

doi:10.3969/j.issn.1003-3114.2025.05.024

引用格式:相东升,李成,陈浩,等.基于目的地意图学习的重点人员长期轨迹区域预测模型[J].无线电通信技术,2025,51(5):1113-1127.[XIANG Dongsheng, LI Cheng, CHEN Hao, et al. Long-term Trajectory Area Prediction Model for Key Individuals Based on Destination-intent Learning [J]. Radio Communications Technology, 2025, 51(5): 1113-1127.]

## 基于目的地意图学习的重点人员长期轨迹区域预测模型

相东升<sup>1</sup>, 李成<sup>2</sup>, 陈浩<sup>2</sup>, 陈澄<sup>3</sup>, 李波<sup>4\*</sup>, 韩楠<sup>5</sup>,  
谢添丞<sup>1</sup>, 杨春芳<sup>6,7</sup>, 乔少杰<sup>1</sup>

(1. 成都信息工程大学 软件工程学院, 四川 成都 610225;

2. 成都市公安局 科技信息化处, 四川 成都 610017;

3. 成都市公安信息技术研究所, 四川 成都 610017;

4. 四川省妇幼保健院 信息部, 四川 成都 610041;

5. 成都信息工程大学 管理学院, 四川 成都 610225;

6. 网络空间安全教育部重点实验室, 河南 郑州 450001;

7. 河南省网络空间态势感知重点实验室, 河南 郑州 450001)

**摘要:**重点人员轨迹预测对于预防潜在犯罪行为、优化应急响应及情报分析具有重要作用,公安部门应用该项技术能够维护社会稳定、提高城市管理效率、促进经济发展。现有技术面临动态环境适应、忽视社交影响范围、周围移动对象影响量化等挑战,基于此提出一种新型融合时空查询的基于目的地意图学习的重点人员长期轨迹区域预测模型。针对移动对象轨迹时空特征难以捕获的问题,提出一种基于时空多注意力(Spatio-Temporal Multiple Attention, STMA)机制的重点人员轨迹预测模型,通过时间注意力模块和空间注意力模块分别捕捉时间依赖关系和空间交互关系,提高模型对行为特征变化的敏感性;针对社交影响力难以量化的问题,构建社会力函数模拟行人社交影响力,通过虚拟轮廓构建方法和社会力函数,准确模拟动态行为并提高影响力捕捉效率。实验基于真实世界交通数据集进行,结果表明:与当前先进的轨迹预测算法相比,STMA在长短期预测方面均展现出更高的准确率和可靠性,在长期预测方面,其平均准确率为54.3%,比代表性算法Sophie、行人轨迹图表示的时空卷积网络(Social Spatio Temporal Graph Convolutional Neural Network, S-STGCNN)和条件生成神经网络(Conditional Generative Neural System, CGNS)分别高29.3%、13.4%和36.8%。

**关键词:**时空轨迹;轨迹预测;多注意力机制;虚拟轮廓;社会力函数

中图分类号:TP311 文献标志码:A

开放科学(资源服务)标识码(OSID):

文章编号:1003-3114(2025)05-1113-15



收稿日期:2024-11-07

**基金项目:**国家自然科学基金(62272066);四川省科技计划资助(2025ZNSFSC0044, 2025YFHZ0194, 2024YFFK0413);成都市技术创新研发项目(2024-YF05-01217-SN);成都市区域科技创新合作项目(2025-YF11-00050-HZ);成都市重点研发支撑计划产业链协同创新项目(2025-XT00-00005-GX);成都市技术创新研发项目重点项目(2025-YF08-00016-GX);网络空间安全教育部重点实验室及河南省网络空间态势感知重点实验室开放基金(KLCS20240106);网络空间大数据智能安全教育部重点实验室开放基金(CBDIS202404)

**Foundation Item:** National Natural Science Foundation of China (62272066); Sichuan Science and Technology Program (2025ZNSFSC0044, 2025YFHZ0194, 2024YFFK0413); Chengdu Technological Innovation Research and Development Project(2024-YF05-01217-SN); Chengdu Regional Science and Technology Innovation Cooperation Project(2025-YF11-00050-HZ); Industry Chain Collaborative Innovation Project of Chengdu Key Research and Development Support Program(2025-XT00-00005-GX); Chengdu Technological Innovation Research and Development Major Project(2025-YF08-00016-GX); Open Foundation of Key Laboratory of Cyberspace Security, Ministry of Education of China and Henan Key Laboratory of Cyberspace Situation Awareness(KLCS20240106); Open Research Fund of Key Laboratory of Cyberspace Big Data Intelligent Security, Ministry of Education of China(CBDIS202404)

# Long-term Trajectory Area Prediction Model for Key Individuals Based on Destination-intent Learning

XIANG Dongsheng<sup>1</sup>, LI Cheng<sup>2</sup>, CHEN Hao<sup>2</sup>, CHEN Cheng<sup>3</sup>, LI Bo<sup>4\*</sup>, HAN Nan<sup>5</sup>, XIE Tiancheng<sup>1</sup>,  
YANG Chunfang<sup>6,7</sup>, QIAO Shaojie<sup>1</sup>

(1. School of Software Engineering, Chengdu University of Information Technology, Chengdu 610225, China;

2. Science and Technology Information Division, Chengdu Municipal Public Security Bureau, Chengdu 610017, China;

3. Chengdu Public Security Information Technology Research Institute, Chengdu 610017, China;

4. Information Department, Sichuan Provincial Maternity and Child Health Care Hospital, Chengdu 610041, China;

5. School of Management, Chengdu University of Information Technology, Chengdu 610225, China;

6. Key Laboratory of Cyberspace Security, Ministry of Education of China, Zhengzhou 450001, China;

7. Henan Key Laboratory of Cyberspace Situation Awareness, Zhengzhou 450001, China)

**Abstract:** Predicting trajectories of key individuals plays an important role in preventing potential criminal activities, optimizing emergency response, and intelligence analysis. Application of this technology by public security departments helps maintain social stability, improve urban management efficiency, and improve economic development. However, existing techniques face challenges in adapting to dynamic environments, neglecting the scope of social influence, and influence quantification of neighborhood moving objects. A novel model for predicting long-term trajectory areas of key individuals based on destination-intention learning by integrating spatio-temporal queries, is proposed. Firstly, aiming to solve the problem of capturing the spatio-temporal features of moving object trajectories, a key individuals trajectory prediction model called Spatio-Temporal Multiple Attention (STMA) is introduced. It can enhance the model sensitivity to the change of behavioral features by capturing temporal dependencies and spatial interactions through temporal and spatial attention modules, respectively. Secondly, in order to cope with the problem of quantifying the social influence, a social force function is constructed to simulate the social influence of pedestrians. The virtual contour construction method and the social force function can accurately simulate dynamic behaviors and improve the efficiency of influence capture. Experiments based on real-world traffic datasets show that, compared to the state-of-the-art trajectory prediction algorithms, STMA demonstrates higher accuracy and reliability in long-term and short-term trajectory prediction. In terms of long-term forecasting, the STMA model achieves an average accuracy rate of 54.3%, outperforming Sophie by 29.3%, Social Spatio Temporal Graph Convolutional Neural Network (S-STGCNN) by 13.4%, Conditional Generative Neural System (CGNS) by 36.8%.

**Keywords:** spatio-temporal trajectory; trajectory prediction; multi-attention mechanism; virtual contour; social force function

## 0 引言

在信息技术和大数据技术迅猛发展的推动下,人类行为轨迹数据呈现出前所未有的丰富性。在庞大的数据资源中,重点人员的行为轨迹数据因其具有重要的战略价值和社会影响力,成为了公共安全、城市管理等行业密切关注的核心。鉴于此,探索一种高效的方法来预测重点人员的长期轨迹区域,已成为学术界和业界研究的热点课题。本文针对上述热点研究内容,提出了一种新型基于目的地意图学习的重点人员长期轨迹区域预测模型。

目的地意图作为影响个体行为轨迹的核心要素,揭示了人员移动的根本目的和内在动机。通过对目的地意图的深入学习,能够更为精确地预见人

员未来的行为路径。然而,现行轨迹预测技术往往未能充分重视目的地意图的作用,或者仅限于短期轨迹的预测,未能对长期轨迹区域进行深入探究。上述局限性在一定程度上降低了预测模型在实际应用中的准确度和实用性。目前,准确建模和预测重点人员的未来轨迹面临诸多挑战:

① 在动态和密集的人群环境中,重点人员的运动通常受到其自身历史路径以及与邻近人员的相互作用的影响。

② 无法处理移动对象之间的复杂关系。轨迹预测中,移动对象之间存在复杂的交互,预测某个重点人员的轨迹需要考虑其与行人、车辆以及环境之间的互动。

针对现有轨迹预测方法的不足,本文创新之处体

现在以下方面:首先,建立了一个综合历史轨迹数据、社会环境信息以及个体行为特征的目的地图学习框架,旨在深入剖析重点人员的移动目的;其次,提出一种新型长期轨迹区域预测模型,该模型全面考虑了时间依赖性、空间相关性以及个体行为的多样性,显著提升了预测的精确度和持久性;最后通过实验验证了模型的实效性。本文的主要贡献包括:

① 针对移动对象轨迹的时空依赖性,提出了一种基于 STMA 机制的重点人员轨迹预测模型,该模型引入了时空多重注意力机制,并对交互场景中其他移动对象的时空关系进行加权,使模型不仅能够捕获不同时间步的历史轨迹依赖关系,还能获得其他移动对象的空间距离依赖关系,使模型能够有效预测移动目标在空间中的移动趋势。

② 通过考虑现实复杂场景中重点人员与行人之间的交互、重点人员与环境之间的交互以及重点人员与车辆之间的交互,提出了一种情境感知驱动的多影响力轨迹预测方法,并基于虚拟轮廓技术,构建了考虑多方面因素的社会力函数,有效提升了模型轨迹预测的准确率。

③ 在公开数据集上进行实验,证明了本文方法的有效性,通过与经典轨迹预测方法以及先进轨迹预测方法的对比实验,验证了本文所提模型的性能优势。

## 1 相关工作

### 1.1 轨迹预测基本概念

轨迹预测是指通过移动对象过去和当前的位置数据来预测其在未来某个时间点或时间段内的位置<sup>[1]</sup>。这涉及到对移动对象如人、车辆等在空间上已移动的路径进行分析和推理,路径序列信息通常表现为:已知一个移动对象在时间  $t_1, t_2, \dots, t_{obs-1}, t_{obs}$  的位置序列为  $(x_1, y_1), (x_2, y_2), \dots, (x_{obs-1}, y_{obs-1}), (x_{obs}, y_{obs})$ , 轨迹预测的目的是根据移动目标这一序列的轨迹预测该目标未来的轨迹,即需要得出该目标在时间  $t_{obs+1}, t_{obs+2}, \dots, t_n$  的位置序列  $(x_{obs+1}, y_{obs+1}), (x_{obs+2}, y_{obs+2}), \dots, (x_n, y_n)$ , 其中  $obs$  为观测周期的最后一个时间节点。通过上述一系列轨迹数据可以描述出一个移动对象的轨迹,如在  $t$  时刻,移动定位设备返回一个移动对象的数据是  $(x, y)$ , 说明  $t$  时刻对象所在位置为  $(x, y)$ 。每次间隔同一段时间记录该对象的位置数据则可在地图上标出一系列的轨迹点,将轨迹点按出现时间的先后顺序连接在一起构成移动对象的一条线性轨迹。

如图 1 所示,若这一系列轨迹点在某一处的间隔较小,则可推测目标附近存在一个减速点,如目标在公园散步或是在车道上经过堵车区域等;若某一处的轨迹点间隔特别密集,则可推测该目标附近存在一个停留点,如目标可能在餐厅堂食用餐或者在公司办公等情况;若某一处的轨迹点间隔非常稀疏,则可推测目标附近存在移动模式改变的变化点,如目标可能从步行变为了乘坐车辆等情况,此时需要使用不同的数学变量表达不同的移动模式。

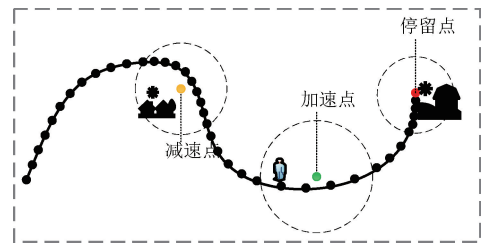


图 1 简单轨迹举例

Fig. 1 Example of a simple trajectory

### 1.2 传统轨迹预测

传统的轨迹预测模型通常基于物理定律、基于经典机器学习和基于情景交互。轨迹预测方法通常假设历史轨迹数据与预测轨迹数据之间存在一定的相关性,通过海量的历史轨迹数据构建相应的数学模型,进而实现对移动对象运动趋势的判断<sup>[1]</sup>。常用的方法包括卡尔曼滤波器、隐马尔可夫模型、高斯混合模型、贝叶斯网络等。Prevost 等<sup>[2]</sup>率先将卡尔曼滤波器用于轨迹预测领域,用在短距离的车辆轨迹预测。乔少杰等<sup>[3]</sup>提出了一种基于卡尔曼滤波的动态轨迹预测算法,利用前一时刻的估计值和当前时刻的观测值更新对状态变量的估计,进而实现对下一时刻的轨迹位置预测,获得了较好的预测效果。Qiao 等<sup>[4]</sup>使用隐马尔可夫模型,设计实现了一种可以自适应调整轨迹预测参数的动态预测算法,根据不同类型轨迹数据预测最佳路线,但是这一模型没有考虑大数据环境下算法的运行时间等性能问题。此后,乔少杰等<sup>[5]</sup>在此基础上提出了一种自适应参数选择的高阶隐马尔可夫轨迹预测模型,以解决大数据环境下移动对象轨迹预测问题。夏卓群等<sup>[6]</sup>提出了一种变分高斯混合模型,在传统的高斯混合模型的基础上引入变分贝叶斯近似推理方法和自适应参数选择方法,解决了高斯混合模型的权重、均值和精度这 3 个参数的估计问题并完成了轨迹预测。李万高等<sup>[7]</sup>将历史轨迹分解后,对相邻位置信息进行处理,提升了贝叶斯模型的时效性和精准度。

### 1.3 深度学习类神经网络

传统的轨迹预测方法主要依赖于理论建模的框架,其预测能力主要取决于模型提前预设的机制,且这些机制通常由一些固定公式或参数组成,需要进行校准和验证,因此传统方法预测精度较低。基于深度学习方法的轨迹预测方法能够更准确地学习历史轨迹数据的特征,根据特征预测移动目标的移动趋势,轨迹预测的深度学习方法有以下几种:循环神经网络(Recurrent Neural Network, RNN)、长短期记忆网络(Long Short-Term Memory, LSTM)、卷积神经网络(Convolutional Neural Network, CNN)等。

Min 等<sup>[8]</sup>基于 RNN 对车辆的障碍物路径进行预测,通过预测路径的不确定性来预测车辆的未来轨迹。理论上,基于 RNN 的模型可以保留数据集中所有的历史信息,但其在处理长期依赖关系方面仍有所欠缺。Alahi 等<sup>[9]</sup>利用 LSTM 捕捉移动对象在时间序列上的移动特征,在模型中构建了社交池化层,通过计算拥挤空间中各移动目标对其他对象的交互影响,使模型能够根据这些影响力预测各移动对象后续的移动趋势。Nikhil 等<sup>[10]</sup>提出了一种基于 CNN 的轨迹预测模型,使用端到端卷积结构进行轨迹预测。模型使用堆叠卷积层以增强时间的一致性,并通过一个完全连接层计算其位置信息。

### 1.4 注意力机制

注意力机制<sup>[11]</sup>可以使轨迹预测模型在复杂或拥挤场景中处理时空序列信息时对特定的动态特征提高关注度,赋予高权重,通过模拟真实场景中移动对象在移动过程中的关注度分配能使模型在复杂的场景中做出合理的预测。Liang 等<sup>[12]</sup>为解决复杂道路网络结构中难以捕获车辆轨迹依赖性的问题,提出了一种基于深度学习的序列到序列(Sequence to Sequence, Seq2Seq)轨迹预测模型,构建了空间和时间注意力层用于模拟道路网络中轨迹数据的动态空间依赖以及短期轨迹与长期轨迹的依赖关系。Chen 等<sup>[13]</sup>提出了一种包含注意力机制的 LSTM 轨迹预测模型,该模型在编码阶段输出特定的高权重部分,通过计算解码、编码时刻的注意力权重,对编码器输出进行加权求和,生成的上下文信息能够使轨迹预测更加准确。

### 1.5 情景感知技术

重点人员在移动过程中会伴随着避开障碍物、

适当远离其他移动对象的行为,因此在轨迹预测中要使模型具有类似人类的思考模式,需要感知周围的情景。考虑场景中其他的障碍物或是移动对象的轨迹预测模型往往预测能力会高于传统的轨迹预测模型。Li 等<sup>[14]</sup>提出了一种基于 Transformer 网络的多模态轨迹预测框架,其情境感知模块分别是碰撞感知区、周围感知区以及场景感知区,并运用了多头线性全局注意力机制,减轻了 Transformer 框架的运行负载。Zhao 等<sup>[15]</sup>提出了一种多智能体张量融合网络,将周围移动目标轨迹与交互场景环境编码成一个张量以模拟人与环境之间的相互作用,结合 U-Net CNN 获取情境感知信息,使用生成对抗网络模拟行人的随机性行为,通过 LSTM 解码未来的轨迹来完成预测。Liang 等<sup>[16]</sup>为预测视频中的行人轨迹,提出了名为 Next 的多任务学习模型,通过提取行人关键行为和动作特征的人物行为模块以及分析行人与周围环境交互的人物交互模块使模型具有丰富的视觉信息,采用包含注意力机制的 LSTM 编解码器生成行人未来的轨迹,且利用曼哈顿网络预测行人未来的行动。Beregruer 等<sup>[17]</sup>提出了一种用于行人轨迹预测的情景交互隐藏变量模型,通过使用一个 RNN 编码器和 2 个 RNN 的解码器这种复合结构捕捉行人轨迹的属性,引入注意力机制和 CNN 分析情景互动以及周围环境与行人之间的交互信息,使复杂交互场景中也能具有较好的预测性能。

## 2 模型架构

### 2.1 基于 STMA 机制的轨迹预测模型构建

融合 STMA 目标重点人员轨迹预测模型架构如图 2 所示。首先,采用目标检测算法以及目标跟踪算法对视频数据进行结构化处理,以确保从视频帧中精准提取每个移动目标的位置信息及运动轨迹等状态信息。随后,编码器对各对象的状态信息进行编码生成隐藏状态,并将其输入 STMA 模块。其中,由空间注意力模块计算重点人员交互场景的空间信息的重要性,用于提取移动对象之间的潜在互动信息,为每个移动对象分配不同的权重。由时间注意力模块计算各移动目标每个时间步的重要性,捕捉单个重点人员不同时间步长之间的潜在依赖关系,为其输入序列分配不同的权重。最后,与高斯噪声一同输入解码器,生成预测轨迹。

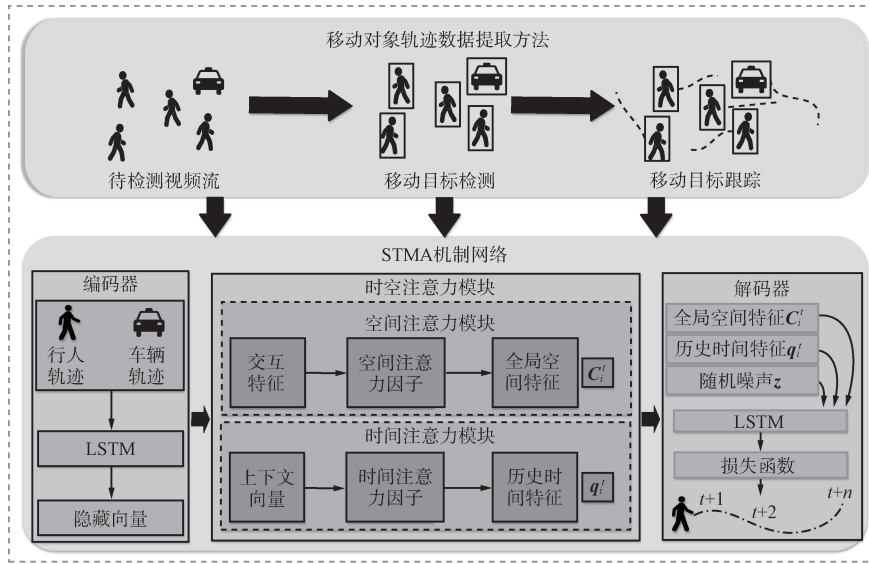


图2 融合STMA目标重点人员轨迹预测模型架构

Fig. 2 Architecture of the target key individuals trajectory prediction with integrated STMA

### 2.2 移动对象轨迹数据提取技术

STMA模型采用YOLOv5目标检测算法以及DeepSORT目标跟踪算法对视频数据进行结构化处理。首先,利用YOLOv5从视频帧中提取出每个移动目标的位置信息,并为每个目标分配不同且唯一的标识符。然后,通过DeepSORT算法对各移动目标进行跟踪,将同一个标识符的目标连续帧的位置坐标连接形成该目标的运动轨迹。关键技术包括:

#### (1) 图像预处理

帧提取是对视频数据的特有处理方式,由于目标检测通常是对静态图像中的物体进行识别,在处理视频数据时需要先将其转化为图片,通常采用帧提取方法将视频拆解为单独的帧,以便逐帧分析。帧提取如式(1)所示:

$$F(t) = V(t \cdot \Delta t), \quad (1)$$

式中: $F(t)$ 表示视频中时间点为 $t$ 的帧, $V$ 表示总视频流, $\Delta t$ 为提取视频帧的时间间隔。

通过对单帧的图像进行去噪,减少图像传输过程中的干扰,提高信噪比。噪声通常表现为图像中分布不均的明暗点或是颗粒纹理,在目标检测和跟踪过程中可能会成为影响最终预测效果的因素。在抑制噪声方面使用高斯滤波方法,通过高斯函数计算图像中每个像素点周围像素的加权平均以达到去噪的效果,如式(2)所示:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad (2)$$

式中: $G(x, y)$ 为坐标 $(x, y)$ 处的权重, $x$ 和 $y$ 为图像像素点的横坐标和纵坐标, $\sigma$ 为高斯分布的标准差参数,其大小决定了权重分布的范围。

#### (2) 移动目标检测技术

YOLOv5会使算法在将预处理后的图像平均划分为若干个栅格(Grid Cell),每个栅格通过深度卷积神经网络提取其中图像的特征,识别和检测它覆盖区域内可能存在的对象,生成多个边界框(Bounding Boxes)尝试覆盖该区域内可能存在的对象。边界框各自具有不同的比例和尺寸,使模型能够捕捉到不同形状和大小的目标。对于每个边界框,YOLOv5不仅需要预测其位置和尺度,还需要预测一个表示边界框内包含目标的概率的置信度(Confidence Score),以及表示该目标属于每个可能类别的类别概率。模型的预测结果通过非极大值抑制进行优化,去除重叠的边界框,确保每个目标匹配唯一的标识。

#### (3) 移动目标跟踪技术

DeepSORT<sup>[18]</sup>是一种基于深度学习技术对视频流中的目标进行实时跟踪的算法,优势在于结合了深度学习的外观特征,使得算法在目标外观发生变化、目标密集或遮挡的情况下仍能稳定跟踪,进而提升匹配准确性。其为视频中每一个移动对象目标分配一个独特的标识符并在连续的视频帧中维持该目标标识符的一致性。DeepSORT工作原理如图3所示。

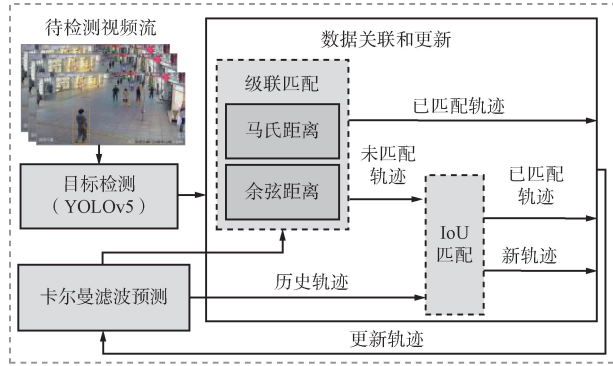


图3 DeepSORT 工作原理

Fig. 3 DeepSORT working schematic diagram

2.3 移动对象状态信息构建

STMA 模型将重点人员轨迹预测建模为一个时间序列预测问题,输入涵盖重点人员以及周围行人和车辆的轨迹、状态信息,输出则为重点人员未来一段时间内的轨迹序列。本文研究的混合交通场景假设包含多个行人以及单个汽车,重点人员、车辆轨迹及状态信息,用位置坐标、速度大小和方向角表示,其中,方向角为速度方向与  $x$  轴之间的夹角,重点人员与车辆的轨迹表示如式(3)~式(5)所示:

$$p_i^t = \{(x_i^t, y_i^t) \mid t = 1, 2, \dots, obs\}, \quad (3)$$

$$p_v^t = \{(x_v^t, y_v^t) \mid t = 1, 2, \dots, obs\}, \quad (4)$$

$$p_i^{pred} = \{(x_i^t, y_i^t) \mid t = obs + 1, obs + 2, \dots, obs + pred\}, \quad (5)$$

式中:  $p_i^t$ 、 $p_v^t$ 、 $p_i^{pred}$  分别为重点人员的历史轨迹信息、车辆的历史轨迹信息以及预测的重点人员未来轨迹信息,  $obs + pred$  为预测周期的最后一个时间节点,  $pred$  为预测持续时间。

2.4 基于 STMA 模块的编解码器

STMA 模型在时间和空间维度上模拟重点人员与普通行人之间、重点人员与车辆之间的交互。本节介绍一种包含 STMA 模块的编解码器架构,其网络结构如图 4 所示。编码器首先对普通行人、重点人员和车辆的轨迹和状态信息进行编码,生成特征向量。然后,将通过 LSTM 生成的隐藏状态输入 STMA 模块。该模块包括 2 个子组件:空间注意力模块与时间注意力模块,空间注意力模块计算每个移动对象的空间信息对于目标重点人员的重要性,时间注意力模块则计算目标历史轨迹中各时间步对当前时刻行为状态的影响权重,从而提升模型对动态行为模式的捕捉能力。最后,与高斯噪声一同输入解码器中进行轨迹预测,生成一组目标重点人员未来的轨迹坐标序列。

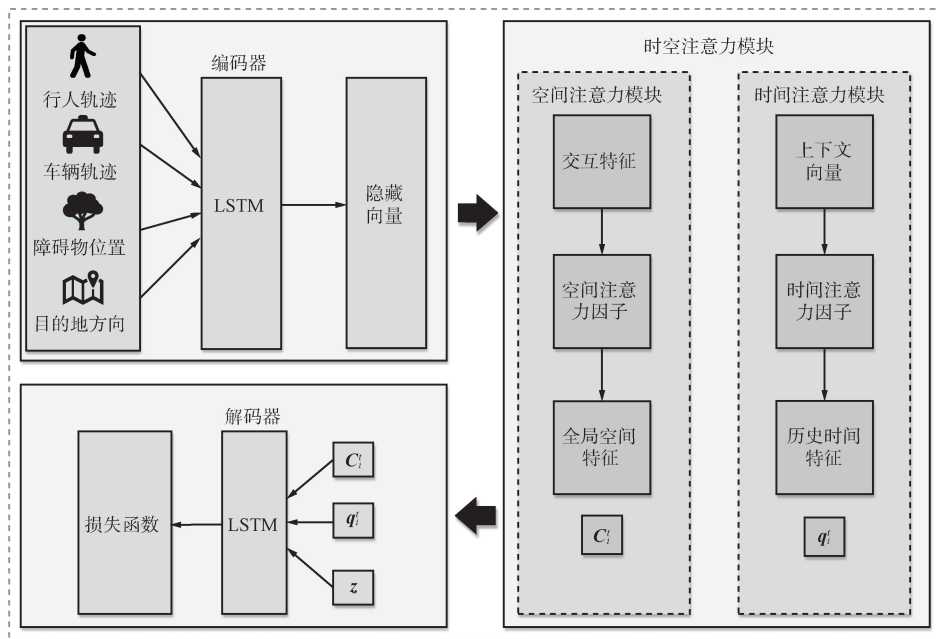


图4 STMA 编解码器架构

Fig. 4 Architecture of STMA encoder-decoder

2.4.1 编码器模块

编码器网络结构如图 5 所示,编码器通过 LSTM

对所有移动人员和车辆的轨迹数据进行预处理,并将轨迹特征编码为隐藏状态。

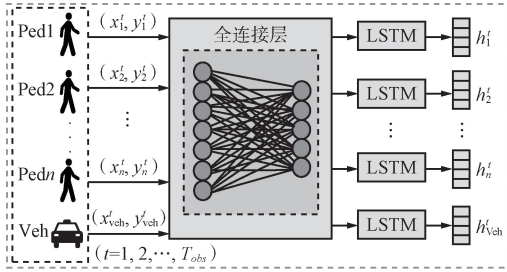


图5 编码器网络结构

Fig. 5 Network structure of the encoder

定义重点人员  $i$  与车辆在  $t$  时刻的位置坐标  $p_i^t = (x_i^t, y_i^t)$  以及  $p_v^t = (x_v^t, y_v^t)$ , 将重点人员和车辆坐标拼接可得到  $t$  时刻观测到的信息向量  $X_i^t$ :

$$X_i^t = \phi(p_i^t, p_v^t), \quad (6)$$

式中:  $\phi(\cdot)$  为向量拼接函数。编码器将拼接后的信息向量  $X_i^t$  嵌入到全连接层中得到语义向量  $e_{iv}^t$ , 并输入到 LSTM 网络中计算当前时刻的隐藏状态, 其计算如式(7)所示:

$$\begin{cases} h_{i,en}^t = LSTM(h_{i,en}^{t-1}, e_{iv}^t; F_i^t, W_F; W_{en}, b_{en}) \\ e_{iv}^t = FC(X_i^t, W_{FC1}) \end{cases} \quad (7)$$

式中:  $FC$  为全连接层,  $W_{FC1}$  为编码器全连接层的权重,  $F_i^t$  为  $t$  时刻重点人员  $i$  所受的社会影响力,  $h_{i,en}^{t-1}$  为  $t$  时刻重点人员  $i$  的隐藏状态,  $h_{i-1,en}^{t-1}$  为前一时刻重点人员  $i$  的隐藏状态,  $W_{en}$  为 LSTM 网络的权重,  $b_{en}$  为偏置。

先拼接各时刻的信息向量  $X_i^t$  得到重点人员的特征向量, 单个重点人员的隐藏状态通过每个时刻该重点人员的隐藏状态依次拼接而成; 再将若干个重点人员的隐藏状态进行拼接进而得到所有重点人员的隐藏状态, 即编码器的输出, 如式(8)所示:

$$\begin{cases} X_i = \phi(X_i^1, X_i^2, \dots, X_i^{obs-1}, X_i^{obs}) \\ H_{i,en} = \phi(h_{i,en}^1, h_{i,en}^2, \dots, h_{i,en}^{obs-1}, h_{i,en}^{obs}) \\ H'_{en} = \phi(H_{1,en}, H_{2,en}, \dots, H_{N-1,en}, H_{N,en}) \end{cases}, \quad (8)$$

式中:  $X_i$  为重点人员  $i$  的单人特征向量,  $H'_{en}$  为重点人员  $i$  的单人聚合隐藏状态,  $H_{en}$  为所有重点人员的聚合隐藏状态,  $N$  为所有人员的数量。

### 2.4.2 时间注意力模块

时间注意力模块旨在为时间序列中不同的时间步给予不同的关注度, 以解决序列特征时间依赖度较低的问题, 其结构如图6所示,  $T_{ob}$  为轨迹观测时长 (Trajectory Observation Length)。

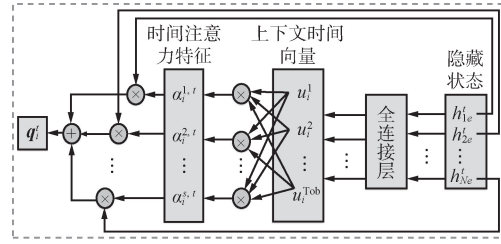


图6 时间注意力模块结构

Fig. 6 Structure of the temporal attention module

时间特征向量计算如式(9)所示:

$$\begin{cases} u_i^t = FC(h_{i,en}^t; W_a, b_a) = \tanh((W_a)^T h_{i,en}^t + b_a) \\ a_i^t = \frac{e^{(u_i^t)^T u_i^t}}{\sum_{k=1}^{obs} e^{(u_i^k)^T u_i^k}} \end{cases}, \quad (9)$$

式中:  $u_i^t$  为上下文时间向量,  $a_i^t$  为时间注意力特征向量,  $\tanh$  为激活函数,  $W_a$  和  $b_a$  为时间注意网络参数,  $T$  为转置操作。

通过结合隐藏状态及时间特征向量计算历史时间特征:

$$q_i^t = \sum_{i=1}^{obs} a_i^t h_{i,en}^t, \quad (10)$$

式中:  $q_i^t$  为  $t$  时刻重点人员  $i$  的历史时间特征, 包含可观测时间内不同时刻之间的关系, 并对不同时刻的轨迹信息赋予了相应的权重。

### 2.4.3 空间注意力模块

空间注意力模块用于分析重点人员周围的行人和车辆的空间影响, 根据周围移动目标的物理数据过滤出有利用价值的交互信息, 其模块结构如图7所示。

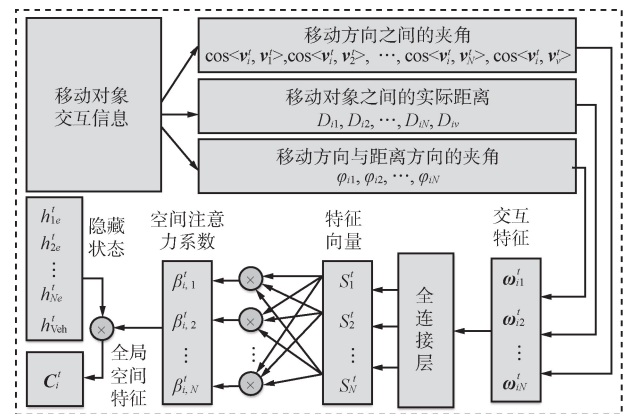


图7 空间注意力模块结构

Fig. 7 Structure of the spatial attention module

空间交互特征主要分为交互距离和交互角度2类,交互距离包括重点人员和车辆的交互距离以及重点人员和行人之间的交互距离:

$$\begin{cases} D_{ij}^t = \sqrt{(x_i^t - x_j^t)^2 + (y_i^t - y_j^t)^2} \\ D_{iv}^t = \sqrt{(x_i^t - x_v^t)^2 + (y_i^t - y_v^t)^2} \end{cases}, \quad (11)$$

式中: $D_{ij}^t$ 表示 $t$ 时刻 $i$ 与行人 $j$ 之间的实际距离, $D_{iv}^t$ 为 $t$ 时刻 $i$ 与车辆之间的实际距离。上述距离均不包含虚拟轮廓,即为目视距离。交互角度包括了行人移动方向以及车辆移动方向,其特征向量计算如下所示:

$$\begin{cases} \omega_{ij}^t = \phi(D_{ij}^t, \cos \varphi_{ij}^t, \cos \langle \mathbf{v}_i^t, \mathbf{v}_j^t \rangle) \\ \omega_{iv}^t = \phi(D_{iv}^t, 0, \cos \langle \mathbf{v}_i^t, \mathbf{v}_v^t \rangle) \end{cases}, \quad (12)$$

式中: $\mathbf{v}_i^t$ 、 $\mathbf{v}_j^t$ 以及 $\mathbf{v}_v^t$ 为 $t$ 时刻 $i$ 、 $j$ 以及车辆的移动方向向量, $\omega_{ij}^t$ 为 $t$ 时刻 $i$ 与 $j$ 的交互角度特征向量, $\omega_{iv}^t$ 为 $t$ 时刻 $i$ 与车辆的交互角度特征向量。聚合所有交互对象的特征向量可得总交互特征向量,并通过全连接层拼接,计算如下:

$$\omega_i^t = \phi(\omega_{i1}^t, \omega_{i2}^t, \dots, \omega_{ij}^t, \dots, \omega_{iN}^t, \omega_{iv}^t), \quad (13)$$

$$\mathbf{S}_i^t = FC(\omega_i^t; W_{FC2}), \quad (14)$$

式中: $\omega_i^t$ 为 $t$ 时刻 $i$ 周围所有交互对象的总交特征向量, $\mathbf{S}_i^t$ 为 $t$ 时刻 $i$ 的全连接空间特征向量, $W_{FC2}$ 为空间注意力模块的全连接连通层参数。根据全连接空间特征向量和聚合隐藏状态计算可得空间注意力系数:

$$\beta_{ij}^t = \frac{e^{(S_i^t, h_{j, en}^t)}}{\sum_{k=1, k \neq i}^N e^{(S_i^t, h_{k, en}^t)}}, \quad (15)$$

式中: $\beta_{ij}^t$ 为 $t$ 时刻 $i$ 与 $j$ 之间交互作用的空间注意力系数。对其进行归一化处理,计算出 $i$ 的全局空间特征:

$$\mathbf{C}_i^t = \phi\left(\sum_{j=1}^N \beta_{ij}^t \mathbf{H}_{i, en}^t\right), \quad (16)$$

式中: $\mathbf{C}_i^t$ 为重点人员 $i$ 的全局空间特征。对全局空间特征和历史时间特征进行对应元素相加可得重点人员 $i$ 在 $t$ 时刻的时空注意力特征向量,其计算如式(17)所示:

$$\mathbf{v}_i^t = \mathbf{C}_i^t \oplus \mathbf{q}_i^t, \quad (17)$$

式中: $\mathbf{v}_i^t$ 为 $t$ 时刻重点人员 $i$ 的时空注意力特征向量,包含了时间、空间特征,可以有效地描述人车之间以及重点人员和普通行人之间的交互作用,使用解码器对其进行解码可得重点人员 $i$ 的后续轨迹。

#### 2.4.4 解码器模块

解码器网络结构如图8所示,为模拟重点人员轨迹在交通场景移动时的不确定性,解码器的输入除时空注意力特征向量外还引入了高斯分布采样的随机噪声,用于模拟交通场景边界噪声。

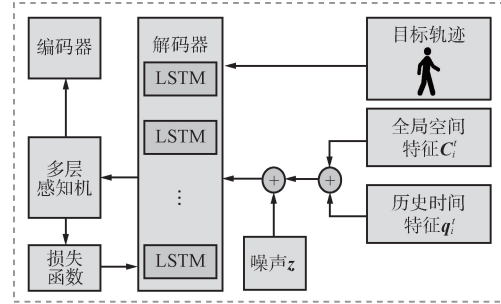


图8 解码器网络结构

Fig. 8 Network structure of the decoder

解码器定义如式(18)~式(21)所示:

$$\mathbf{r}_i^t = \mathbf{v}_i^t \oplus \mathbf{z}_i^t, \quad (18)$$

$$\mathbf{h}_{i, de}^{t+1} = LSTM(\mathbf{h}_{i, de}^t, \mathbf{r}_i^t; W_{de}), \quad (19)$$

$$\mathbf{h}_{i, de}^{obs} = \mathbf{h}_{i, en}^{obs}, \quad (20)$$

$$(x_{i, t+1}^t, y_{i, t+1}^t) = \delta(\mathbf{h}_{i, de}^{t+1}; W_M), \quad (21)$$

式中: $\mathbf{z}_i^t$ 为高斯噪声, $\mathbf{r}_i^t$ 为输入解码器的状态向量, $\mathbf{h}_{i, de}^t$ 为解码器 $t$ 时刻的隐藏状态, $\delta$ 为多层感知机, $W_{de}$ 和 $W_M$ 分别为解码器LSTM和多层感知机的权重参数,通过多层感知机处理解码器当前时刻的隐藏状态即可获得目标下一时刻位置坐标 $(x_{i, t+1}^t, y_{i, t+1}^t)$ 。

轨迹预测模型的损失函数通常使用均方误差损失函数,而并没有考虑到每个时间步之间的差异性,随着该误差的积累会导致有效预测长度逐渐缩短。为了解决上述问题采用最小化多模态损失函数并设置与时间相关的参数项控制每个误差在总损失中的比例,定义如式(22)所示:

$$Loss = \frac{\sum_{i=1}^N \sum_{t=obs+1}^{obs+pred} \|\mathbf{Y}_i^t - \hat{\mathbf{Y}}_i^t\|_2 \cdot e^{\lambda t}}{N \cdot pred}, \quad (22)$$

式中: $\hat{\mathbf{Y}}_i^t$ 为预测轨迹, $\mathbf{Y}_i^t$ 为实际轨迹, $\lambda$ 为控制时间依赖性强度的参数,用于调整不同时间步长的重要性。

### 3 情境感知驱动的多影响力轨迹预测

伴随交互环境复杂程度的上升,重点人员在移动过程中所需要考虑的因素也越来越多。在移动

时,不仅需要观察周围其他移动对象的位置、轨迹变化来了解他人的移动趋势,还需要避让障碍物等其他场景对象。当其他移动对象或障碍物距离重点人员越近时,其对重点人员造成的影响力也会越大。现有的方法大多数侧重于行人之间的交互,而忽略了重点人员在移动过程中所受的其他影响力。为了量化这种影响力,提出了一种情境感知驱动的多影响力轨迹预测方法,构建了一种基于移动对象虚拟轮廓的社会力函数用于优化 STMA 对重点人员所受影响力的捕捉。

STMA 的主要创新点在于:① 结合移动对象在移动过程中的物理因素为每个移动对象构建了虚拟轮廓,模拟真实交互场景中移动对象之间的交互范围;② 提出衰减函数和各向异性函数并结合其功能设计了一种新型社会力函数以整合重点人员移动时所受到的影响力,用于优化 STMA 对行人所受影响力的捕捉。

本文设计的社会力函数构建如图 9 所示。首先,提取各移动对象的行为模式,如位置坐标、移动方向等信息。其次,对周围移动环境进行情境感知,识别交互场景中的障碍物等干扰因素。然后,基于移动对象在移动过程中的物理因素,为每个移动对象都构建虚拟轮廓,以模拟移动对象可能造成或受到影响的范围。最后,设计了各向异性函数以及衰减函数,对移动对象的交互进行建模,整合为社会力函数。



图 9 社会力函数构建

Fig. 9 Construction of social force functions

### 3.1 移动对象运动建模

#### (1) 各向异性函数以及衰减函数的设计

通过设计合适的各向异性函数和衰减函数能够描述移动对象之间相互作用力的特性。各向异性函数描述不同方向上相互作用力的效果,例如位于重点人员正前方行人的影响力会明显大于重点人员左右两侧行人的影响力。衰减函数描述相互作用力在不同的交互距离下的影响,例如一辆车在远距离对重点人员的影响会远远低于近距离车辆的影响,并

且不同的交互场景下影响随距离的变化率也不一样。

各向异性函数描述和处理具有方向依赖型的动态行为,为不同的方向赋予不同的权重和属性<sup>[18]</sup>。行走时会受到如其他移动对象或障碍物的阻碍力,面对这些阻碍时重点人员的前进方向会偏离这些阻碍所在的方向,可以看作重点人员受到的阻碍力会随着前进角度偏离阻碍方向而减小,当前进方向正对阻碍时受到的阻碍力是最大的。各向异性函数在描述重点人员行走时受到周围环境的影响力大小随着影响方向变化而变化,其计算如式(23)所示:

$$\begin{cases} An_{lin}(\varphi, \lambda) = \max\left\{1 - \lambda \cdot \frac{|\varphi|}{\pi}, 0\right\} \\ An_{cos}(\varphi, \lambda) = \lambda + (1 - \lambda) \cdot \frac{1 + \cos|\varphi|}{2} \\ An_{exp}(\varphi, \lambda) = \frac{1}{e^{\lambda \cdot |\varphi|}} \end{cases}, \quad (23)$$

式中:  $An_{lin}$  为线性各向异性函数,  $An_{cos}$  为余弦各向异性函数,  $An_{exp}$  为指数型各向异性函数,  $\varphi$  为相互作用方向与目标行走方向的夹角,  $\lambda$  为各向异性函数的特征参数。

在社会力模型中衰减函数用于描述力随距离增加而减少的情况,模拟移动对象之间交互力如何随距离的变化而变化。当重点人员靠近其他移动目标时所受到的排斥力会随着相距间隔的增大而减小,直至在一定距离外可忽略不计,这个距离为阈值距离,即力的作用距离。衰减函数分为指数衰减函数以及线性衰减函数,其表示如式(24)所示:

$$\begin{cases} f_{exp}(d, A, B) = Ae^{-Bd} \\ f_{lin}(d, d_0, M, \sigma) = \frac{M}{2d_0} \cdot (d_0 - d + \sqrt{(d_0 - d)^2 + \sigma}) \end{cases}, \quad (24)$$

式中:  $f_{exp}$  为指数衰减函数,适用于需要快速衰减影响力的场景,如紧急避让等情况;  $f_{lin}$  为线性衰减函数,适用于衰减速率较为平均的场景,如人员之间或是与障碍物的交互等情况;  $d$  为目标与其他对象之间的距离,参数  $A$  与  $B$  为衰减函数比例因子,  $A$  和  $M$  控制距离  $d=0$  时的影响力大小,  $B$  控制衰减函数的衰减速率,  $d_0$  为阈值距离,即影响力作用范围,  $\sigma$  为平滑参数,调节阈值距离时的平滑过渡。

#### (2) 移动对象虚拟轮廓构建

在复杂的拥挤环境下,为了更准确地模拟和预测移动对象的动态行为,建立移动对象的虚拟轮廓

至关重要<sup>[19]</sup>。移动对象的虚拟轮廓有助于模拟重点人员与行人之间的交互,避免碰撞以及确保整体流动的顺畅。通过依据行人与车辆的影响范围构建虚拟轮廓,可以更好地模拟各个移动对象在各种场景下的行为和轨迹,以提升重点人员轨迹预测的准确性。行人虚拟轮廓算法工作流程如算法 1 所示。

算法 1 行人虚拟轮廓算法

```

输入:重点人员位置序列  $M_i(x_i, y_i)$ , 周围行人位置序列  $M_{ped}(x_{ped}, y_{ped})$ 
输出:重点人员与周围行人的虚拟轮廓边界距离  $d_{i,ped}$ 
1.  $d_{i,ped} = \emptyset$ ;
2. for  $t$  in range(0, len( $M_{ped}$ ))
3.     for  $k$  in range(0, len( $M_{ped}^t$ ))
4.         for  $j$  in range(0, len( $M_{ped}^t$ ) - 1)
5.             if ( $x_{ped(j)}^t > x_{ped(j+1)}^t$ )
6.                  $M_{ped(j)}^t, M_{ped(j+1)}^t = M_{ped(j+1)}^t, M_{ped(j)}^t$ ;
7.             end if
8.         end for
9.     end for
10.    for  $j$  in range(0, len( $M_{ped}^t$ ))
11.         $R_{i,ped(j)}^t = k_{i,ped(j)}^t (\Delta l_i^t + \Delta l_{i,ped(j)}^t)$ ;
12.         $d_{i,ped(j)}^t = \sqrt{(x_{ped(j)}^t - x_i^t)^2 + (y_{ped(j)}^t - y_i^t)^2} - R_{i,ped(j)}^t$ ;
13.     $d_{i,ped} \leftarrow d_{i,ped(j)}^t$ ;
14.    end for
15. end for
16. return  $d_{i,ped}$ ;
    
```

算法复杂性分析:该行人虚拟轮廓算法具有时间复杂度  $O(nm(m-1))$ , 主要由 3 层嵌套循环引起,其中  $n$  表示重点人员数量,  $m$  表示周围行人数量。空间复杂度为  $O(n+m)$ , 体现了算法在存储重点人员与周围行人位置序列时所需的内存空间,反映了其在资源占用上的效率。

在实际情况下,拥挤环境下人体会存在碰撞,此时无法将移动对象视为粒子,人员之间的边界距离与其虚拟轮廓的半径大小有关,而虚拟轮廓的半径大小随人员移动速度的增大而增大,人员的移动速度又可用单位时间内的位移量表示,其计算如式(25)~式(27)所示:

$$d_{ij}^t = \sqrt{(x_j^t - x_i^t)^2 + (y_j^t - y_i^t)^2} - R_{ij}^t, \quad (25)$$

$$R_{ij}^t = k_{ij}^t \times (\Delta l_i^t + \Delta l_j^t), \quad (26)$$

$$\Delta l_i^t = \sqrt{(x_i^t)^2 + (y_i^t)^2} - \sqrt{(x_i^{t-1})^2 + (y_i^{t-1})^2}, \quad (27)$$

式中:  $d_{ij}^t$  为  $t$  时  $i$  与  $j$  之间虚拟轮廓的边界距离,视为  $i$  与  $j$  之间的距离,  $x_i^t$  与  $y_i^t$  为  $t$  时刻  $i$  位置的横坐标与纵坐标,  $x_j^t$  与  $y_j^t$  为  $t$  时  $j$  位置的横坐标与纵坐标。由于在实际拥挤情况中行人之间存在推挤,无法忽视体积,需要为行人构建虚拟轮廓,  $R_{ij}^t$  为  $t$  时  $i$  与  $j$  的虚拟轮廓半径之和,  $k$  为可调节参数,  $\Delta l_i^t$  为  $t$  时刻上一时间步内  $i$  的位移变化量。

### 3.2 社会力函数构建

社会力函数适用于计算复杂场景中重点人员移动时所受的影响力,模型采用社会力函数计算混合交通场景中周围行人、车辆以及障碍物对重点人员的影响,重点人员受力示意如图 10 所示。

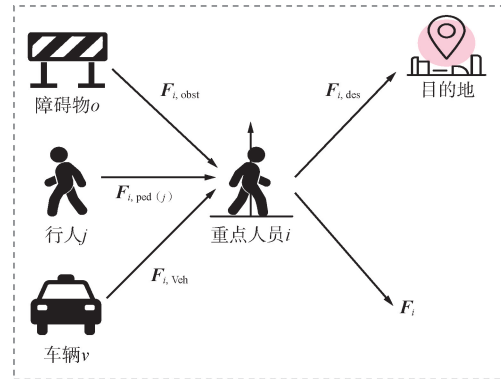


图 10 目标重点人员在移动时的受力示意

Fig. 10 Schematic diagram of the force of the target key individuals when moving

本小节主要研究移动对象交互场景,当忽略重点人员体积将其视为粒子时,社会力函数的基本形式如式(28)所示:

$$F_i^t = F_{i,ped}^t + F_{i,veh}^t + F_{i,obst}^t + F_{i,des}^t, \quad (28)$$

式中:  $F_i^t$  为  $t$  时刻  $i$  在移动时所受的总力,  $F_{i,ped}^t$  为  $t$  时刻  $i$  受其他行人的影响力,  $F_{i,veh}^t$  为  $t$  时刻车辆对  $i$  的影响力,  $F_{i,obst}^t$  为  $t$  时刻障碍物对  $i$  的影响力,这 3 种力通常表现为斥力,  $F_{i,des}^t$  是  $t$  时刻目的地对该  $i$  的力,通常表现为吸引力。

#### (1) 行人的影响力

重点人员与普通行人之间的相互作用力可表示为无接触的排斥力以及有接触的碰撞力,其计算如式(29)所示:

$$F_{i,ped}^t = F_r^t + F_c^t, \quad (29)$$

式中:  $F_r^t$  为  $t$  时刻  $i$  受周围行人的排斥力,  $F_c^t$  为  $t$  时刻  $i$  受周围行人的碰撞力。这些力的大小随距离的增加而减少,排斥力和碰撞力还具有不同的各向异

性,即力的方向随角度的变化而变化,所以这2种力的计算需要加入衰减函数和各向异性函数。

针对不同的交通场景,需要合理设计衰减函数和各向异性函数的形式和特征参数,衰减函数根据重点人员与周围行人之间的距离计算周围行人对重点人员的影响,影响大小随距离的增加而减小<sup>[20]</sup>。各向异性函数根据重点人员移动的方向与周围行人所在位置方向的夹角输出周围行人的影响,影响大小随着角度的增加而减少。经过分析,线性的各向异性函数和衰减函数更适合计算本场景中重点人员与行人之间的影响,并通过迭代优化确定最终函数特征参数,排斥力和碰撞力计算如式(30)~式(31)所示:

$$F_r^t = - \sum_{j \in N_{ped}} (f_{lin}(d_{ij}^t, d_r^0, M_r, \sigma_r) \cdot An_{lin}(\varphi_{ij}^t, \lambda_r) \cdot e_{ij}^t), \quad (30)$$

$$F_c^t = - \sum_{j \in N_{ped}} (f_{lin}(d_{ij}^t, d_c^0, M_c, \sigma_c) \cdot e_{ij}^t), \quad (31)$$

式中: $f_{lin}$ 为线性衰减函数, $An_{lin}$ 为各向异性函数, $e_{ij}^t$ 为方向单位向量, $d_{ij}^t$ 为 $t$ 时刻 $i$ 与 $j$ 相距的距离, $N_{ped}$ 为周围行人的总数量。

### (2) 车辆的影响力

车辆对重点人员的影响力与重点人员和车辆的虚拟轮廓边界距离相关,而车辆与重点人员之间的交互过程通常是短暂而迅速的,并且相较于重点人员和行人之间的交互来说,其危险性更大。针对人车交互环境的场景下,车辆对重点人员的潜在危险或影响力通常会随着二者之间的距离增加而快速减少,并且与重点人员正面相对的车辆可能比从侧面接近的车辆具有更高的危险性,因此使用指数衰减函数和余弦各向异性函数构建车辆对重点人员的影响力较为合适。基于车辆虚拟轮廓计算车对人的影响力 $F_{i,Veh}^t$ 表示如式(32)所示:

$$F_{i,Veh}^t = f_{exp}(d_{iv}^t, A_p, b_p) \cdot An_{cos}(\varphi_{iv}^t, \lambda_p) \cdot e_{iv}^t, \quad (32)$$

式中: $f_{exp}$ 为指数型衰减函数, $An_{cos}$ 为余弦各向异性函数, $e_{iv}^t$ 为 $t$ 时刻 $i$ 位置指向汽车位置的单位向量, $A_p$ 和 $b_p$ 为指数衰减函数的特征参数, $d_{iv}^t$ 为 $t$ 时刻 $i$ 与车辆轮廓之间的距离, $\varphi_{iv}^t$ 为 $t$ 时刻 $i$ 行走的方向与 $i$ 到车辆之间连线的夹角。

### (3) 障碍物的影响力

重点人员与障碍物的交互与距离和交互场景有关,通过引入障碍物的力能够更准确地模拟重点人员在复杂环境中的行为<sup>[21]</sup>。考虑障碍物的力可以帮助预测模型更好地理解重点人员如何导航避开障

碍物,如墙壁、树木或障碍物,这样可以提高轨迹预测的准确性和现实性。障碍物的力反映了重点人员为了避免与障碍物碰撞所产生的排斥效果,这种力对重点人员的最终轨迹有直接的影响,障碍物对 $i$ 的力 $F_{i,obst}^t$ 如式(33)所示:

$$F_{i,obst}^t = - \sum_{o \in N_{obst}} (f_{lin}(d_{io}^t, d_{obst}^0, M_{obst}, \sigma_{obst}) \cdot An(\varphi_{io}^t, \lambda_{obst}) \cdot e_{io}^t), \quad (33)$$

式中: $e_{io}^t$ 为 $t$ 时刻 $i$ 位置指向障碍物的单位向量, $d_{io}^t$ 为 $t$ 时刻 $i$ 与障碍物 $o$ 之间的距离, $\sigma_{obst}$ 为衰减参数,控制障碍物力随距离变化的速率, $\varphi_{io}^t$ 为 $t$ 时刻 $i$ 行走的方向与 $i$ 到障碍物 $o$ 之间连线的夹角。

### (4) 移动对象虚拟轮廓构建

目的地的吸引力反映了重点人员对特定方向的内在驱动力,这种力量能引导重点人员绕过障碍、避免碰撞,并最终到达预定的目的地<sup>[22]</sup>。在重点人员轨迹预测模型中,引入目的地吸引力能够使轨迹预测模型更加符合现实场景的行为模式。例如,重点人员通常不是漫无目的地随意行走,而是有特定的方向的,并且通常会朝着该方向选择最直接的路线前进,但如果路上有障碍,他们会选择避开障碍。因此,一个有效的轨迹预测模型不仅要考虑各移动对象之间的排斥力,也要考虑目的地的吸引力。目的地对 $i$ 的吸引力 $F_{i,des}^t$ 表示如式(34)所示:

$$F_{i,des}^t = f_{lin}(0, d_{des}^0, M_{des}, \sigma_{des}) \cdot An_{lin}(0, \lambda_{des}) \cdot e_{id}^t, \quad (34)$$

式中: $e_{id}^t$ 为 $t$ 时刻 $i$ 位置指向行人当前行走方向的单位向量, $\sigma_{des}$ 为衰减参数,决定了吸引力随着距离增加而减小的速率。

## 4 实验结果与分析

### 4.1 实验环境与数据集描述

实验中所有算法均在一致的软硬件设置中执行,以保障实验结果的可比性,详细配置参数如表1所示。

表1 实验环境配置

Tab. 1 Configuration of experimental environment

软硬件环境	配置
CPU	英特尔至强 W-2245 3.90 GHz
内存/GB	64
GPU	GeForce RTX 3090
显存/GB	24
操作系统	Ubuntu 22.04

本文选用了轨迹预测领域中常用的数据集 UCY 和 ETH,其中包含了 5 个子数据集:ZARA1、ZARA2、UNIV、ETH 以及 HOTEL。ZARA 系列数据集提供了商业街区内行人在高密度环境下的互动轨迹,主要关注行人在拥挤环境中的交互以及避让行为。UNIV 数据集则捕捉了校园环境 中行人在开放空间中自然行走的轨迹<sup>[23]</sup>。ETH 和 HOTEL 数据集分别反映了城市中心和酒店外部环境中行人的行为轨迹,为研究城市环境下的行人动态提供了丰富数据。其中,ZARA1、ZARA2、UNIV、ETH 以及 HOTEL 主要包含 1 586 个行人、57 个障碍物以及少量汽车。这些数据集提供了多样化的场景和行为模式,有助于全面评估模型性能。

#### 4.2 实验设置

在本文实验设置中,对数据集视频流统一采用了 0.4 秒/帧的采样频率提取各数据集中的轨迹信息,以确保信息跨度以及精确度的一致性。实验主要分为 2 种预测模式:基于目标前 8 帧轨迹数据预测该目标后 8 帧轨迹数据,以及基于目标前 8 帧轨迹数据预测该目标后 12 帧轨迹数据。第一种预测 8 帧轨迹点的预测模式本文将其视为短期预测,并将这种模式标记为  $obs = 8, pred = 8$  s。第二种预测 12 帧轨迹点的预测模式本文将其视为长期预测,并将这种模式标记为  $obs = 8, pred = 12$  s。

本文方法中 LSTM 学习率设置为 0.01,批处理大小为 64,训练轮数为 200。

#### 4.3 评价指标

在行人轨迹预测领域,性能评价指标主要用于衡量预测轨迹与真实轨迹之间的差异。实验共使用了 3 种性能评价指标:平均位移误差(Average Displacement Error, ADE)、最终位移误差(Final Displacement Error, FDE)以及准确度(Accuracy, ACC)。

ADE 为衡量预测轨迹与实际轨迹之间平均误差的指标,重点关注的是模型整体预测精度的水平,而不是整个预测过程中某个特定的时间点,其表示如式(35)所示:

$$ADE = \frac{\sum_{i=1}^N \sum_{t=obs+1}^{obs+pred} \| Y_i^t - \hat{Y}_i^t \|_2}{N \cdot pred}, \quad (35)$$

式中: $N$  为轨迹样本总数。ADE 数值越小则表明预测轨迹与真实轨迹越接近,即模型性能越好。

FDE 为衡量所有预测轨迹最后一个时间点的位置坐标与实际观测轨迹最后一个时间点的位置坐

标之间的平均误差,直观地反映了模型在长期预测模式中的准确度,其表示如式(36)所示:

$$FDE = \frac{\sum_{i=1}^N \| Y_i^t - \hat{Y}_i^t \|_2}{N}. \quad (36)$$

FDE 数值越小则表明最后一刻时间点预测的位置坐标与实际的位置坐标越接近。

ACC 为衡量模型预测精度的指标,用于计算所有预测轨迹坐标点相对于实际轨迹坐标点的准确率。实验中,若预测轨迹点与实际轨迹点之间的距离小于等于 0.5 m 则算成功预测,指示函数记为 1,反之则记为 0,其表示如式(37)~式(38)所示:

$$ACC = \frac{\sum_{i=1}^N \sum_{t=obs+1}^{obs+pred} hit(\| Y_i^t - \hat{Y}_i^t \|_2)}{N \cdot pred}, \quad (37)$$

$$hit(\| Y_i^t - \hat{Y}_i^t \|_2) = \begin{cases} 1, & \| Y_i^t - \hat{Y}_i^t \|_2 > 0.5 \\ 0, & \| Y_i^t - \hat{Y}_i^t \|_2 \leq 0.5 \end{cases}, \quad (38)$$

式中: $hit(\cdot)$  为指示函数,当预测点与真实点之间的距离小于等于 0.5 m 时,返回 1,否则返回 0。ACC 数值越大则说明预测位置与实际位置越接近,模型预测准确率越高。

#### 4.4 对比方法

为了评估 STMA 在轨迹预测方面的性能,本研究在同一实验框架下,对以下几类具有代表性的先进基准方法进行了比较分析:

时空网络(Spatio-Temporal Network, STNet)<sup>[24]</sup>利用图神经网络和树结构探索潜在路径,提供全面轨迹预测。社会生成对抗网络(Social Generative Adversarial Network, Social-GAN)<sup>[25]</sup>通过池化层和对抗训练捕捉重点人员与普通行人的互动,生成多模态轨迹。社会-物理注意力网络(Social and Physical Attention Network, SoPhie)<sup>[26]</sup>结合 GAN 和注意力机制,全面考虑社会关系和环境,生成真实轨迹。社交时空图卷积神经网络(Social Spatio-Temporal Graph Convolutional Neural Network, S-STGCNN)<sup>[27]</sup>结合时空图卷积网络,模拟重点人员与环境交互。条件生成神经网络(Conditional Generative Neural System, CGNS)<sup>[28]</sup>结合条件潜空间学习和软注意力,生成多样且可行的轨迹假设,并通过正则化方法确保样本可行性。

#### 4.5 实验结果

对比实验使用 ADE 和 FDE 两个指标进行性能

评估,这2个指标分别衡量预测轨迹与真实轨迹在时间平均和最终位置的误差<sup>[29]</sup>。根据输入的8帧

前序轨迹数据来预测目标接下来12帧的位置数据,其实验结果如表2所示。

表2 不同的评价指标比较不同模型的预测误差 ( $pred=12\text{ s}$ )

**Tab. 2 Prediction errors comparison of different models under different evaluation measurements ( $pred=12\text{ s}$ )**

评价指标	数据集	STNet	CGNS	SoPhie	Social-GAN	S-STGCNN	STMA
ADE	ETH	0.33	0.62	0.70	0.81	0.64	<b>0.26</b>
	HOTEL	<b>0.16</b>	0.70	0.76	0.72	0.49	0.25
	UNIV	0.32	0.48	0.54	0.60	0.44	<b>0.19</b>
	ZARA1	<b>0.23</b>	0.32	0.30	0.34	0.34	0.28
	ZARA2	<b>0.18</b>	0.35	0.38	0.42	0.30	0.23
AVG	ALL	0.29	0.52	0.58	0.60	0.46	<b>0.26</b>
FDE	ETH	0.47	1.40	1.43	1.52	1.11	<b>0.37</b>
	HOTEL	<b>0.25</b>	0.93	1.67	1.61	0.85	0.35
	UNIV	0.55	1.22	1.24	1.26	0.75	<b>0.42</b>
	ZARA1	<b>0.41</b>	0.59	0.63	0.69	0.53	0.54
	ZARA2	<b>0.34</b>	0.71	0.78	0.84	0.48	<b>0.34</b>
AVG	ALL	0.45	1.00	1.18	1.24	0.81	<b>0.41</b>

由于 ETH 数据集和 HOTEL 数据集的移动对象表现出了非线性的运动趋势以及较长的停留时间,且移动路线具有较强的多模态性,因此考虑到移动对象之间全局交互的模型,其表现相对于其他模型来说更为良好。UNIV 数据集中,行人密度较高,呈现群体性移动的现象,包含大量的群体行为和社交交互,因此能捕捉到复杂社会力交互的模型在该数据集上具有较好的性能<sup>[30]</sup>。ZARA1 数据集和 ZARA2 中的行人密度较低,并且场景为开阔商业街区,行人路径选择多变,其中在 ZARA1 数据集中,行人更为稀疏,基于图理论的方法在该情况下表现更优。在 ADE 和 FDE 方面,STMA 优于最先进的基线模型 STNet,预测性能平均提高了 11.5% 和 9.7%。可见,其通过将移动对象与情景之间的影响力进行社会力量化,以及运用时空多注意力机制对历史轨

迹时间步与空间距离信息加权,能够获得更全面的交互特征,尤其是在包含人车交互的场景中,STMA 相比于其他基线模型来说会具有更大的优势。

预测 ACC 的数值是衡量轨迹预测模型性能最直观的体现,本实验比较了不同模型在不同数据集中以  $pred=8\text{ s}$ 、 $pred=12\text{ s}$  两种不同预测模式下的预测精度,以了解模型在不同时间跨度下的预测能力。实验结果如图 11 和图 12 所示。可以看出,随着预测长度的增加,模型的预测精度普遍有所下降。

在短期预测模式下,本文模型 STMA 的平均预测准确率为 63.7%,比 SoPhie 高 23.9%,比 S-STGCNN 高 14.9%,比 CGNS 高 27.4%,比 STNet 高 5.5%,比 Social-GAN 高 20%。实验结果表明 STMA 在处理短期预测任务时具有显著的优势。

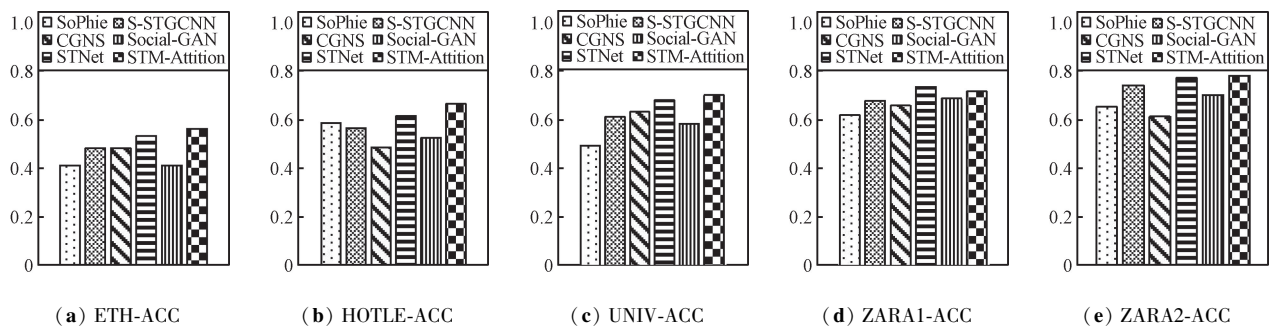


图 11 短期预测模式下各模型 ACC 对比 ( $pred=8\text{ s}$ )

Fig. 11 ACC comparison of different models in short-term prediction mode ( $pred=8\text{ s}$ )

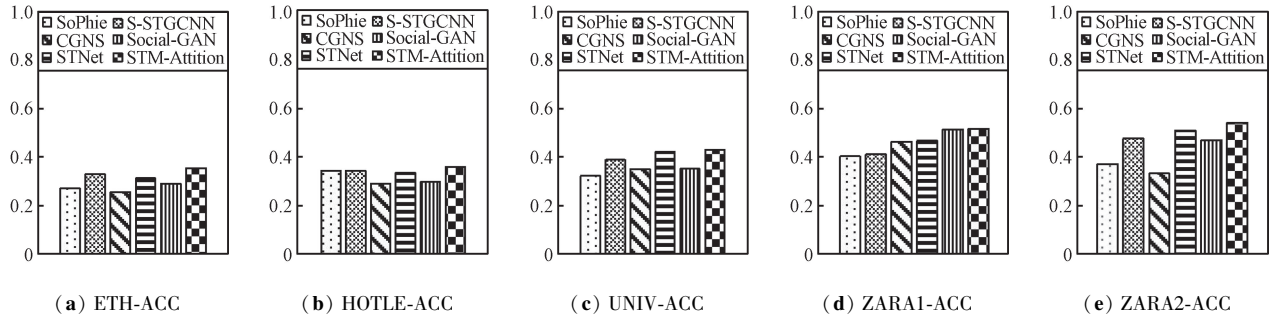


图 12 长期预测模式下各模型 ACC 对比( $pred=12\text{ s}$ )

Fig. 12 ACC comparison of different models in long-term prediction mode( $pred=12\text{ s}$ )

在长期预测模式下,STMA 模型的平均准确率为 54.3%,比 SoPhie 高 29.3%,比 S-STGCNN 高 13.4%,比 CGNS 高 36.8%,比 STNet 高 8%,比 Social-GAN 高 13.1%。由上述实验结果可知,STMA 模块在重点人员轨迹预测中充分提取了交互特征,情境感知模块也量化了环境信息对重点人员的影响力,有效地提升了模型预测准确率。

### 5 结束语

本文在重点人员轨迹预测方面进行了研究,为了提高移动对象轨迹中时空特征的捕获效率,提出一种基于 STMA 机制的重点人员轨迹预测模型,并设计了一种基于 STMA 模块的编解码器。此外,为了更加准确地评估目标重点人员在复杂交通场景中所受影响力,提出了一种情境感知驱动的时空多影响力轨迹预测方法。实验表明,本文所提方法在长短期预测准确率等方面均优于现有的轨迹预测模型。

未来工作可以在以下几个方面进一步优化:

- ① 融合多模态数据,如视频、雷达、激光等传感器的数据,以提高对移动对象运动状态的感知能力,进而提高轨迹预测准确性。
- ② 考虑更多的情景因素,比如注重交通规则、道路情况、天气等情景因素对移动对象运动轨迹的影响,以提高模型的适应性。
- ③ 提升预测效率以及实时性,满足多领域应用场景的需求。

#### 参考文献

[1] 乔少杰,金琨,韩楠,等.一种基于高斯混合模型的轨迹预测算法[J].软件学报,2015,26(5):1048-1063.  
 [2] PREVOST C G, DESBIENS A, GAGNON E. Extended Kalman Filter for State Estimation and Trajectory Prediction of a Moving Object Detected by an Unmanned Aerial Vehicle[C]//Proceedings of the American Control Conference. New York:IEEE,2007:1805-1810.

[3] 乔少杰,韩楠,朱新文,等.基于卡尔曼滤波的动态轨迹预测算法[J].电子学报,2018,46(2):418-423.  
 [4] QIAO S J, SHEN D Y, WANG X T, et al. A Self-adaptive Parameter Selection Trajectory Prediction Approach via Hidden Markov Models[J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 16(1):284-296.  
 [5] 乔少杰,李天瑞,韩楠,等.大数据环境下移动对象自适应轨迹预测模型[J].软件学报,2015,26(11):2869-2883.  
 [6] 夏卓群,胡珍珍,罗君鹏,等.不确定环境下移动对象自适应轨迹预测方法[J].计算机研究与发展,2017,54(11):2434-2444.  
 [7] 李万高,赵雪梅,孙德厂.基于改进贝叶斯方法的轨迹预测算法研究[J].计算机应用,2013,33(7):1960-1963.  
 [8] MIN K, KIM D, PARK J, et al. RNN-based Path Prediction of Obstacle Vehicles with Deep Ensemble[J]. IEEE Transactions on Vehicular Technology, 2019, 68(10):10252-10256.  
 [9] ALAHI A, GOEL K, RAMANATHAN V, et al. Social LSTM: Human Trajectory Prediction in Crowded Spaces [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Nevada: IEEE, 2016: 961-971.  
 [10] NIKHIL N, TRAN M B. Convolutional Neural Network for Trajectory Prediction [C]//Proceedings of the European Conference on Computer Vision (ECCV) Workshops. New York: ECCV, 2019: 186-196.  
 [11] NIU Z Y, ZHONG G Q, YU H. A Review on the Attention Mechanism of Deep Learning [J]. Neurocomputing, 2021, 452(10):48-62.  
 [12] LIANG Y B, ZHAO Z. NetTraj: A Network-based Vehicle Trajectory Prediction Model with Directional Representation and Spatiotemporal Attention Mechanisms [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(9):14470-14481.

- [13] CHEN K, SONG X, YUAN H T, et al. Fully Convolutional Encoder-Decoder with an Attention Mechanism for Practical Pedestrian Trajectory Prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11):20046-20060.
- [14] LI Z N, CHEN Z W, LI Y J, et al. Context-aware Trajectory Prediction for Autonomous Driving in Heterogeneous Environments [J]. Computer-Aided Civil and Infrastructure Engineering, 2024, 39(1):120-135.
- [15] ZHAO T Y, XU Y F, MONFORT M, et al. Multi-agent Tensor Fusion for Contextual Trajectory Prediction [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 12126-12134.
- [16] LIANG J W, JIANG L, NIEBLES J C, et al. Peeking into the Future: Predicting Future Person Activities and Locations in Videos [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 5725-5734.
- [17] BERENQUER A D, ALIOSCHA-PEREZ M, OVENEKE M C, et al. Context-aware Human Trajectories Prediction via Latent Variational Model [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 31(5): 1876-1889.
- [18] AZHAR M, ZAMAN F H K, TAHIR N M, et al. People Tracking System Using DeepSORT [C] // Proceedings of the 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE). New York: IEEE, 2020: 137-141.
- [19] HUANG Y J, DU J T, YANG Z R, et al. A Survey on Trajectory-prediction Methods for Autonomous Driving [J]. IEEE Transactions on Intelligent Vehicles, 2022, 7(3): 652-674.
- [20] 全浩宇, 张青苗, 赵军辉. 移动边缘计算中的智能任务迁移技术: 现状与展望 [J]. 无线电通信技术, 2022, 48(5): 804-811.
- [21] 李静宜, 丁飞, 张楠, 等. 基于深度 LSTM 与遗传算法融合的短期交通流预测模型 [J]. 无线电通信技术, 2022, 48(5): 836-843.
- [22] WANG S, BAO Z F, CULPEPPER J S, et al. A Survey on Trajectory Data Management, Analytics, and Learning [J]. ACM Computing Surveys (CSUR), 2021, 54(2): 1-36.
- [23] 黄峻, 田永林, 戴星原, 等. 基于深度学习的自动驾驶多模态轨迹预测方法: 现状及展望 [J]. 智能科学与技术学报, 2023, 5(2): 180-199.
- [24] HAN J, ZHENG H, CHEN D Z, et al. STNet: An End-to-End Generative Framework for Synthesizing Spatiotemporal Super-resolution Volumes [J]. IEEE Transactions on Visualization and Computer Graphics, 2021, 28(1): 270-280.
- [25] GUPTA A, JOHNSON J, LI F F, et al. Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks [C] // Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City: IEEE, 2018: 2255-2264.
- [26] SADEGHIAN A, KOSARAJU V, SADEGHIAN A, et al. SoPhie: An Attentive Gan for Predicting Paths Compliant to Social and Physical Constraints [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2019: 1349-1358.
- [27] XU Z F, LIU Y, CHU X, et al. Aircraft Trajectory Prediction for Terminal Airspace Employing Social Spatiotemporal Graph Convolutional Network [J]. Journal of Aerospace Information Systems, 2023, 20(6): 319-333.
- [28] LEGENSKY S, EDWARDS D, BUSH R, et al. CFD General Notation System (CGNS)-status and Future Directions [C] // Proceedings of the 40th AIAA Aerospace Sciences Meeting & Exhibit. Reno: AIAA, 2002: 752.
- [29] 毛焯, 陈宇麟, 严舒婷, 等. 结合重识别特征和运动预测的多目标跟踪方法 [J]. 无线电通信技术, 2023, 49(4): 740-745.
- [30] 魏凌, 杨宏伟. 民用无人机飞行轨迹测量方法及评估 [J]. 无线电通信技术, 2023, 49(4): 725-732.

#### 作者简介:

**相东升** 男, (2001—), 硕士研究生。主要研究方向: 轨迹预测、相似性检索、生物信息学。

**李成** 男, (1987—), 硕士, 副高级工程师。主要研究方向: 警用卫星定位系统应用、公安信息化建设。

**陈浩** 男, (1982—), 副高级工程师。主要研究方向: 警务视频大数据应用、警务地理信息系统应用。

**陈澄** 女, (1996—), 硕士。主要研究方向: 警务视频大数据应用、警务地理信息系统应用。

(\*通信作者)**李波** 男, (1984—), 硕士, 高级工程师。主要研究方向: 医疗大数据分析、医疗数据分级分类、人工智能。

**韩楠** 女, (1984—), 博士, 副教授。主要研究方向: 时空数据库、人工智能。

**谢添丞** 男, (1997—), 硕士。主要研究方向: 轨迹预测。

**杨春芳** 男, (1983—), 博士, 副教授。主要研究方向: 信息隐藏、多媒体智能理解、网络安全。

**乔少杰** 男, (1981—), 博士后, 教授, 博士生导师。主要研究方向: 轨迹预测、时空数据库、人工智能数据库。