An improved genetic algorithm for path-planning of unmanned surface vehicle [J]. *Sensors*, 2019, 19(11): 2640/1-23.
- [12] PADEN B, ČÁP M, YONG S Z, et al. A survey of motion planning and control techniques for self-driving urban vehicles [J]. *IEEE Transactions on Intelligent Vehicles*, 2016, 1(1): 33-55.
- [13] DUBINS L. On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents [J]. *American Journal of Mathematics*, 1957, 79(3): 497-516.
- [14] REEDS J, SHEPP L. Optimal paths for a car that goes both forward and backward [J]. *Pacific Journal of Mathematics*, 1991, 145(2): 367-393.
- [15] WANG S, WANG Y, TAN M, et al. Path planning of underwater biomimetic robot [M]//*Underwater Biomimetic Vehicle-Manipulator System*. Singapore: Springer, 2023: 89-125.
- [16] LIANG Z, ZHENG G, LI J. Automatic parking path optimization based on Bezier curve fitting [C]//Pro-

- ceedings of the 2012 IEEE International Conference on Automation and Logistics. Zhengzhou: IEEE, 2012: 583-587.
- [17] ZIEGLER J, BENDER P, DANG T, et al. Trajectory planning for Bertha—a local, continuous method [C]//Proceedings of the 2014 IEEE Intelligent Vehicles Symposium. Dearborn: IEEE, 2014: 450-457.
- [18] WERLING M, KAMMEL S, ZIEGLER J, et al. Optimal trajectories for time-critical street scenarios using discretized terminal manifolds [J]. The International Journal of Robotics Research, 2012, 31(3): 346-359.
- [19] YU L, SHAO X, WEI Y, et al. Intelligent land-vehicle model transfer trajectory planning method based on deep reinforcement learning [J]. Sensors, 2018, 18(9): 2905/1-22.
- [20] CHAI R, LIU D, LIU T, et al. Deep learning-based trajectory planning and control for autonomous ground vehicle parking maneuver [J]. IEEE Transactions on Automation Science and Engineering, 2023, 20(3): 1633-1647.
- [21] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of 31st Conference on Neural Information Processing Systems. Long Beach: NIPS, 2017: 6000-6010.
- [22] BARMPOUNAKIS E, GEROLIMINIS N. On the new era of urban traffic monitoring with massive drone data: the pNEUMA large-scale field experiment [J]. Transportation Research Part C: Emerging Technologies, 2020, 111: 50-71.
- [23] WERLING M, ZIEGLER J, KAMMEL S, et al. Optimal trajectory generation for dynamic street scenarios in a frenét frame [C]//Proceedings of the 2010 IEEE International Conference on Robotics and Automation. Anchorage: IEEE, 2010: 987-993.
- [24] MCNAUGHTON M, URMSON C, DOLAN J, et al. Motion planning for autonomous driving with a conformal spatiotemporal lattice [C]//Proceedings of the 2011 IEEE International Conference on Robotics and Automation. Shanghai: IEEE, 2011: 4889-4895.
- [25] PIVTORAIKO M, KELLY A. Differentially constrained motion replanning using state lattices with graduated fidelity [C]//Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems. Nice: IEEE, 2008: 2611-2616.
- [26] ZIEGLER J, BENDER P, SCHREIBER M, et al. Making bertha drive—an autonomous journey on a historic route [J]. IEEE Intelligent Transportation Systems Magazine, 2014, 6(2): 8-20.
- [27] ZHOU Z, YE L, WANG J, et al. HiVT: Hierarchical vector transformer for multi-agent motion prediction [C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 8813-8823.
- [28] MERCY T, van PARYS R, PIPELEERS G. Spline-based motion planning for autonomous guided vehicles in a dynamic environment [J]. IEEE Transactions on Control Systems Technology, 2017, 26(3): 1-8.
- [29] LI A, XU Z, LI W, et al. Urban signalized intersection traffic state prediction: a spatial-temporal graph model integrating the cell transmission model and transformer [J]. Applied Sciences, 2025, 15(5): 2377/1-23.
- [30] HAMED M, AL-MASAEID H, SAID Z. Short-term prediction of traffic volume in urban arterials [J]. Journal of Transportation Engineering, 1995, 121(3): 249-254.
- [31] DRUCKER H, BURGESS C, KAUFMAN L, et al. Support vector regression machines [J]. Advances in Neural Information Processing Systems, 1997, 9: 779-784.
- [32] ATHEY S, TIBSHIRANI J, WAGER S. Generalized random forests [J]. The Annals of Statistics, 2019, 47(2): 1148-1178.
- [33] YE Q, WONG S, SZETO W. Short-term traffic speed forecasting based on data recorded at irregular intervals [C] //Proceedings of 2010 13th International IEEE Annual Conference on Intelligent Transportation Systems. Madeira Island: IEEE, 2010: 1541-1546.
- [34] XU Y, LU Y, JI C, et al. Adaptive graph fusion convolutional recurrent network for traffic forecasting [J]. IEEE Internet of Things Journal, 2023, 10(13): 11465-11475.
- [35] WU Z, PAN S, LONG G, et al. Graph wavenet for deep spatial-temporal graph modeling [EB/OL]. (2019-03-31)[2025-11-10]. arXiv preprint, 2019: arXiv:1906.00121.
- [36] LOPEZ P, BEHRISCH M, BIEKER-WALZ L, et al. Microscopic traffic simulation using SUMO [C]//Proceedings of the 2018 21st International Conference on

- Intelligent Transportation Systems. Maui: IEEE, 2018: 2575-2582.
- [37] DOSOVITSKIY A, ROS G, CODEVILLA F, et al. CARLA: an open urban driving simulator [C]//Proceedings of the 1st Conference on Robot Learning. Mountain View: PMLR, 2017: 1-16.
- [38] FAN H, ZHU F, LIU C, et al. Baidu Apollo EM motion planner [EB/OL]. (2018-07-20) [2025-11-10]. arXiv preprint, 2018: arXiv: 1807.08048.
- [39] WESTHOFEN L, NEUROHR C, KOOPMANN T, et al. Criticality metrics for automated driving: a review and suitability analysis of the state of the art [J]. Archives of Computational Methods in Engineering, 2023, 30(1): 1-35.

Predictive Motion Planning for Intelligent Connected Vehicles Based on Transformer Architecture

LI Anran^{1,2} PAN Yuyan³ XU Zhenlin⁴ GAO Bolin² LI Yongxing¹ YU Hongsheng⁵ CHEN Yanyan¹

(1. College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China; 2. School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China; 3. Department of Civil and Environmental Engineering, Pennsylvania State University, PA 16802, USA; 4. Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft 2826 CN, Netherlands; 5. Institute of Electronic Computing Technology, China Railway Academy Group Co., Ltd., Beijing 100081, China)

Abstract: Efficient and safe motion planning for intelligent connected vehicles in complex traffic scenarios remains a pivotal challenge in the field of autonomous driving. This research proposed ST-Trans traffic prediction model based on the Transformer architecture and developed a predictive motion planner for intelligent connected vehicles leveraging ST-Trans. The ST-Trans model utilizes the Transformer architecture to mine spatial-temporal evolution patterns from real-time vehicle data and lane segment structural information provided by dynamic high-definition maps, thereby predicting future traffic states of lane segments. It further enhances prediction accuracy by incorporating lane segment connectivity and intersection signal phase information. The model adopts an encoder-decoder framework, where a lane encoder fuses vehicle and lane features, a road encoder models dynamic topological relationships, and a decoder iteratively generates future traffic state sequences. Experimental results demonstrate that ST-Trans outperforms the optimal baseline model by 12.2%, 12.1%, and 3.55 percentage points in terms of mean absolute error(MAE), root mean square error(RMSE), and accuracy, respectively. Based on the predictions from ST-Trans, the proposed predictive motion planner employs a two-layer structure. The lower-layer path planner dynamically selects target points and integrates dynamic programming with quadratic programming to generate smooth paths. The upper-layer speed planner constructs spatio-temporal corridors to compress the solution space and similarly combines dynamic programming and quadratic programming to generate safe efficient, and comfortable speed profiles. This structure significantly reduces the computational complexity of the motion planning task. Simulation experiments were conducted using SUMO and CARLA to evaluate the predictive motion planner. The results indicate that the ST-Trans-based predictive motion planner successfully implements predictive path and speed planning, and outperforms traditional motion planners in terms of safety, efficiency, comfort, and computational speed. The experiments verify that the proposed method effectively shortens the duration of high-risk states, improves traffic throughput and maintains low computational latency.

Key words: dynamic high-definition map; motion planner; traffic prediction model; deep learning