

基于船舶运动行为与时序图神经网络的 轨迹预测研究

魏昊坤, 陈金勇, 刘敬一, 楚博策, 张文宝, 姜岩松, 郭琦, 裴新宇

(中国电子科技集团公司 第 54 研究所, 石家庄 050081)

摘要: 随着我国“海洋强国”战略的提出, 航运业飞速发展; 海上交通量迅猛增长, 碰撞事故频发, 同时积累了丰富的船舶航行数据, 亟需在此数据基础上, 对船舶的航行位置进行长时序的预测, 加强对海域的整体交通状况的认知, 降低船舶碰撞事故率; 为此在实验中首先对 AIS 数据进行预处理, 剔除其中的易于去除的异常点, 提出基于船舶航行特征的动态轨迹去纠缠方法去除纠缠点; 其次, 依据船舶航行特征提出顾及行为语义约束的时空轨迹密度自适应聚类方法对船舶运动模式进行挖掘, 得到船舶典型运动行为模式; 最后, 针对船舶轨迹以及船舶运动模式, 提出一种基于运动模式的时序图神经网络轨迹预测模型, 对轨迹进行长时序预测, 选取粤港澳大湾区作为实验海域, 经对比实验验证, 该模型在长时序预测上效果优于传统模型。

关键词: 船舶自动识别系统数据; 船舶轨迹预测; 轨迹预处理; 轨迹聚类; 图神经网络; 门控循环单元

Research on Trajectory Prediction Based on Ship Motion Behavior and Temporal Graph Neural Network

WEI Haokun, CHEN Jinyong, LIU Jingyi, CHU Boce, ZHANG Wenbao,

JIANG Yansong, GUO Qi, PEI Xinyu

(The 54th Research Institute of CETC, Shijiazhuang 050081, China)

Abstract: With the promotion of the “Marine Power” strategy in China, the shipping industry is rapidly developing. Rapid growth in maritime traffic leads to frequent ship collisions. There is an urgent need for the long-term prediction of ship trajectories based on accumulated ship navigation data to enhance the awareness of maritime traffic conditions and reduce collision rates. Firstly, the paper pre-processes the automatic identification system (AIS) data to eliminate easily removable outliers, and proposes a dynamic trajectory decorrelation method based on ship characteristics to remove decorrelation points. Then, a self-adaptive spatiotemporal trajectory clustering method with behavioral semantic constraints based on the characteristics of ship navigation is proposed to mine ship motion patterns and obtain typical ship behavior patterns. Finally, for the ship trajectories and motion patterns, a motion pattern-based temporal graph neural network model is presented. For long-term prediction of ship trajectories and patterns, the Guangdong-Hong Kong-Macao Greater Bay Area is selected as a test region, comparative experiments validate that the proposed model outperforms the traditional models in long-term prediction.

Keywords: AIS data; ship trajectory prediction; trajectory preprocessing; trajectory clustering; graph neural network; gated recurrent unit

0 引言

在 21 世纪, 海洋成为了人类发展的重要舞台。随着“海洋强国”、“交通强国”以及“海上丝绸之路”等国家战略的相继提出, 我国的航运产业迎来了蓬勃的发展。随着全球贸易的增加以及我国经济的崛起, 航运在我国的全球贸易中处于十分重要的地位, 据联合国贸易和发展会议 2023 年海运评书报告指出航运贸易占世界贸易的 80% 以上, 并且连年增长, 使得船舶的需求不断上升, 船舶逐渐

向智能化、大型化、多样化的方向的发展, 这就导致了海运航道上船舶数量的不断增加, 尤其是在在港口周边海域以及近岸海域水域承载负荷增大, 船舶碰撞事故频发, 据广东省海事局统计, 粤港澳大湾区周边海域, 船舶碰撞事故高发, 船舶碰撞事故数量在当月事故量中占比高达 21.7% 以上^[1], 图 1 为广东省 2023 年第三、四季度海上碰撞事故占比图。因此, 研究船舶未来轨迹减少船舶碰撞事故发生率逐渐成为热点研究方向。

收稿日期: 2024-05-28; 修回日期: 2024-07-07。

基金项目: 中国博士后科学基金项目(2021M703021); 河北省重点研发计划项目(22340301D); 河北省博士后基金项目(B2021003031)。

作者简介: 魏昊坤(1998-), 男, 硕士。

通讯作者: 陈金勇(1970-), 男, 硕士, 研究员级高工。

引用格式: 魏昊坤, 陈金勇, 刘敬一, 等. 基于船舶运动行为与时序图神经网络的轨迹预测研究[J]. 计算机测量与控制, 2023, 32(10): 39

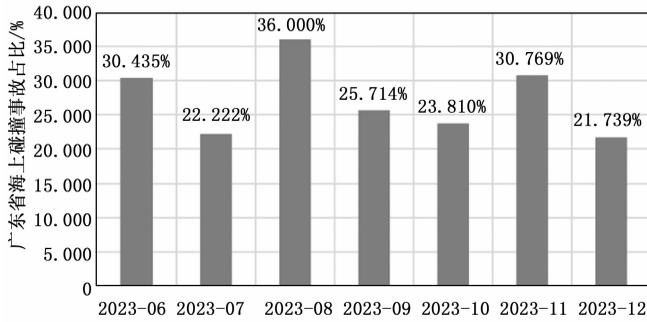


图 1 广东省 2023 年第三、四季度海上碰撞事故占比图

在船舶航行智能化方面,目前 100 吨以上的船舶必须配备船舶自动识别系统(AIS, automatic identification system)^[2], AIS 能够提供丰富的船舶海上航行数据,为海上船舶的交通管理以及轨迹预测的自动化、智能化提供了强大的数据支撑,充分利用 AIS 提供的船舶航行数据对于船舶避碰研究具有重要的指导意义。现有的对于船舶避碰的研究主要聚焦于船舶的轨迹预测,基于船舶的历史航行轨迹对未来的船舶运动轨迹进行预测,对于海域中可能出现的船舶碰撞事故做出提前预警^[3]。

目前对于海上船舶轨迹的预测主要是预测船舶在未来较短时间内的运动轨迹,由于海上船舶尺寸大,加之近岸区域船舶数量多,在短时间内难以做出迅速的航行调整,难以做出有效预警,因此亟需对船舶未来运动轨迹进行长时序预测。

海上船舶轨迹预测现有的方法主要分为 3 大类:基于传统的物理模型的方法^[4-6]、基于统计与机器学习的方法^[7-8]以及基于深度学习的方法。基于传统的物理模型的方法和基于统计与机器学习的方法在预测精度上基本能够满足一些船舶轨迹短时序的预测场景,但预测精度相对于基于深度学习的方法较低,在对海量轨迹数据的建模能力不足,对其轨迹点间的长时序依赖关系上不能挖掘出其中存在的复杂关系,并且面对海量 AIS 数据中所包含的船舶航行静态信息、动态信息等,不能够对其中的多船关联关系进行建模^[9-10]。但基深度学习的方法能够充分挖掘其中的关联关系,能够在轨迹的长时序预测中提高预测精度。

基于深度学习的船舶轨迹预测方法基本思想是采用深度神经网络模型对船舶的历史轨迹进行建模预测,由于海上船舶航行多变,船舶轨迹之间关系并非线性,复杂多样,利用其强大的特征提取能力和非线性性关系建模能力对船舶历史轨迹进行预测^[11-12],根据模型的架构可以将模型分为基于循环神经网络模型^[13-16]和基于图卷积神经网络模型的方法。基于循环神经网络(RNN, recurrent neural network)模型的船舶轨迹预测方法主要包含 RNN、长短期记忆网络(LSTM, long short-term memory)、门控循环单元(GRU, gated recurrent unit)及其变体。由于梯度消失的问题,RNN 通常仅具备短期记忆能力。为了解决这一问题,LSTM 模型应运而生,它是对 RNN 的改进与拓展,通

过专门的门控机制有效结合短期与长期记忆,从而在一定程度上缓解了梯度消失问题。在船舶轨迹预测领域,研究人员针对复杂的轨迹数据特点,提出了基于 LSTM 的集成模型,例如文献[17]结合了野生自枚举技术与 LSTM,学习船舶轨迹的时空模式,并应捕捉船舶的位置与空间和时间的关系对船舶未来时刻的位置信息以及异常流量进行预测。文献[18]设计了上下文感知 LSTM 模型,该模型可以考虑模型类型等上下文变量。该方法在保持逻辑一致性的同时,通过递归反馈环来提高预测精度。但是其预测误差在长时序预测上仍有不足,无法融合更高维度的附加知识,而图神经网络则能够在一定程度上解决这一问题。基于图卷积神经网络模型的方法主要利用图神经网络对船舶轨迹之间的关联关系进行特征提取位置预测。主要将轨迹视为图结构数据,轨迹点作为图节点,轨迹点的位置信息以及其他属性作为节点特征。文献[19]将周边对象的速度和属性标签构建带有外部知识的图邻接矩阵,利用时空图卷积神经网络(STGCN, spatio-temporal graph convolutional network)来学习运动特征和对象之间的交互特征来预测轨迹的分布。文献将内河中航行的船舶之间的交互作用力作为船舶之间的关联信息,从船舶之间的运动交互中,提取高精度特征用于生成船舶未来轨迹。文献[20]提出了 Social-STGCNN,它将轨迹建模为时空图,将对象作为顶点,将对象之间的交互力作为边,然后构建权重矩阵,以便更准确地预测轨迹并减少参数。此技术提高了预测精度,同时还加快了计算速度。为了更准确地预测单个轨迹。

针对以上问题,本文旨在将运动行为知识融入到船舶轨迹预测当中,利用时序图卷积网络多维特征的提取和学习能力,依据聚类结果对轨迹之间进行关联建模,进一步在轨迹关联关系上提取相同模式下的轨迹之间的特征交互,在时序上学习船舶轨迹点间的长期依赖关系,从多维度对船舶轨迹进行长时序预测,解决预测时间短、可解释性差的问题。

1 基于运动学约束的船舶轨迹处理

船舶自动识别系统(AIS)在船舶航行监控和管理中发挥着日益重要的作用。AIS 系统能够实时提供船舶的位置、速度、航向等关键信息,然而,在实际应用中,由于 AIS 数据可能受到多种因素的影响,如设备故障、信号干扰或数据传输延迟等,原始轨迹数据中往往存在异常点或数据缺失等问题,这直接影响了对船舶轨迹的准确分析和预测。因此首先要对轨迹中的异常点进行处理。异常点主要分为重复点、漂移点和纠缠点。

重复点是指那些在轨迹中出现多次的相同点或空间距离十分接近的点,这些点在数据接收或者发送的过程当中出现设备故障或者网络拥挤等问题造成数据的重复传输导致的,重复点的空间位置、对地航速和对地航向相同,UnixTime 时间戳不同。图 2 为重复点示意图。

漂移点是指那些在轨迹中明显偏离正常轨迹路径的点,

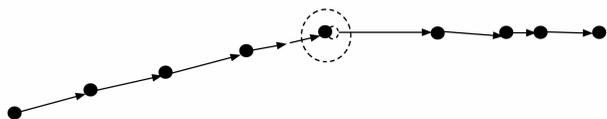


图 2 重复点示意图

这些点通常是由于 AIS 设备故障、解码错误或定位误差导致的。漂移点所在的航段的方向与前后正常航段存在明显的折断不连贯, 无法构成连续的轨迹。图 3 为漂移点示意图。

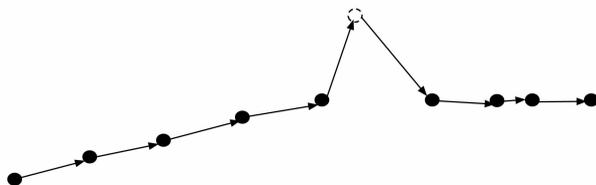


图 3 漂移点示意图

船舶的最大航行速度通常以“节”(knots)作为计量单位。不同类型的船舶, 由于船舶设计、推进系统功率以及船体形状的差异, 其最大航行速度存在一定区别。商业运输船舶的最大速度一般在 10~25 节范围内。部分高速客轮或军用舰船, 其主推进系统功率较大, 最大航速可以达到 30 节以上, 因此本实验中最大速度阈值为 30 节。

纠缠点是指那些在空间位置上保持正确但是在时间戳上排序混乱的点, 与漂移点、重复点不同之处在于纠缠点通常是在时间上时序相邻的多个点共同构成纠缠轨迹段, 以轨迹段的形式出现。纠缠点产生的原因是在 AIS 传输或接受的过程中部分出现异常或者错误, 导致点的信息在解码存入数据库使轨迹点在时序上存储出现混乱, 导致轨迹在地图上展现出来部分轨迹点出现纠缠反复的现象。其中船舶在航行时会按照 $P_0 \rightarrow P_1 \rightarrow P_2 \rightarrow P_3 \rightarrow P_4 \rightarrow P_5$ 的航行时序进行正常航行, 但是在图中由于轨迹点 P_2 和 P_3 的接收解码顺序不同会导致轨迹在轨迹点 P_2 和 P_3 之间发生轨迹点折返情况, 在轨迹结构上表现为轨迹缠绕的现象也就是本文所定义的纠缠现象。图 4 为正常轨迹示意图, 图 5 为纠缠轨迹示意图。去纠缠的整体流程如图 6 所示。

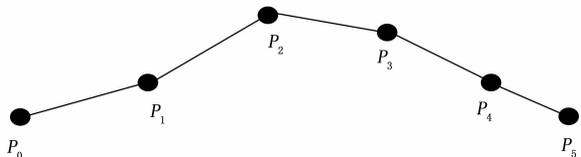


图 4 正常轨迹图

本文基于上述去除轨迹中明显的漂移点和重复点的方法, 对水上移动业务标识码 (MMSI, maritime mobile service identity) 为 100872924 船舶轨迹进行去除修复, 图 7 为轨迹中明显的漂移点能够很好地 将轨迹中偏离轨迹的点以及近岸重复点进行识别去除。

为去除轨迹中以直接去除的纠缠点, 本文选取 MMSI

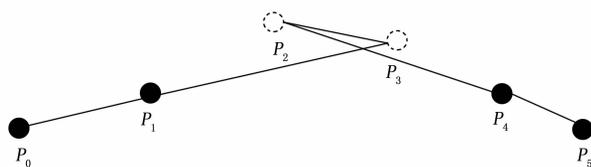


图 5 纠缠轨迹示意图

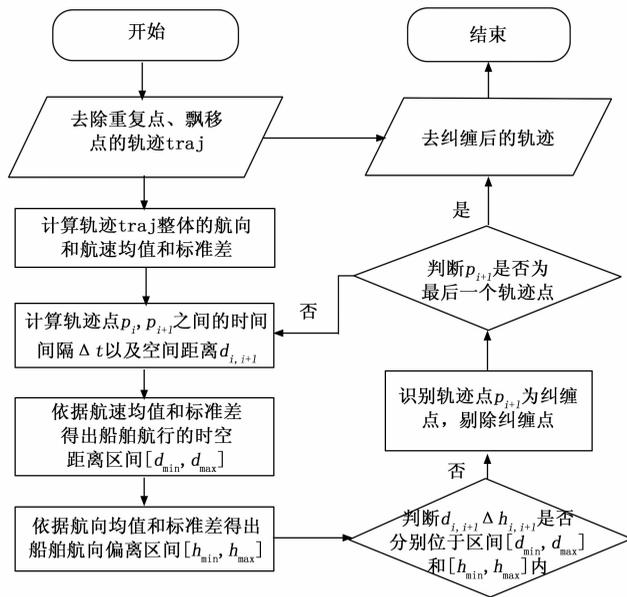
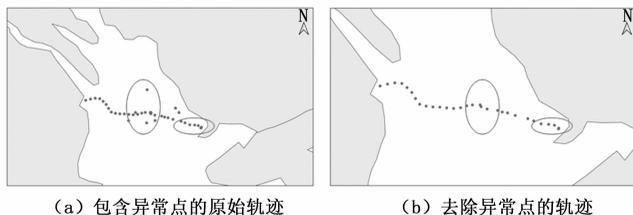


图 6 轨迹去纠缠整体流程图



(a) 包含异常点的原始轨迹 (b) 去除异常点的轨迹

图 7 轨迹异常点去除对比图

为 412475330 的轨迹进行去纠缠实验, 图 8 为原始轨迹中两部分的纠缠轨迹图, 经过本文所提方法对轨迹进行去纠缠之后得到的轨迹如图 9 所示, 对比两图中 I、II 区域部分, 可展现出本方法能够对其中反复纠缠的轨迹进行有效的去除, 在保证轨迹整体航向的基础上, 对杂乱纠缠轨迹进行处理, 形成高质量轨迹。

2 基于运动模式的船舶轨迹聚类挖掘

对船舶轨迹进行处理之后, 得到了高质量的船舶轨迹数据。在此数据基础上, 针对船舶的基本行为模式进行定义、识别, 将轨迹赋予语义维度上特征, 再将轨迹应用船舶轨迹聚类挖掘算法对轨迹进行聚类分析^[21], 揭示不同船舶在粤港澳大湾区特定海域的运动模式规律。

2.1 船舶运动行为语义定义

船舶的运动行为语义的定性描述以及定量表达是对船舶运动模式分析和挖掘的关键。本文结合民用船舶的共性

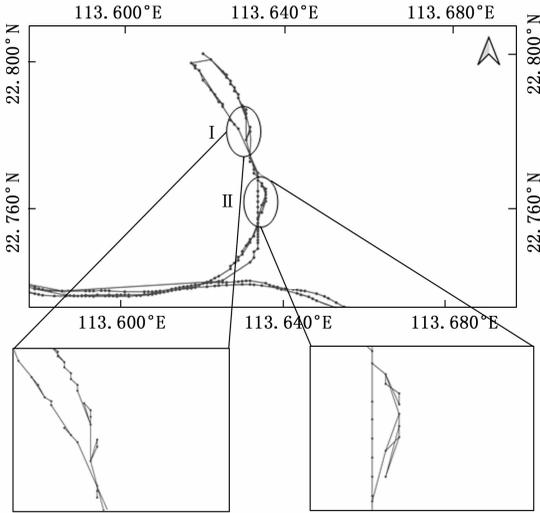


图 8 原始轨迹纠缠区域图

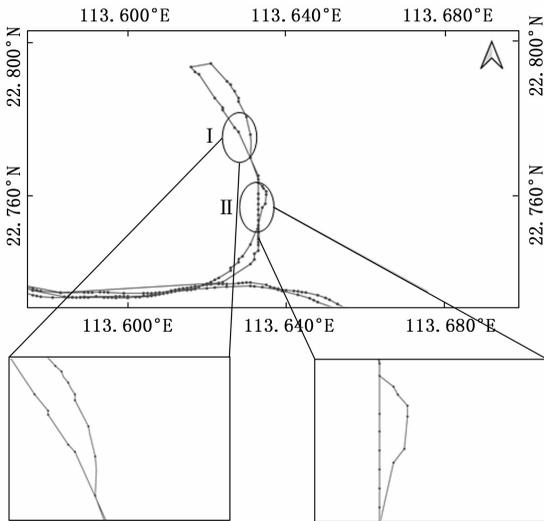


图 9 去纠缠效果图

航行特征, 定义开始航行、正常航行、绕行、折返、环绕、停留、结束航行 7 种船舶基本行为语义, 定义如下:

1) 开始航行: 指船舶从某一港口或锚地起航, 开始新的一段航程。主要位于港口或锚地附近, 轨迹起始部分集中在有限的港区范围内, 如图 10 (a) 所示。

2) 正常航行: 指船舶沿着预定航线以常规工作状态航行。具有该语义的轨迹点在空间位置上, 主要表现为各点之间空间位置不固定, 整体轨迹较为平直, 如图 10 (b) 所示。

3) 绕行: 指船舶由于障碍物 (如岛屿、浅滩等) 而改变航向进行绕行。具有该语义的轨迹点主要表现在轨迹点位于轨迹的中间且空间位置不固定, 取决于障碍物的具体位置和船舶的来向, 轨迹呈现一定的弯曲或偏转形状, 绕开障碍物的路径, 如图 10 (c) 所示。

4) 折返: 指船舶在航行过程中, 由于某些原因改变方向, 返回原出发点或中途某一点。具有该语义的轨迹点, 在空间位置上轨迹呈现明显的 U 形或类似曲线。在运动学

上, 折返行为表现为航向发生近 180 度的大角度转变, 航速可能会降低。在时间属性上, 折返通常持续一段时间处于轨迹中间时刻。当轨迹点满足上述定义, 则被标记为折返语义, 如图 10 (d) 所示。

5) 环绕: 指船舶在相对固定的区域内, 按圆形或类似轨迹环绕运动。具有该语义的轨迹点, 如图 10 (e) 所示。

6) 停留: 指船舶在某一固定位置长时间停驻。具有该语义的轨迹点, 在空间位置上, 在一定区域内轨迹比较集中, 轨迹变化很小, 如图 10 (f) 所示。

7) 结束航行: 指船舶到达目的地港口或锚地, 结束当前的一段航程。具有该语义的轨迹点, 轨迹的终点位于该目的地区域的特定地理位置, 如图 10 (g) 所示。

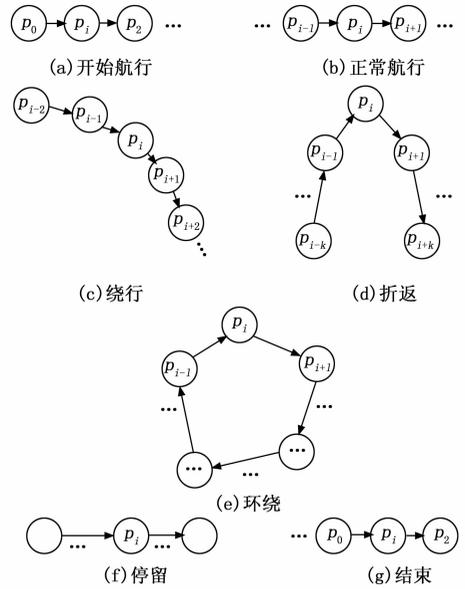


图 10 船舶行为语义示意图

2.2 船舶轨迹聚类挖掘

为了深入挖掘船舶轨迹数据中的潜在价值, 本文进一步开展了轨迹相似度计算和聚类分析。轨迹相似度计算是衡量不同轨迹间相似程度的关键手段, 其准确性直接影响到聚类分析的效果。通过相似度计算, 可以将船舶轨迹之间的关联程度进行定量表示, 通过时空轨迹聚类可以将行为特征相近的船舶轨迹聚集在一起, 形成不同的轨迹类别, 这对于分析船舶的航行模式、长时序的船舶轨迹预测具有重要意义。

本文针对各轨迹中主要从空间、行为语义和时间相似性 3 个维度对轨迹点之间的相似性进行度量, 其流程如图 11 所示。

2.2.1 轨迹点相似性度量

1) 空间相似性度量: 在计算轨迹点之间的相似度时, 本文采用计算不同船舶轨迹中任意两个轨迹点之间的空间距离的方法。首先, 将轨迹点的 WGS84 地理坐标系下的经纬度位置信息转为 Web 墨卡托投影下的真实坐标。然后, 利用计算出两轨迹点之间的空间距离。在此基础上, 设定

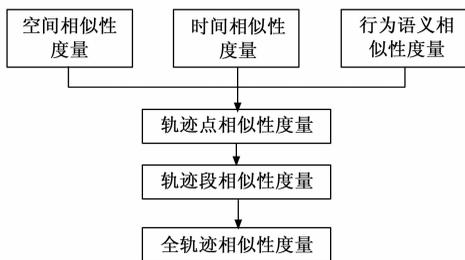


图 11 轨迹相似性度量流程图

空间距离阈值, 该阈值是通过统计粤港澳大湾区及相关内部航道宽度综合得出的。若两轨迹点之间的空间距离小于该阈值, 则判定这两个轨迹点在空间上相似。

2) 行为语义相似性度量: 在轨迹点进行语义赋值后, 每个轨迹点都包含相应的行为语义信息。本文采用简单的判断方式, 即若两个轨迹点具有相同的行为语义, 则判定它们相似; 反之, 则不相似。这种方法能够直观地反映船舶在航行过程中的行为特征。

3) 时间相似性度量: 船舶的航线受气候等多种因素影响, 呈现出周期性的航行行为。因此, 在判断轨迹点相似性时, 时间因素不可忽视。本文设定时间判定的阈值为两个月, 若两个轨迹点之间的时间差小于该阈值, 则判定它们在时间上相似。这种度量方法能够反映船舶在不同时间段内的航行规律。

2.2.2 轨迹段相似性度量

在轨迹点相似性度量的基础上, 本文进一步研究了轨迹段的相似性度量方法。通过计算由相邻轨迹点构成的轨迹段之间的相似度, 不仅能够考虑轨迹点之间的相关性, 还能顾及轨迹段之间的方向和运动趋势。这种度量方法能够更好地刻画船舶局部的航行特征。在判断轨迹段相似性时, 本文设定了前提条件, 构成两个相似轨迹段的两个起始对应轨迹点必须相似。

2.2.3 全轨迹相似性度量

全轨迹相似性度量是本文研究的最终目标。通过寻找不同轨迹中具有相同连续轨迹段的结构, 即最大相似子轨迹, 本文实现了对更为复杂的轨迹进行更抽象的表达。这种方法能够在理解轨迹整体模式的基础上对轨迹的相似度

进行度量, 从而更好地表达船舶在特定海域的航行模式, 揭示船舶的复杂航行行为。在计算最大相似子轨迹时, 本文首先计算两条轨迹的长度, 然后判断轨迹中相似轨迹段的连续结构, 最后通过计算包含连续轨迹段数最多的子轨迹作为最大相似子轨迹, 并利用 Jaccard 系数作为轨迹间的相似度量指标, 全轨迹相似性计算公式如下:

$$S(T_i, T_j) = \frac{s}{m+n-s} \quad (1)$$

其中: S 为相似度, T_i, T_j 分别为不同轨迹, s 为最大相似子轨迹长度, m 和 n 为两轨迹长度。

2.3 顾及行为语义约束的时空轨迹密度自适应聚类

基与上述时空语义融合方法, 构建出轨迹之间的相似度矩阵是船舶轨迹聚类的主要数据基础。聚类分析作为无监督学习的一种重要方法, 能够将具有相似性的船舶轨迹自动划分为同一类别, 从而揭示数据的内在结构和规律, 在众多聚类算法中, DBSCAN 以其对噪声和异常值的鲁棒性以及对其任意形状簇的适应能力而备受关注, 本文提出基于 DBSCAN 的顾及行为语义约束的时空轨迹密度自适应聚类方法, 对轨迹进行聚类, 聚类过程如下:

1) 根据轨迹的密度, 建立密度可达图。对于每条轨迹, 与其密度更高或者密度相等的邻居估计构成一个连接。这些连接可以形成一个或多个密度可达图;

2) 采用层次聚类的策略合并密度可达图, 构建聚类层次结构。通过计算轨迹之间的距离和密度的关系, 将轨迹分配到合适的聚类簇中。较高密度的轨迹被认为是聚类的核心轨迹, 较低密度的轨迹可以被连接到核心轨迹所在的聚类簇;

3) 轨迹簇的稳定性评估。通过对聚类结果进行多次采样和重复实验, 评估不同聚类簇的稳定性。稳定的聚类簇意味着在不同的采样数据上都能得到相似的聚类结果;

4) 标记噪声轨迹;

5) 得到粤港澳大湾区的轨迹模式分类。

本文选取南海数据中香港南部的外伶仃岛与南丫岛之间的船舶轨迹进行聚类测试, 数据空间范围为 $22.038 3 \sim 22.198 6^\circ N, 114.013 8 \sim 114.321 8^\circ E$, 时间范围为 2019 年 11 月 1 日至 2019 年 11 月 30 日。聚类结果包含 4 类典型运动模式, 如图 12 所示。

3 基于运动模式的时序图神经网络预测模型

3.1 模型架构

针对船舶轨迹长时序预测, 本文结合挖掘得到的粤港澳大湾区的运动模式提出了一种融合船舶历史运动模式的时序图神经网络轨迹预测模型。

为了学习数据集中相同运动模式下各轨迹间的关联关系, 本文采用 GCN 对其复杂拓扑关系特征进行提取。将当前船舶轨迹与历史模式中其他船舶轨迹之间的动态特征,

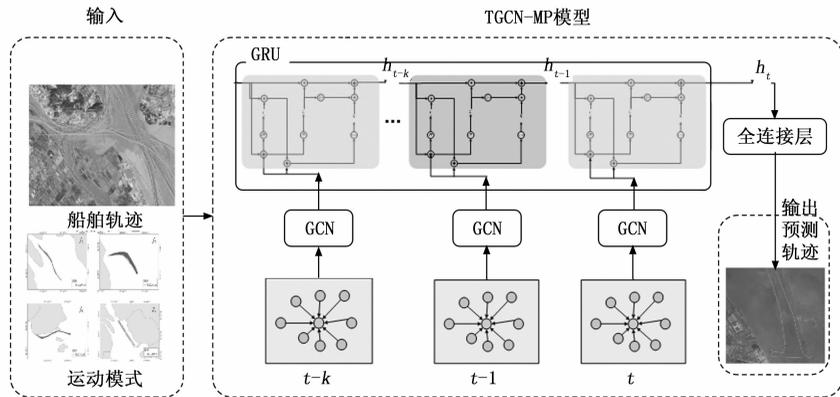


图 13 TGCN-MP 模型架构图

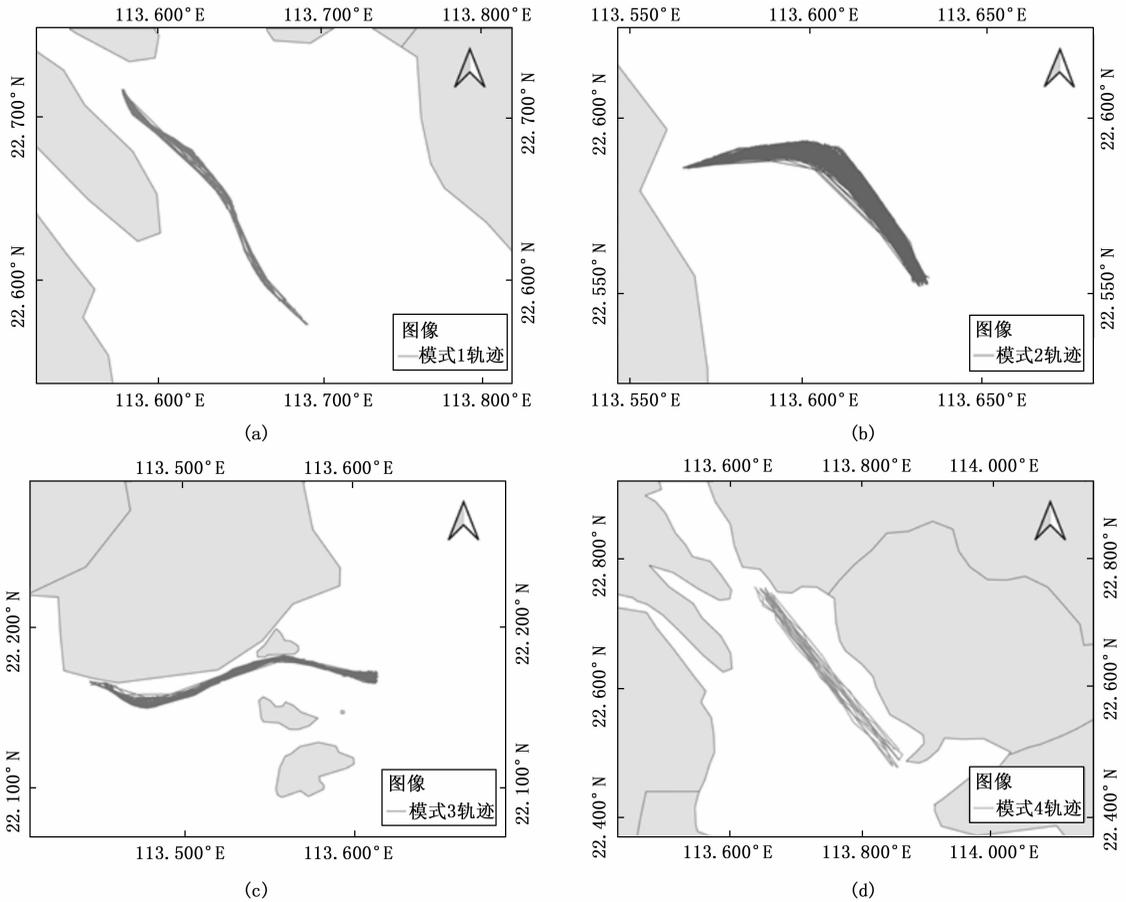


图 12 船舶轨迹聚类结果

如经纬度信息、对地航速、对地航向以及航行行为等，进行综合建模。通过 GCN 对这些时空关联特征进行建模和聚合，学习出不同海域、不同时刻下隐藏的轨迹关联规律。并且将相邻轨迹点之间聚合的特征作为该轨迹点的高维特征。为了将 GCN 提取的各轨迹点高维特征融合时序信息，采用 GRU 对图卷积网络学习到的复杂特征进行建模。GRU 能够从上一时刻的隐藏状态中学习到下一时刻多属性特征的演化规律，在关注当前船舶动态变化的同时，也能够保留其中长期的时序依赖信息。通过这种时序建模方式，本文所提出的模型能够从历史运动模式、时间序列特征和多轨迹之间的关联关系这三个关键维度对船舶未来轨迹进行全面预测。这种融合建模的方法，可以充分挖掘数据中隐藏的关联规律，为精准预测船舶轨迹提供有力支撑。

在图卷积神经网络模块，将每个时间片的邻接矩阵和各节点的特征值输入到 GCN 中利用一层卷积层，多节点之间的关系进行聚合提取：

$$f(\mathbf{X}, \mathbf{A}) = \text{LeakyRelu}(\hat{\mathbf{A}}\mathbf{X}\mathbf{W}_0) \quad (2)$$

其中： \mathbf{X} 代表特征矩阵， \mathbf{A} 代表邻接矩阵。

$$\hat{\mathbf{A}} = \tilde{\mathbf{D}}^{-1/2}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-1/2} \quad (3)$$

$\tilde{\mathbf{D}}$ 是图的度矩阵。

$$\tilde{\mathbf{D}} = \sum_j \mathbf{A}_{ij} \quad (4)$$

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I} \quad (5)$$

公式表示在图的结构中让每个节点针对于自身添加一个环路， \mathbf{I} 为单位矩阵。 \mathbf{W}_0 表示 GCN 的权重矩阵。LeakyRelu 为该层的激活函数。

在 GRU 模块中，分别计算该状态的隐藏门、更新门，获得船舶轨迹在运动上的依赖和特征，获取当前时刻 t 的隐藏态 h_t 。

$$u_t = \sigma(\mathbf{W}_u[f(\tilde{\mathbf{A}}, \mathbf{X}_t), h_{t-1}] + b_u) \quad (6)$$

$$r_t = \sigma(\mathbf{W}_r[f(\tilde{\mathbf{A}}, \mathbf{X}_t), h_{t-1}] + b_r) \quad (7)$$

$$h' = \tanh(\mathbf{W}_h[f(\mathbf{A}, \mathbf{X}_t), (r_t * h_{t-1})] + b_h) \quad (8)$$

$$h_t = u_t * h_{t-1} + (1 - u_t) * h'_t \quad (9)$$

3.2 评估指标

为了验证本文提出模型的有效性，并且与其他模型进行对比，本文提出利用 MSE ， ADE ， FDE ， ACC 作为评价指标，并且将 MSE 作为损失函数。

1) 均方误差 (MSE , mean square error):

$$MSE = \frac{1}{n} \sum_{i=1}^n (p_k - \hat{p}_k)^2 \quad (10)$$

其中： $p_i = (lon_i, lat_i)$ 代表真实轨迹经纬度， $\hat{p}_i = (\widehat{lon}_i, \widehat{lat}_i)$ 代表预测的轨迹点经纬度。

2) 平均距离误差 (ADE , average displacement error):

$$ADE = \frac{1}{n} \sum_{i=1}^n \sqrt{(p'_i - \hat{p}_i)^2} \quad (11)$$

其中： $p'_i = (x_i, y_i)$ 代表由经纬度转换到 Web 墨卡托投

影下的真实空间位置, $\hat{p}_i = (\hat{x}_i, \hat{y}_i)$ 代表预测的轨迹点在 Web 墨卡托投影下的真实空间位置。

3) 最终距离误差 (FDE, final displacement error):

$$FDE = \sqrt{(p'_n - \hat{p}'_n)^2} \quad (12)$$

4) 准确率 (ACC, accuracy):

准确率表示对船舶未来时刻预测的轨迹点在误差允许的范围内的比例, 其中允许的误差范围的阈值为 δ_{acc} , 此阈值能够直观的反应出轨迹预测的有效性, 能够直接表明在预测误差允许的范围船舶。允许范围的误差设置是参考海港总体设计规范^[22]中船舶的制动水域和船舶的回旋水域设计要求, 结合船舶自身长度和宽度得出, 船舶的最小应急范围为 4~5 倍的船长, 经统计, 本文选取的船舶轨迹长度区间为 50~450 m, 因此在综合考虑实验船长的情况下, 设定动态阈值, 每条船的依据其船长设定其对应的阈值。

$$\zeta = \frac{\lambda_i}{\lambda_p} \times 100\% \quad (13)$$

其中: ζ 为准确率, λ_i 表示预测点的误差小于 δ_{acc} 的数量, λ_p 为预测的点的总数。

MSE, ADE, FDE 的值越小, 代表预测值约接近真实值, ACC 的值越大, 表明预测轨迹点准确率越高, 可信度更高。

4 实验结果与分析

4.1 实验数据

AIS 数据中包含丰富的船舶航行信息, 例如水上移动通信业务标识码、船长、对地速度等船舶依据数据类型可分为 3 类, 静态信息、动态信息以及航程信息, 详细信息如表 1 所示。

表 1 船舶 AIS 数据组成

| 数据类型 | 字段 |
|------|---|
| 静态信息 | 水上移动通信业务标识码, 船舶名称, 船舶类型, 船长, 船宽 |
| 动态信息 | 船舶状态, 对地航速, 对地航向, 经度, 纬度, 转向率, 更新时间, UnixTime 时间戳 |
| 航程信息 | 吃水深度, 目的地, 预计到达时间 |

本实验选取 (105~125°E, 6~24°N) 区域内的轨迹进行实验, 如图 14 所示, 此范围涵盖了南海大部分区域以及台湾海峡, 时间范围为 2019 年 1 月到 2019 年 12 月, 包含货船、客船、捕捞船、拖轮、执法艇、油轮等 6 种船舶类型, 涉及在航、锚泊、失控、操纵受限、吃水受限、靠泊、搁浅、捕捞作业等 8 种船舶状态, 涉及到的船舶数量达到数万艘, 总体数据量达到 3 TB。

4.2 实验结果

本文对粤港澳大湾区船舶航行轨迹进行预测, 首先对模型与 GRU, Bi-GRU, CNN-GRU 三种典型预测方法在不同预测时长上进行对比, 寻找船舶轨迹预测的最长有效时

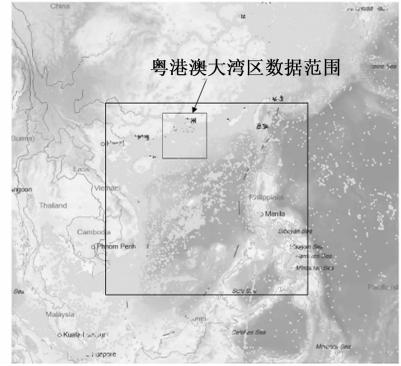


图 14 实验数据范围

间, 其中设定各模型训练参数相等, 输入轨迹的长度为 20, batchsize=1 024, 学习率 $lr=0.001$ 。结果如表 2 所示。实验中在与其他模型对比方面, 在预测时长为 10 min 时, TGCN-MP 模型在各项指标上优于 GRU 和 Bi-GRU 模型, 但是低于 CNN-GRU 模型。在预测时长为 20 min 时, TGCN-MP 模型在各项指标上仍优于 GRU 和 Bi-GRU 模型, 但是各项指标评分低于 CNN-GRU 模型, 模型之间的各项指标差距有所降低, 相较于与预测时长为 10 min 相比, TGCN-MP 模型与 CNN-GRU 模型平均距离误差的差距由 168 m 减小到了 88 m。在预测时长为 30 min 时, TGCN-MP 模型在各项指标上开始优于其他模型。在预测时长为 40 min 时, TGCN-MP 模型预测精度达到最优, 且各项指标优于预测时长为 30 min 的模型。但是在预测时长为 50 min 时, 本方法预测精度开始略微下降。

表 2 模型预测精度对比

| 预测时长 / min | 模型 | MSE | ADE/m | FDE/m | ACC/% |
|------------|---------|--------------------|--------------|--------------|--------------|
| 10 | GRU | 0.000 487 5 | 4 432 | 4 421 | 43.63 |
| | Bi-GRU | 0.000 318 2 | 2 784 | 2 169 | 69.90 |
| | CNN-GRU | 0.000 047 1 | 1 126 | 920 | 86.61 |
| | TGCN-MP | 0.000 0582 | 1 294 | 1 275 | 85.19 |
| 20 | GRU | 0.000 516 2 | 5 381 | 4 324 | 38.62 |
| | Bi-GRU | 0.000 458 5 | 3 074 | 3 514 | 54.32 |
| | CNN-GRU | 0.000 053 9 | 1 342 | 1 599 | 90.60 |
| | TGCN-MP | 0.000 066 1 | 1 412 | 1 954 | 87.54 |
| 30 | GRU | 0.000 674 3 | 8 736 | 7 325 | 32.72 |
| | Bi-GRU | 0.000 588 3 | 7 498 | 9 454 | 46.33 |
| | CNN-GRU | 0.000 078 3 | 1 739 | 2 014 | 88.60 |
| | TGCN-MP | 0.000 076 4 | 1 293 | 1 345 | 96.34 |
| 40 | GRU | 0.000 945 4 | 10 736 | 11 576 | 19.32 |
| | Bi-GRU | 0.000 348 3 | 8 498 | 10 634 | 75.32 |
| | CNN-GRU | 0.000 173 9 | 2 739 | 2 454 | 80.60 |
| | TGCN-MP | 0.000 068 3 | 1 194 | 1 071 | 98.73 |
| 50 | GRU | 0.001 077 8 | 12 736 | 12 437 | 12.74 |
| | Bi-GRU | 0.000 348 3 | 10 298 | 11 734 | 24.50 |
| | CNN-GRU | 0.000 273 9 | 3 467 | 4 453 | 75.45 |
| | TGCN-MP | 0.000 073 8 | 1 285 | 2 103 | 94.66 |

综上所述可得, TGCN-MP 模型在短时间预测能力上弱于轨迹预测的主流算法, 然而在预测时长为 30~40 min 时预测能力有所提升, 特别地, 在预测时长为 40 分钟时模型预测的平均位移偏差最小, 充分表明 TGCN-MP 模型能够更好地识别和利用数据的全局结构或周期性规律。在 40 min 的预测时长情况下, 模型更容易捕捉轨迹的典型运动模式规律, 从而使得其预测精度相较于其他时长, 表现更为出色。此外, TGCN-MP 模型能够综合船舶运动模式、轨迹关联关系, 展现出了良好的适应性, 能够很好的适应于长时序的轨迹预测。

5 结束语

随着全球航运贸易的繁荣发展, 海域承载负荷逐渐增大, 船舶碰撞事故频发。基于 AIS 的船舶轨迹预测能够对船舶的轨迹进行长时序的预测, 降低船舶碰撞事故率。本文提出的基于 AIS 数据预处理以及基于船舶运动行为的轨迹挖掘分析方法能够得到高质量的 AIS 数据, 并且能够挖掘出船舶的特定航行规律。在此数据基础之上, 设计了基于运动模式的时序图神经网络预测模型, 该模型不仅能够捕捉船舶与历史运动模式之间的关联, 还能够分析轨迹在时间维度上的依赖关系, 实现轨迹的长时序精准预测, 并且相比于传统模型效果最优。

本实验对船舶之间的挖掘分析对船舶和周边环境的交互信息较少, 并且在预测模型设计时, 网络特征的提取能力仍由优化提升的空间。因此, 未来我们将致力于研究如何利用更全面、更丰富的信息挖掘到船舶潜在的、更为典型的运动模式, 引入更先进的图神经网络结构, 增强船舶之间的关系特征提取能力, 或者结合其他机器学习方法, 如深度学习、强化学习等, 以进一步提高预测精度和效率。

参考文献:

- [1] 中华人民共和国广东海事局. 2023 年 12 月海上接警处置和险情分析 [R/OL]. (2024-01-09) [2024-01-21]. <https://www.gd.msa.gov.cn/2e8964e1ecc1c7dbc7db1c4555003600/10dacla23f9f4f9999b43bfe744cb52a.jhtml>.
- [2] 陈星森. 中国海事局印发《国内航行船舶船载电子海图系统和自动识别系统设备管理规定》[J]. 中国海事, 2011 (7): 75.
- [3] 陈柏果. 基于 AIMM-SCKF 的船舶轨迹预测和状态估计 [D]. 大连: 大连海事大学, 2022.
- [4] ABDELAAL M, FRENZLE M, HAHN A. Nonlinear model predictive control for trajectory tracking and collision avoidance of underactuated vessels with disturbances [J]. Ocean Engineering, 2018, 160: 168-180.
- [5] ZHANG Z, YIN J, WANG N, et al. Vessel traffic flow analysis and prediction by an improved PSO-BP mechanism based on AIS data [J]. Evolving Systems, 2019, 10 (3): 397-407.
- [6] RONG H, TEIXEIRA A P, SOARES G C. Ship trajectory uncertainty prediction based on a Gaussian Process model [J]. Ocean Engineering, 2019 (182): 499-511.
- [7] TONG X P, XU C, SANG L, et al. Vessel trajectory prediction in curving channel of inland river [C] // International Conference on Transportation Information & Safety. IEEE, 2015.
- [8] 王腾飞. 混合场景下船舶碰撞风险识别与避碰智能算法研究 [D]. 武汉: 武汉理工大学, 2020.
- [9] FENG H, CAO G, XU H, et al. IS-STGCNN: An Improved Social spatial-temporal graph convolutional neural network for ship trajectory prediction. [J]. Ocean Engineering, 2022, 266 (3): 1-15.
- [10] ATLURI G, KARPATNE A, KUMAR V. Spatio-temporal data mining: a survey of problems and methods [J]. ACM Computing Surveys, 2018, 51 (4): 83-95.
- [11] MA H, ZUO Y, LI T. Vessel navigation behavior analysis and multiple-trajectory prediction model based on AIS data [J]. Journal of Advanced Transportation, 2022 (1): 6622862.
- [12] MA J, JIA C, SHU Y, et al. Intent prediction of vessels in intersection waterway based on learning vessel motion patterns with early observations [J]. Ocean Engineering, 2021, 232: 109154.
- [13] WANG C, ZHU M, OSEN O, et al. AIS data-based probabilistic ship route prediction [C] // 2023 IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). IEEE, 2023, 6: 167-172.
- [14] LI W, ZHANG C, MA J, et al. Long-term vessel motion prediction by modeling trajectory patterns with AIS data [C] // International Conference on Transportation Information and Safety (ICTIS). IEEE, 2019: 1389-1394.
- [15] MURRAY B, PERERA L P. Ship behavior prediction via trajectory extraction-based clustering for maritime situation awareness [J]. Journal of Ocean Engineering and Science, 2022, 7 (1): 1-13.
- [16] TIAN L, JIAN M. Ship navigation behavior prediction based on AIS Data [J]. IEEE Access, 2022, 10: 47997-48008.
- [17] VENSUS J, TREIGYS P, MARKEVICIUTE J. Unsupervised marine vessel trajectory prediction using LSTM network and wild bootstrap techniques [J]. Nonlinear Analysis: Modelling and Control, 2021, 26 (4): 718-737.
- [18] MEHRI S, ALESHEIKH A A, BASIRI A. A contextual hybrid model for vessel movement prediction [J]. IEEE Access, 2021, 9: 45600-45613.
- [19] RAINBOW, B, MEN Q, SHUM H. Semantics-STGCNN: a semantics-guided spatial-temporal graph convolutional network for multi-class trajectory prediction [C] // IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2021: 324-331.
- [20] ABDUALLAH M, KUN Q, MOHAMED E, et al. Social-STGCNN: a social spatio-temporal graph convolutional neural network for human trajectory prediction [C] // Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020: 527-535.
- [21] 董婉婷. 基于 AIS 的渔船轨迹相似性分析技术研究 [D]. 大连: 大连海洋大学, 2020.
- [22] JTS 165-2013. 海港总体设计规范 [S]. 中华人民共和国交通运输部, 2013.11.12.