

船舶活动知识图谱可视分析方法

刘建湘^{1,2)}, 陈晓慧^{2)*}, 刘海砚²⁾, 张兵²⁾, 叶林²⁾, 易万潇³⁾, 刘涛²⁾

¹⁾ (中国人民解放军 93116 部队 沈阳 110000)

²⁾ (中国人民解放军战略支援部队信息工程大学数据与目标工程学院 郑州 450000)

³⁾ (中国人民解放军 61578 部队 十堰 442000)

(cxh_vrlab@163.com)

摘要: 随着全球海事贸易的高速发展, 对船舶活动进行可视分析的需求越来越大, 用户在没有非常专业的知识背景下, 难以快速地选择合适的可视化方法进行船舶活动分析, 面临“冷启动”问题. 为此, 提出船舶活动知识图谱可视分析方法. 基于多源异构的船舶活动数据, 构建包含时间、地点、船舶和活动 4 类船舶活动要素的船舶活动知识图谱, 并结合船舶活动分析需求, 设计船舶对象特征分析、船舶活动事件分析、船舶活动关联分析和船舶活动时空分析 4 种分析任务; 在船舶活动知识图谱的基础上, 提出基于活动要素的船舶活动可视分析方法, 并结合过滤、关联和导航等交互方式, 设计船舶活动知识图谱可视分析模型. 以 AIS 数据和“沸腾环贸”海上事件数据作为船舶活动数据, 基于所提出的船舶活动知识图谱可视分析模型, 开发了船舶活动知识图谱可视分析原型系统, 实验结果表明, 该系统可以为用户推荐相对合适的可视化方法, 为船舶活动的分析提供更多的视角; 也验证了如禁渔期等特殊情况下造成船舶活动规律变化等结论.

关键词: 船舶活动; 船舶事件; 知识图谱; 时空活动; 可视分析

中图法分类号: TP391.41

DOI: 10.3724/SP.J.1089.2024.19918

Research on Visual Analysis Method of Ship Activity Knowledge Graph

Liu Jianxiang^{1,2)}, Chen Xiaohui^{2)*}, Liu Haiyan²⁾, Zhang Bing²⁾, Ye Lin²⁾, Yi Wanxiao³⁾, and Liu Tao²⁾

¹⁾ (93116 Unit of the PLA, Shenyang 110000)

²⁾ (School of Data and Target Engineering, PLA Strategic Support Force Information Engineering University, Zhengzhou 450000)

³⁾ (61578 Unit of the PLA, Shiyan 442000)

Abstract: With the rapid development of global maritime trade, the demand for visual analysis of ship activities is growing. Without very professional knowledge, users are difficult to quickly select appropriate visual methods to analyze ship activities and face the problem of “cold start”. Based on multisource heterogeneous ship activity data, this paper constructs a ship activity knowledge graph including four types of ship activity elements of “time, place, ship and activity”. Combining with the needs of ship activity analysis, four analysis tasks are designed, including ship object feature analysis, ship activity event analysis, ship activity correlation analysis and ship activity space-time analysis; Besides, based on the knowledge graph of ship activities, the visual analysis method of ship activities based on activity elements is designed, and the visual analysis model of ship activity knowledge graph is designed combined with the interactive modes of filtering, association, and navigation. In this paper, the AIS data and the “boiling ring trade” maritime event data are used as the ship ac-

收稿日期: 2022-08-18; 修回日期: 2022-11-12. 基金项目: 国家自然科学基金(41801313, 41901397, 42171456). 刘建湘(1997—), 男, 硕士研究生, 主要研究方向为时空数据可视化; 陈晓慧(1983—), 女, 博士, 副教授, 论文通信作者, 主要研究方向为时空数据可视化; 刘海砚(1971—), 男, 博士, 教授, 博士生导师, 主要研究方向为数字地图制图、时空数据挖掘; 张兵(1995—), 女, 硕士, 主要研究方向为数据可视化; 叶林(1995—), 男, 博士研究生, 主要研究方向为人工智能; 易万潇(1991—), 男, 硕士研究生, 主要研究方向为网络安全; 刘涛(1999—), 男, 硕士研究生, 主要研究方向为时空数据挖掘.

tivity data. Based on the above model, a prototype system for visual analysis of the ship activity knowledge graph is developed. The results show that the system can recommend relatively appropriate visual methods for users, provide more perspectives for the analysis of ship activities, and also verify the conclusions such as the changes in the laws of ship activities caused by special circumstances such as the closed fishing periods.

Key words: ship activities; ship events; knowledge graph; spatiotemporal activities; visual analysis

随着全球海事贸易的高速发展, 船舶数量近年来持续增长, 海上船舶活动日益频繁, 事故时有发生^[1]. 为了保障海上交通安全、高效运转, 各国海事主管机构和海岸防卫部门联合建立了船舶自动识别系统(automatic identification system, AIS), 以增强用户对全球海上交通的综合态势感知能力^[2]. 运行 30 年来, AIS 积累了海量的船舶定位数据, 还有船讯网^①和沸腾环贸^②等与海上交通事件有关的公开数据源, 为海上船舶活动分析提供了巨大的数据驱动型技术创新空间.

许多研究人员将大数据分析技术应用于船舶活动分析, 助力海上船舶活动的分析向智能化发展. 当前, 对于船舶活动规律的分析大都是挖掘一般性特征, 如船舶活动季节性特征和船舶流量的时空分布特征等; 可视分析方法主要基于密度的船舶活动可视化、基于时空立方体的船舶活动可视化和基于核密度估计的船舶活动可视化等方面, 缺少一个完整的分析流程或者比较全面的分析模型, 人们在没有太多的专业知识情况下, 难以快速地选择比较合适的可视化方法进行深度分析, 面临“冷启动”问题.

本文通过知识图谱^[3]融合多源船舶活动数据, 将船舶活动表达为活动、船舶、时间和地点 4 个核心组成要素, 并结合船舶活动分析需求, 总结出 14 种船舶活动分析场景; 基于上述 4 个船舶活动组成要素设计船舶活动可视分析模型, 以支持对不同场景的可视分析方法进行自动推荐, 包括船舶活动空间可视分析方法、船舶活动时序可视分析方法、船舶活动时空可视分析方法和船舶活动属性可视分析方法等.

1 相关工作

近年来, 已有很多船舶活动可视分析方法, 分为基于密度的可视分析、基于时空立方体的可视分

析和基于关系图的可视分析 3 类.

1.1 基于密度的船舶活动可视分析

Scheepens 等^[4]利用密度图可视化方法描述船舶活动轨迹的空间分布特征. 密度图通过计算船舶的数量或船舶轨迹点的数量获得密度信息; 然后使用颜色映射表示每个空间单元的密度值, 根据密度计算的方式, 将密度图的生成方法分为直接绘制法^[5-7]、基于网格的方法^[8-10]和基于核密度估计的方法^[11-12]. 直接绘制法易于实现, 但依赖于人类观察者的视觉观察来估计轨迹点的数量并得到密度分布, 常常导致轨迹点的严重重叠, 并且很难实现精确的直观估计. 具有均匀网格的密度图易于实现和理解, 但网格的分辨率很难确定, 特别是对于一个差异很大的分布; 使用非均匀网格的密度映射可以对重要的小尺度区域使用细粒度网格, 对非重要区域使用粗粒度网格来提供细节分析粒度级别. 核密度估计是一种常用的估计数据概率密度分布的方法, 与网格划分步骤或目标数据分布的任何先验知识无关^[13], 该方法可以使用各种核函数(如均匀核函数和径向基核函数)对不同场景进行密度估计. 基于核密度估计的密度图是显示船舶密度或船舶活动轨迹密度分布的主要手段. 密度图可以作为一种空间快照, 能够有效地反映船舶活动轨迹的空间分布, 但不能反映船舶活动轨迹的时间特征.

1.2 基于时空立方体的船舶活动可视分析

时空立方体在船舶轨迹可视化中同时考虑空间和时间的特征. 时空立方体^[14]最初用于分析移动的人或物体. Andrienko 等^[13]首次将时空立方体应用于船舶活动轨迹可视化, 以探索船舶活动轨迹之间的时空关系; Bach 等^[14]对现有的时空立方体技术进行系统的调查, 还提出一个描述性框架, 指导适当选择空间-时间立方体技术的目标分析任务. Amini 等^[15]指出, 时空立方体的主要缺点是三维立方体会带来视觉遮挡, 用户必须多次旋转和

① <http://www.shipxy.com>

② <http://www.gogotrade.info>

缩放一个三维立方体,以选择特定的轨迹或识别特定的时空模式,并找到一个满意的视点.为了解决这个问题, Demšar 等^[16]将核密度估计方法与时空立方体相结合,用不同的颜色区分海上交通的峰谷信息; Tominski 等^[17]将船舶活动轨迹可视化叠加的三维轨迹带,减少时空立方体中轨迹的视觉遮挡,并使用颜色表示轨迹属性,如速度、转向角度等,以提高时空立方体的信息表示能力.

1.3 基于关系图的船舶活动可视分析

相关分析在船舶航迹分析中起着至关重要的作用,如探索船舶航速与气象条件之间的关系,研究港口之间的货运关系等.因此,许多为关系表示量身定制的可视化方法经常用于船舶轨迹分析.平行坐标图是一种经典的多维数据可视化方法^[18-19],可以帮助用户探索船舶航行与环境条件之间的内在联系. Lu 等^[20]将船舶轨迹点作为多维数据,并应用平行坐标图分析船舶类型、吨位、速度和其他属性之间的关系; Lundblad 等^[7]使用平行坐标图分析航行和气象参数之间的关系.流地图与和弦图也被广泛用于可视化一对多或多对多的关系^[21-23],它们在海上运输管理中已用于分析船舶运输网络、航行周期等海上交通管理的有价值的信息.

2 船舶活动知识图谱构建与需求概述

2.1 多源船舶活动数据预处理

本文涉及的多源船舶活动数据是有关船舶活动的动态数据,包括描述船舶活动轨迹的 AIS 数据、描述船舶事件的文本数据和与船舶航行区域相关的地理矢量数据.除了地理矢量数据可以通过 OSM(OpenStreetMap)直接获取并入库外,其他数据均需要进行预处理后再入数据库.

2.1.1 AIS 数据预处理

本文中的 AIS 数据是 2020-01-01—2021-09-30 的全球轨迹数据,包含近 2.5 亿条记录. AIS 采用唯一识别码 MMSI 对船舶进行编码,包括船舶的静态数据信息、船舶的动态轨迹信息和航次等相关信息共 30 个字段. AIS 数据具有海量性、高维度、动态性、时空相关性、不确定性和知识密度低等特征;其中,存在空值、坐标缺失和数据不符合规范的各种问题. 本文将 AIS 数据导入 PostgreSQL 数据库后,从去除经纬度空值的数据、去除时间空值的数据、过滤掉不需要的字段和去除静态属性冗余 4 个方面进行清洗,其流程如图 1 所示.

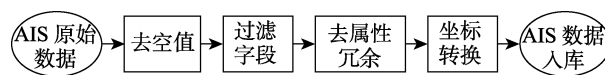


图 1 AIS 数据预处理流程

2.1.2 事件文本数据预处理

“沸腾环贸”网站的主页提供了全球的船舶事故和海盗事件的报道,其中,详情字段描述船舶事故和海盗事件的船舶名称、发生事故的时间、地点和伤亡情况等. 本文使用 Python 语言编写爬虫对该网站的船舶事故和海盗事件进行爬取,得到 460 条船舶事故记录和 537 条海盗事件记录.

针对文本描述类的数据进行事件关系抽取时,可以采用深度学习模型提取文本的词汇级或句子级特征,前提是将文本数据进行预处理,将其表示为特征向量.除了进行中文分词和去停用词等处理,本文通过 Skip-Gram 模型获取词向量. Skip-Gram 模型是一种 word2vortor 的模型,输入当前的词来预测上下文的词,通过 embedding 映射单词与向量,解决因向量稀疏而导致计算量大的问题.

2.2 船舶活动知识图谱构建

本文对船舶活动知识图谱定义如下:以图的形式表达船舶时空活动及其之间的关系的知识库,其中,船舶活动实体为节点,船舶活动实体间的关系为边. 本文采用自顶向下和自底向上相结合的方式构建船舶活动知识图谱,包括知识建模、知识存储、知识抽取和知识应用,其结构如图 2 所示.其中, Protégé 软件是斯坦福大学医学院生物信息研究中心开发的本体编辑和知识获取软件;事件

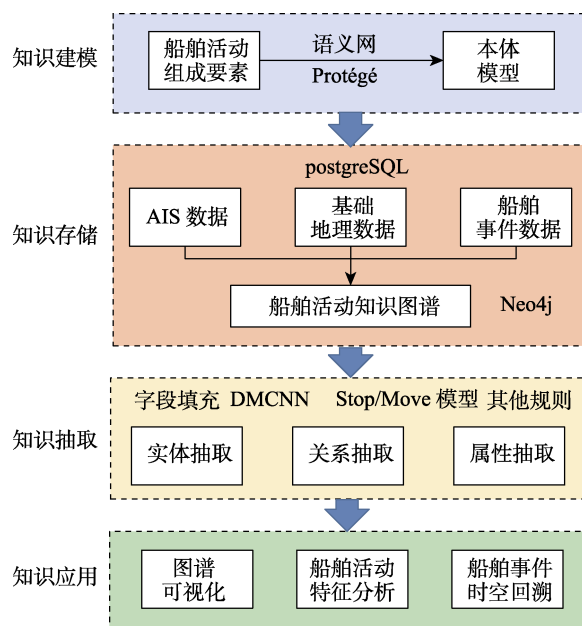


图 2 船舶活动知识图谱构建流程

抽取模型是 DMCNN(dynamic multiPooling convolutional neural networks); Neo4j 是一种图数据库; Stop/Move 模型是一种语义轨迹模型, 用于提取轨迹数据的语义信息。

本文设计船舶活动本体模型(ship activity model based on simple event model, SA-SEM), 将其作为船舶活动知识图谱的本体模型, 如图 3 所示。其中, 绿色椭圆形概念属于核心类, 浅蓝色椭圆形概念属于类型系统, 黄色椭圆形概念属于属性约束。SA-SEM 包括 6 个核心概念: (1) 过程

(Process)。船舶从港口出发直到抵达下一个港口的一次海上航行过程。(2) 事件(Event)。海上船舶状态发生变化的原因, 包括自然灾害事件、海上意外事故和其他事故。(3) 参与者(Actor)。参与事件的主体, 即船舶。(4) 地点(Place)。具有空间位置的实体, 如一个具体的地点名字或者经纬度。(5) 时间(Time)。具有时间特征的实体, 如一个时间点或者一个时间段。(6) 行为(Action)。船舶活动更基础的动作行为, 如起锚, 移动(匀速、加速和减速)和锚定。

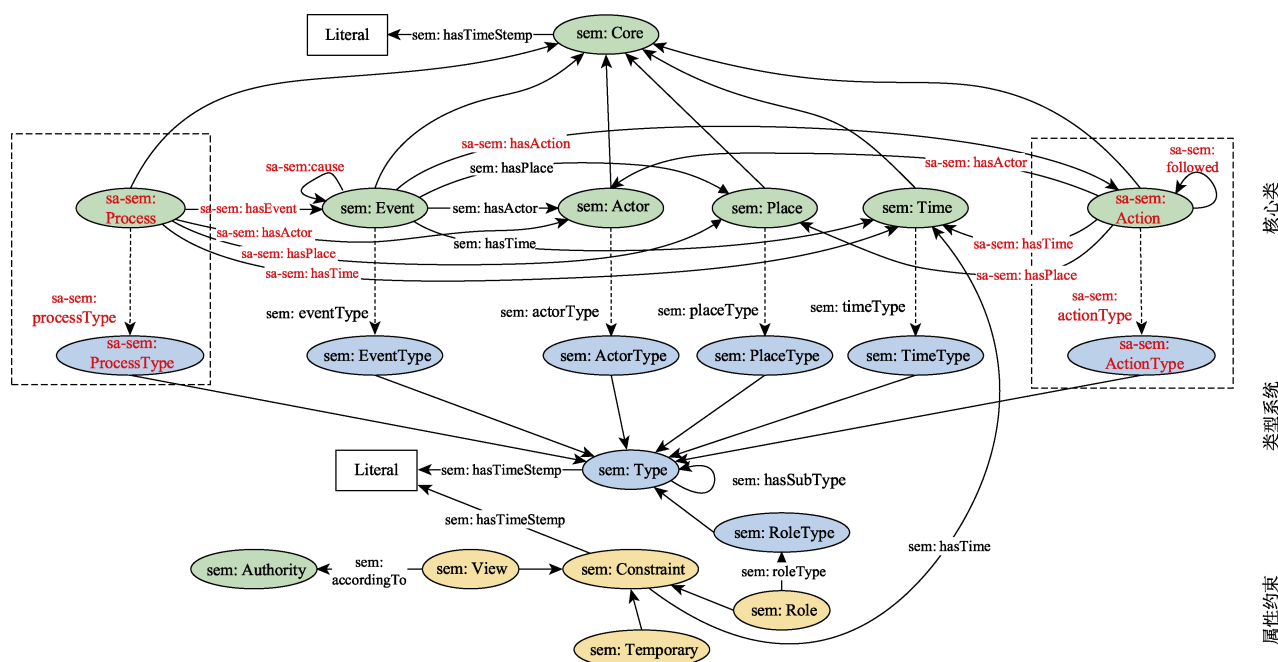


图 3 SA-SEM 关系

除了常规的时序关系和空间关系外, 本文描述船舶活动的实体关联关系还有 3 种: (1) Process, Event, Action 之间的包含关系; (2) Event 之间的因果关系; (3) Action 之间的顺承关系。

2.3 船舶活动分析需求概述

随着全球经济贸易往来越来越频繁, 人们迫切需要掌握海上船舶的活动情况, 如船舶活动轨迹的时空查询和船舶活动事件的时空分析等。本文基于多源船舶活动数据, 发挥知识图谱的查询优势和数据可视化的直观优势, 从 4 个方面梳理船舶活动的分析需求。

分析需求 1. 船舶对象特征分析。由于海上船舶数量庞大, 需要对其进行过滤后再分析。基于特定时间、特定空间的船舶查询结果, 需要对其数量特征、属性特征再进行可视化, 以分析重要局部范围内的海上船舶静态特征。如分析某一年在马六

甲海峡、苏伊士运河等重要区域的船舶数量特征变化和船舶类型分布规律等。因此, 船舶对象特征分析包括船舶的查询、船舶数量特征分析和船舶属性特征分析。

分析需求 2. 船舶活动事件分析。针对关注的特殊船舶, 对其某一时空范围内的活动事件进行分析, 有助于掌握船舶活动事件的态势规律和探究事故发生的原因, 对减少海上事故的发生具有重要意义。如分析某一区域各种类型的船舶发生哪种事故频率最高, 需要围绕船舶的活动事件展开分析。

分析需求 3. 船舶活动关联分析。船舶类型与船舶活动事件之间存在关联关系, 不同类型的船舶在其活动类型的地点和时间都会具有一定的规律。如正常情况下, 渔船会在捕捞期前往允许捕鱼的海域活动, 对于休渔期出现在休渔海域的渔船

活动则是异常行为, 因此可以对船舶类型和船舶活动进行关联分析寻找异常活动。

分析需求 4. 船舶活动时空分析. 分析需求 1 是围绕船舶本身进行分析, 分析需求 2 是围绕船舶事件进行分析, 但这 2 个需求都是基于特定的时空范围, 所以当时空属性变成需要分析的要素时, 就要进行船舶活动的时空回溯, 如分析船舶活动的时序变化特征和空间分布规律。

3 船舶活动知识图谱可视分析模型

3.1 船舶活动分析任务设计

将 SA-SEM 中的 6 个核心概念作为分析要素, 因为船舶过程、船舶事件和船舶行为都属于船舶活动的范围. 为了表示方便, 将 Process, Event 和 Action 统一用 Movement 表示, 即(Process, Event, Action) Movement, 所以分析要素共有 4 个: Time, Place, Actor 和 Movement, 本文将其分别用 T , P , A , M 表示. 针对上述 4 个要素, 由其中一些要素推断出其他要素, 可以转换为简单的组合问题, 即 4 个要素分成 2 部分, 且保证任一部分不能为空, $C_4^1 + C_4^2 + C_4^3 = 14$, 共 14 种分析情况. 本文基于第 2.3 节中的 4 个船舶活动分析需求, 设计 4 个船舶活动分析任务, 将上述 14 种分析情况全部包含进去。

3.1.1 船舶对象特征分析

对于特定的船舶活动, 分析其相关船舶静态特征. 首先查询关注的 Actor; 之后利用 AIS 数据中的静态属性信息对其分类、统计等处理后, 利用数据可视化的方法进行分析. (1) 给定时间约束 $M \times T \rightarrow A \times P$. 即在这一段时间内进行了这些活动的是哪些船舶? 在何地? (2) 给定空间约束 $M \times P \rightarrow A \times T$. 即在这些地方范围内进行了这些活动的是哪些船舶? 在何时? (3) 给定时间和空间约束 $M \times T \times P \rightarrow A$. 即在这段时空范围内, 进行了这些活动的是哪些船舶? (4) 既没有给时间约束, 也没有给空间约束 $M \rightarrow A \times T \times P$. 即何时、何地哪些船舶进行了此类活动?

3.1.2 船舶活动事件分析

分析船舶活动事件. 首先查询关注的事件, 然后围绕事件进行可视分析. (1) 给定时间约束 $A \times T \rightarrow M \times P$. 即某一船舶在这段时间范围内, 于何地进行了什么活动? (2) 给定空间约束 $A \times P \rightarrow M \times T$. 即某一船舶在这些地方范围内, 于何时进行了什么活动? (3) 给定时间和空间约束 $A \times T \times P \rightarrow M$. 即某一船舶在这段时空范围内, 进

行了什么活动? (4) 既没有给时间约束, 也没有给空间约束 $A \rightarrow M \times T \times P$. 即某一船舶于何时、何地进行了什么活动?

3.1.3 船舶活动关联分析

在时空范围已知的前提下, 将不同船舶实体关联的船舶活动进行可视化, 适用于一些重要地缘位置的船舶活动规律分析. (1) 给定时间约束 $T \rightarrow A \times M \times P$. 即在这段时间范围内, 有哪些船舶于何地进行了哪些活动? (2) 给定空间约束 $P \rightarrow A \times M \times T$. 即在这一空间范围内, 有哪些船舶于何时进行了哪些活动? (3) 给定时间和空间约束 $T \times P \rightarrow A \times M$. 即在这段时空范围内, 有哪些船舶进行了哪些活动?

3.1.4 船舶活动时空分析

在船舶和活动已知的前提下, 进行时空回溯是比较常见的分析需求, 往往应用于船舶事故的原因分析. 当某船舶遭遇了某事故后, 为探究原因, 吸取教训, 需要对其时空轨迹进行回溯. 但是, 简单的轨迹回溯不能直接得到事故发生的原因, 需要进一步地挖掘分析. (1) 给定时间约束 $A \times M \times T \rightarrow P$. 即在这一段时间范围内, 这些船舶于何地进行了此类活动? (2) 给定空间约束 $A \times M \times P \rightarrow T$. 即在这一空间范围内, 这些船舶于何时进行了此类活动? (3) 既没有时间约束, 也没有空间约束 $A \times M \rightarrow T \times P$. 即这些船舶在何时、何地进行了此类活动? 上述分析场景推荐的方法如图 4 所示。

3.2 船舶活动可视分析模型设计

可视分析通过人机交互将自动处理和可视分析方法紧密结合^[24]. 在本文进行船舶活动可视分析的流程中, 用户既可以选择自动分析, 也可以选择交互分析. 其中, 自动分析依靠可视化方法的推荐模型完成; 而交互分析能充分地发挥用户在分析流程中的核心作用, 并辅助用户进行决策. 对船舶活动进行可视分析的流程如图 5 所示。

3.3 交互方法设计

在可视分析过程中, 除了视觉方面的可视化图表展示, 用户的交互是核心的要素, 尤其是对于数据量大、结构复杂的系统, 静态可视化图表限制了数据的表达空间. 本文的船舶活动可视分析都是基于船舶活动知识图谱进行的, 在设计交互方式时, 应充分地利用知识图谱的优势. 针对 4 个方面的分析任务, 考虑船舶活动知识图谱数据量大、关联性强和时空相关性等特点, 本文设计了过滤、关联和导航 3 种交互方式。

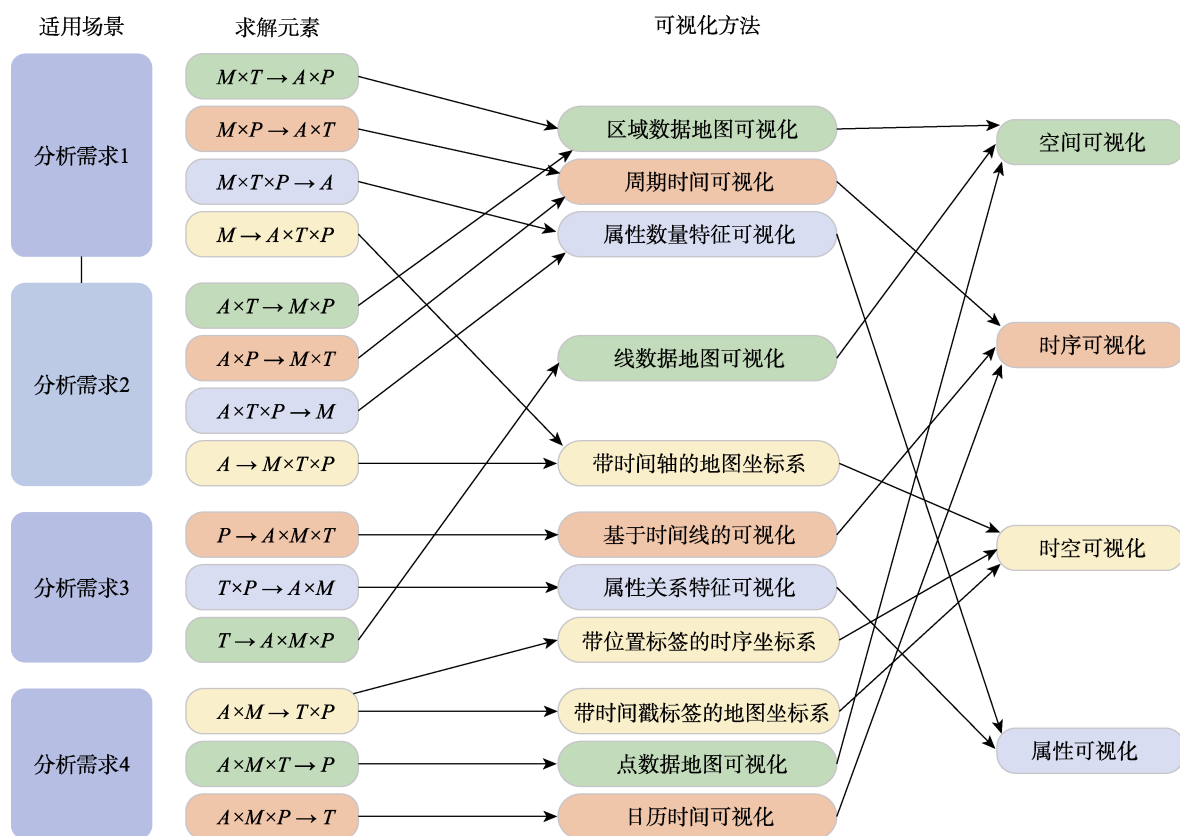


图 4 船舶活动可视化方法推荐

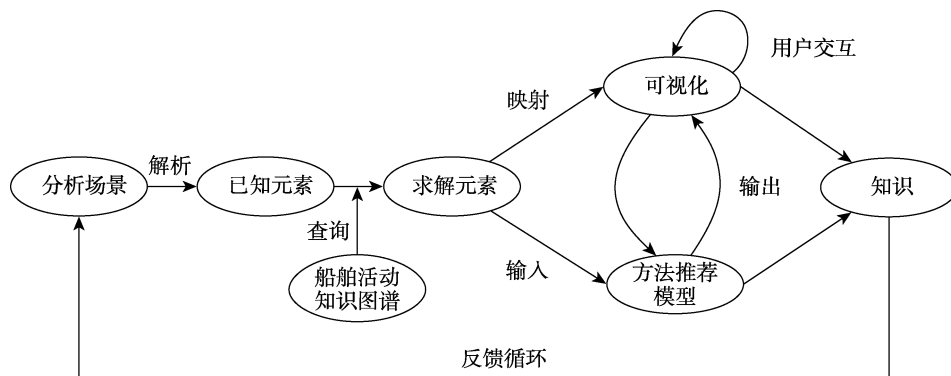


图 5 船舶活动可视分析模型

3.3.1 过滤

通过用户输入关键字实现对图谱的查询, 从本质上过滤了不相关的其他信息, 也可以认为是动态查询的过程中, 使用控件设置其他的筛选条件. 因为船舶活动知识图谱体量较大, 不便展示全部图谱, 所以使用这种方式能帮助用户快速找到感兴趣的船舶活动实体相关的图谱关系.

3.3.2 关联

关联的应用方式有 2 种: (1) 对相同的数据使用不同的视图进行可视化表达; (2) 对不同但关联的数据使用相同的视图进行可视化表达. 用户将多个视图间的数据进行联动, 在图谱中框选一部

分实体, 使得其他视图数据变化为所框选实体关联的数据, 实现视图联动, 便于用户观察数据的不同维度, 进行关联分析等.

3.3.3 导航

由于屏幕空间是有限的, 不能面面俱到地展示图形的所有细节, 因此用户对图形进行缩放和平移等交互操作, “导航”至感兴趣的视点, 实现总体和局部等不同的粒度的分析.

4 船舶活动知识图谱可视分析原型系统

本文开发的船舶活动可视分析系统中, 后端

使用 PostgreSQL 数据库存储 AIS 数据、基础地理数据和船舶事件数据,使用 Neo4j 数据库存储知识图谱数据,采用 Node+React 的全栈框架进行数据处理与传输;前端使用 Echarts, D3 等可视化工具进行数据可视分析,地图是基于百度地图实现的.图 6 所示为船舶活动知识图谱查询界面,图 7 所示为船舶活动知识图谱可视分析界面.



图 6 船舶活动知识图谱查询界面



图 7 船舶活动可视分析界面

船舶活动知识图谱的查询需求大致包括船舶常规活动的查询、突发事件的查询和船舶活动的时空查询.船舶常规活动的查询通过船舶名称、活动时间、活动地点等条件进行查询,在后端使用 Cypher 语句进行查询.如查询船舶名称是 CHANGCHENG9HAO(长城 9 号)的一艘船,近一个月(2021-08—2021-09)在南海海域的活动图谱情况,涉及 Actor, Time 和 Place 这 3 个条件的复合查询,查询结果如图 8 所示.其中,绿色节点表示 Actor,蓝色节点表示 Process,紫色节点表示 Event,粉红色节点表示 Action,红色节点表示 Time,黄色节点表示 Place.

4.1 船舶活动空间分布规律分析

分析船舶活动的空间分布规律时,可以依靠活动的点线面等空间信息进行全方面分析,即船舶活动轨迹点的分析、航线的分析和航行区域的分析.由于 2020 年受到 COVID-19 疫情影响,船舶活动频率大幅降低,对其活动空间分布规律分析的

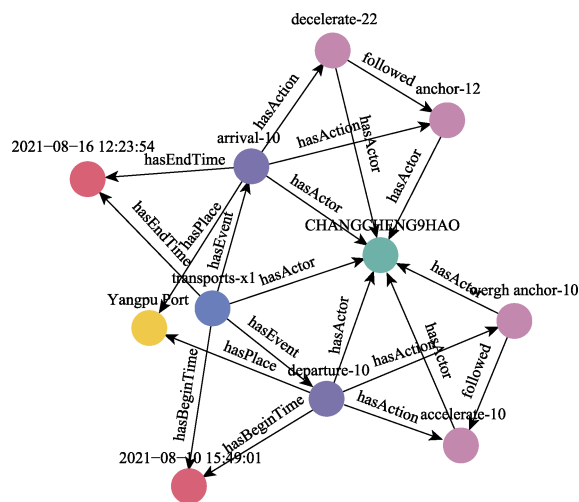
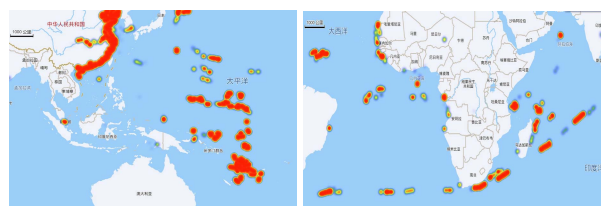


图 8 常规船舶活动查询结果

结果容易失之偏颇,因此本文使用 2016—2017 年的 AIS 数据进行分析.

4.1.1 基于轨迹点的空间分布规律分析

可视分析流程如下:在船舶活动知识图谱中,2017-11-21T00:00:00—2017-11-21T23:59:59 时间段的、船舶过程类型等于“Fishing”且属性“所属国家”等于“China”的船舶(Actor)实体,并将返回实体在 PostgreSQL 数据库中对应的轨迹点利用热力地图进行可视化.图 9 所示为用热力地图表示 2017-11-21 进行捕鱼活动的中国船舶的空间分布情况,用户利用缩放和平移等操作,实现与热力地图的导航交互,从而可以从总体、区域等不同的空间粒度进行分析.可以看出,其主要分布在中国沿海、太平洋地区、非洲东南方向的印度洋地区和非洲西部方向的部分大西洋地区.



a. 中国沿海、太平洋地区

b. 印度洋、大西洋地区

图 9 热力图可视化船舶捕鱼活动

4.1.2 基于航程线的空间分布规律分析

可视分析流程如下:在船舶活动知识图谱中,查询与船舶事件“停靠高雄港”相关的、时间段 2017-07-01T00:00:00—2017-07-31T23:59:59 的船舶(Actor)实体和过程(Process)实体,并根据过程实体的属性“出发地点”和船舶实体的数量利用 FlowMap 进行可视化.图 10 所示为 2017 年各国前

往中国台湾高雄港的船舶活动流量情况,其中,线的宽度表示船舶数量的大小。可以看出,前往高雄港的船舶除了我国台湾外,主要是来自欧洲、非洲和东南亚部分地区的,其中葡萄牙马德拉的船舶最多。

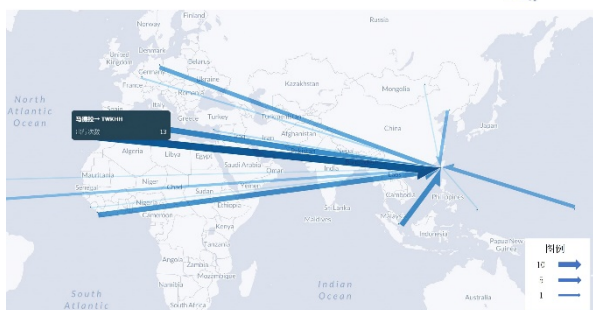


图 10 前往高雄港的船舶活动流量分布情况

4.1.3 基于区域填充的空间分布规律分析

可视分析流程如下:在船舶活动知识图谱中,查询 2017-11-21T00:00:00—2017-11-21T23:59:59 时间段的、与船舶类型是“Military”的船舶(Actor)实体相连接的过程(Process)实体,并将返回实体的“出发地点”所在行政区域利用区域地图进行可视化。图 11 所示为使用区域地图对 2017-11-21 各国船舶进行“军事”类活动的次数进行可视化的结果。可以看出,美国、巴西、澳大利亚、新西兰、英国和尼日利亚的军舰在相关区域活动最频繁。

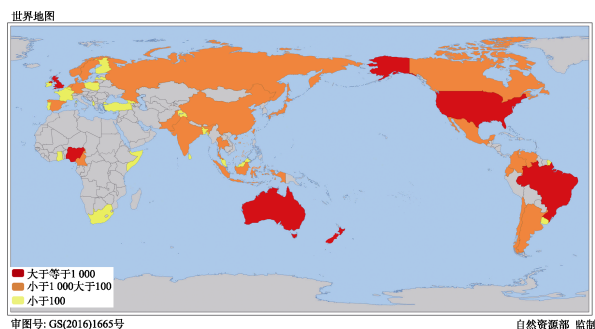


图 11 各国军事类活动频率的空间分布

4.2 船舶活动时序变化规律分析

对船舶活动的时序变化规律分析时,可以依靠船舶活动时序变化的周期性和趋势性进行综合分析。

4.2.1 船舶活动周期性规律分析

可视分析流程如下:在船舶活动知识图谱中,查询与船舶类型是“Fishing”的船舶(Actor)实体相连接的、“出发地点”或“到达地点”这 2 个属性隶属于中国的过程(Process)实体,并将返回实体的数量利用雷达图进行可视化,其中,将渔船在中国沿海

地区进行的活动按照 1 年为周期的频率用雷达图进行可视化,每一个径向的轴表示一个月。周期性规律的单位可以是年、月、周、日、小时等不同的时间粒度。图 12 所示为将出发地和目的地在中国沿海地区且活动类型是客运的船舶,按照 1 年为周期用雷达图进行可视化。可以看出,7 月份左右的渔船活动较少,这与我国沿海地区 7 月份禁渔期相符合,而客船在 2017 年的下半年活动数量出现了倍增的情况。

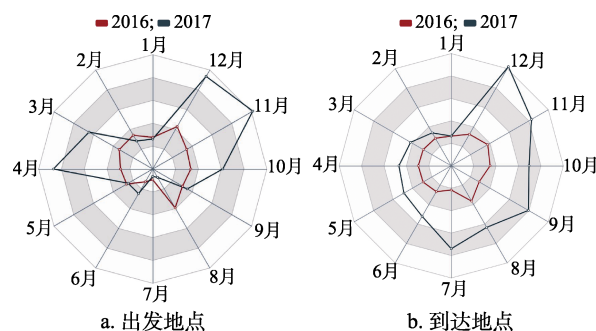


图 12 2016—2017 年中国沿海船舶活动频率变化情况

4.2.2 船舶活动趋势性规律分析

可视分析流程如下:在船舶活动知识图谱中,查询“出发地点”或“到达地点”这 2 个属性隶属于中国的过程(Process)实体,以及返回的过程实体相关联的船舶(Actor)实体和“出发时间”等属性,并将返回实体的数量随时间的变化情况利用河流图进行可视化。在船舶活动知识图谱中,查询时间段 2017-07-01T00:00:00—2017-07-31T23:59:59 的过程(Process)实体,并将返回实体的类型利用河流图进行可视化。图 13 所示为 2017 年 7 月份各类船舶活动次数的变化趋势,总的来说,活动频率趋于稳定。

4.3 船舶活动时空变化规律分析

4.3.1 带位置标签的时序坐标系可视化

可视分析流程如下:在船舶活动知识图谱中,查询与船舶事件“驶入台湾海峡”或“驶离台湾海峡”相关的过程(Process)实体,并将返回的过程实体的“出发时间”“出发地点”或“到达地点”等属性,以及返回实体的数量随时间的变化情况利用气泡图进行可视化。图 14 所示为赤湾港、高雄港、香港港和上海港 2017 年每个月的流量情况,其中,气泡大小表示船舶的数量多少。

4.3.2 带时间戳标签的地图坐标系可视化

可视分析流程与第 4.3.1 节相同。图 15 所示为一段船舶活动的轨迹,将时间属性用时间戳表示并高度还原展示轨迹的位置和运动过程。

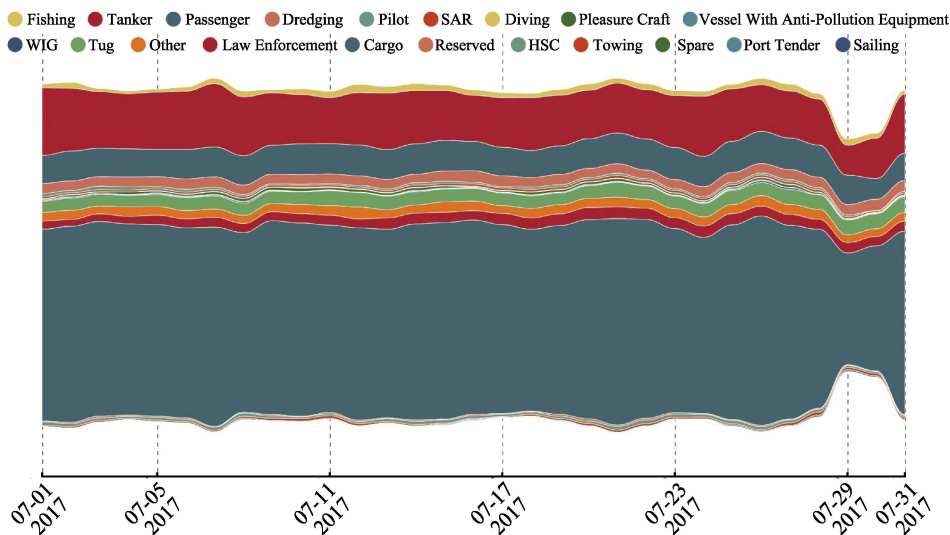


图 13 各类船舶活动次数的趋势变化

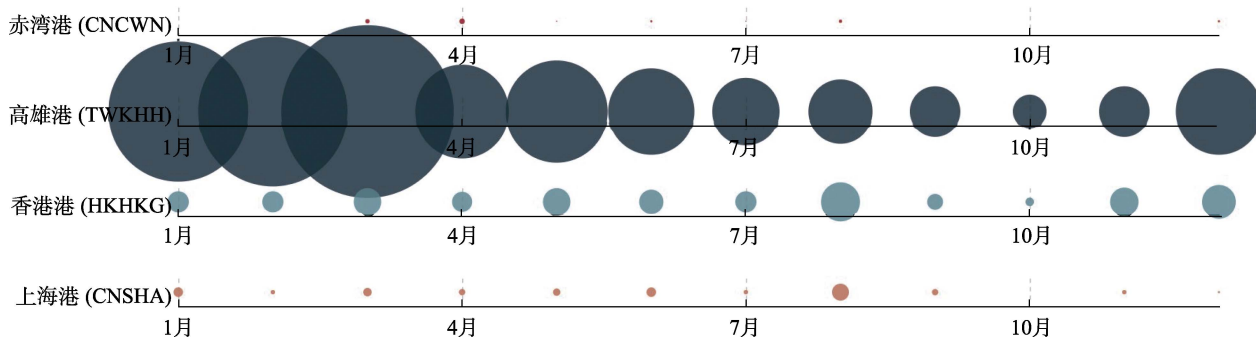


图 14 2017 年我国部分港口船舶流量变化情况

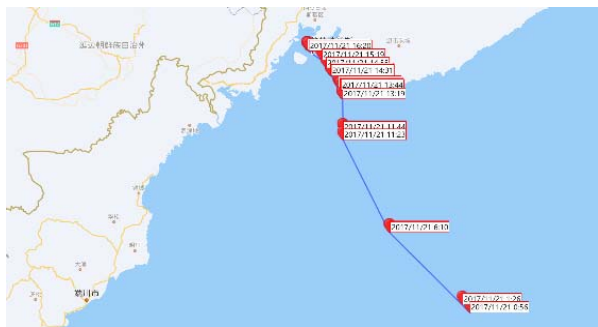


图 15 带时间戳标签的地图可视化

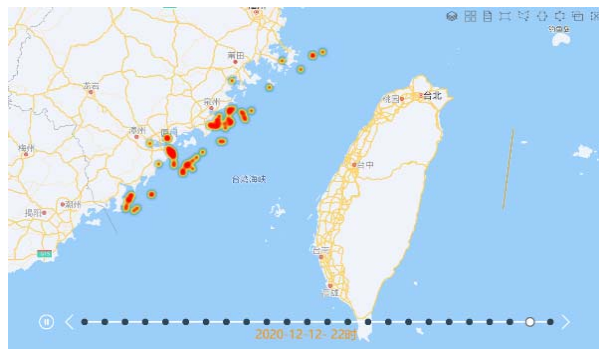


图 16 2017 年 12 月 12 日泉州港附近渔船活动情况

4.3.3 带时间轴的地图坐标系动态可视化

可视分析流程如下：在船舶活动知识图谱中，查询与船舶过程类型等于“Fishing”的过程(Process)实体相关联的船舶(Actor)实体，并将返回船舶实体在 PostgreSQL 数据库中对应的轨迹点进行可视化。图 16 所示为 2017-12-12 泉州港附近的渔船活动情况，使用时间轴控制每小时的数据并以顺序动画的形式展示。

4.4 船舶活动要素属性特征分析

船舶活动要素属性特征包括数量特征和关联

关系特征，其中，关联关系特征指船舶活动对象之间与船舶活动对象的属性之间存在的关联关系。

(1) 船舶活动属性数量特征分析

可视分析流程如下：在船舶活动知识图谱中，查询 2017-11-21T00:00:00—2017-11-21T23:59:59 时间段的、与属性“所属国家”等于“China”的船舶(Actor)实体相连接的过程(Process)实体，并将返回过程实体的类型利用饼图进行可视化。图 17 所示为将中国的船舶类型占比情况用饼图进行展示。

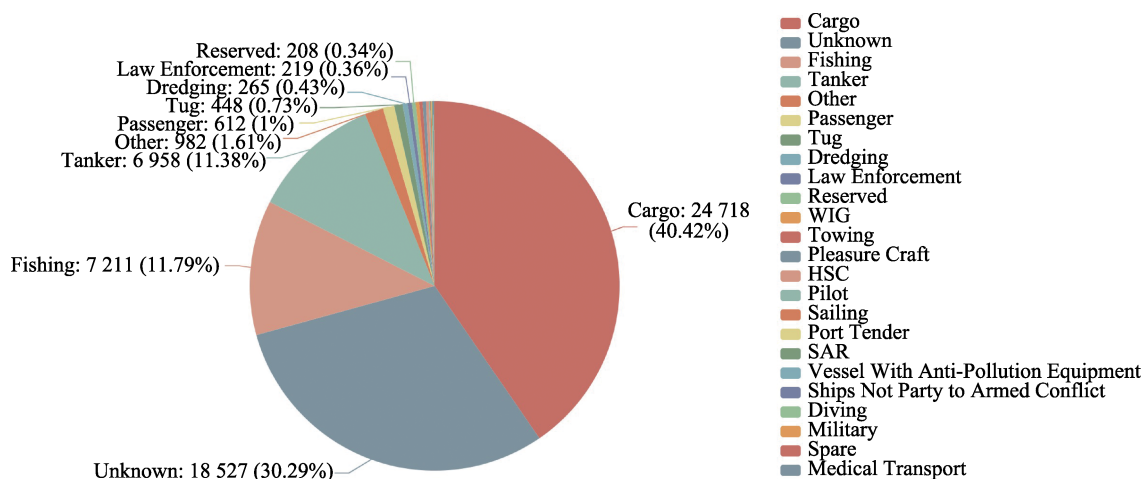


图 17 中国船舶类型占比情况

(2) 船舶活动对象关联特征分析

可视分析流程如下: 在船舶活动知识图谱中, 查询 2016-01-01T00:00:00—2017-12-31T23:59:59 时间段的, 与“达到地点”是高雄港、赤湾港、上海港和香港港 4 个港口的过程(Process)实体相连接的船舶(Actor)实体, 并将返回的船舶实体的“所属国

家”属性利用桑基图进行可视化。图 18 所示为将与“到达高雄港”“到达赤湾港”“到达上海港”和“到达香港港”的事件(Event)实体相关联的过程(Process)实体和船舶(Actor)实体进行关系可视化, 由于实体太多, 因此船舶实体按照“所属国家”进行聚类, 事件实体和过程实体按照“名称”进行聚类。

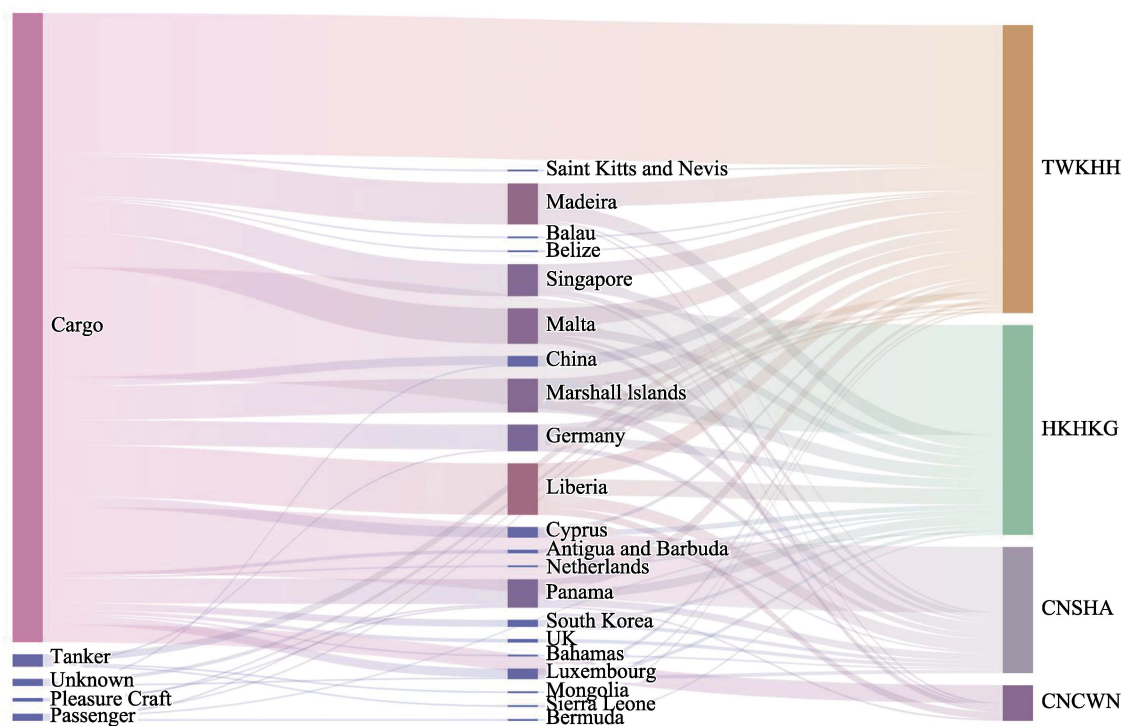


图 18 船舶活动对象关联关系情况

5 结 语

构建知识图谱是一种融合多源数据信息的有效方式, 本文构建的船舶活动知识图谱将多源数

据提取的实体和关系融合在一起, 可以更加高效地分析船舶活动, 将助力海上船舶活动的溯源和分析。本文通过梳理船舶活动分析需求, 根据船舶活动的 4 个核心组成要素“活动、船舶、时间和地

点”设计船舶活动分析模型;使用船舶活动空间可视化、时序可视化、时空可视化和属性可视化等方法,应用于各种船舶活动分析的场景,为船舶活动的管理和监控提供更多的分析视角。

本文的创新点是设计了包含 14 种不同场景的船舶活动可视分析模型,为船舶活动的可视分析任务提供有效的图表和交互支持,并结合过滤、关联和导航等人机交互方式,充分发挥人在可视分析流程中的核心作用,解决分析过程中的“冷启动”问题,同时为船舶活动的分析提供更多的视角。

本文构建了船舶活动知识图谱,但只做了简单的查询和可视化等应用。未来将对船舶活动知识图谱进行推理应用,以及考虑针对不同的船舶活动可视化需求实际情况,对可视化图表进行创新设计,实现更好的可视分析效果。

参考文献(References):

- [1] Zhang Y, Yeung D Y. Overlapping community detection via bounded nonnegative matrix tri-factorization[C] //Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2012: 606-614
- [2] Harati-Mokhtari A, Wall A, Brooks P, *et al.* Automatic identification system (AIS): data reliability and human error implications[J]. The Journal of Navigation, 2007, 60(3): 373-389
- [3] Singhal A. Introducing the knowledge graph: things, not strings[R]. America: Official Blog of Google, 2012
- [4] Scheepens R, Willems N, van de Wetering H, *et al.* Interactive visualization of multivariate trajectory data with density maps[C] //Proceedings of the IEEE Pacific Visualization Symposium. Los Alamitos: IEEE Computer Society Press, 2011: 147-154
- [5] Andrienko G, Andrienko N, Dykes J, *et al.* Geovisualization of dynamics, movement and change: key issues and developing approaches in visualization research[J]. Information Visualization, 2008, 7(3-4): 173-180
- [6] Silveira P A M, Teixeira A P, Soares C G. Use of AIS data to characterise marine traffic patterns and ship collision risk off the coast of Portugal[J]. The Journal of Navigation, 2013, 66(6): 879-898
- [7] Lundblad P, Eurenio O, Heldring T. Interactive visualization of weather and ship data[C] //Proceedings of the 13th International Conference Information Visualisation. Los Alamitos: IEEE Computer Society Press, 2009: 379-386
- [8] Halpern B S, Walbridge S, Selkoe K A, *et al.* A global map of human impact on marine ecosystems[J]. Science, 2008, 319(5865): 948-952
- [9] Adrienko N, Adrienko G. Spatial generalization and aggregation of massive movement data[J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(2): 205-219
- [10] Zhang L Y, Meng Q, Fwa T F. Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters[J]. Transportation Research Part E: Logistics and Transportation Review, 2019, 129: 287-304
- [11] Lampe O D, Hauser H. Interactive visualization of streaming data with kernel density estimation[C] //Proceedings of the IEEE Pacific Visualization Symposium. Los Alamitos: IEEE Computer Society Press, 2011: 171-178
- [12] Wu L, Xu Y J, Wang Q, *et al.* Mapping global shipping density from AIS data[J]. The Journal of Navigation, 2017, 70(1): 67-81
- [13] Andrienko G, Andrienko N, Bak P, *et al.* A conceptual framework and taxonomy of techniques for analyzing movement[J]. Journal of Visual Languages & Computing, 2011, 22(3): 213-232
- [14] Bach B, Dragicevic P, Archambault D, *et al.* A descriptive framework for temporal data visualizations based on generalized space-time cubes[J]. Computer Graphics Forum, 2017, 36(6): 36-61
- [15] Amini F, Rufiange S, Hossain Z, *et al.* The impact of interactivity on comprehending 2D and 3D visualizations of movement data[J]. IEEE Transactions on Visualization and Computer Graphics, 2015, 21(1): 122-135
- [16] Demšar U, Verrantus K. Space-time density of trajectories: exploring spatio-temporal patterns in movement data[J]. International Journal of Geographical Information Science, 2010, 24(10): 1527-1542
- [17] Tominski C, Schumann H, Andrienko G, *et al.* Stacking-based visualization of trajectory attribute data[J]. IEEE Transactions on Visualization and Computer Graphics, 2012, 18(12): 2565-2574
- [18] Pack M L, Wongsuphasawat K, VanDaniker M, *et al.* ICE--visual analytics for transportation incident datasets[C] //Proceedings of the IEEE International Conference on Information Reuse & Integration. Los Alamitos: IEEE Computer Society Press, 2009: 200-205
- [19] Zhao Y, Luo F, Chen M H, *et al.* Evaluating multi-dimensional visualizations for understanding fuzzy clusters[J]. IEEE Transactions on Visualization and Computer Graphics, 2019, 25(1): 12-21
- [20] Lu C, Gao S. Visual analysis of multi-factor association on inland waterway accident[C] //Proceedings of the International Conference on Engineering and Advanced Technology. Paris: Atlantis Press, 2016: 110-115
- [21] Koylu C, Guo D S. Design and evaluation of line symbolizations for origin-destination flow maps[J]. Information Visualization, 2017, 16(4): 309-331
- [22] Phan D, Xiao L, Yeh R, *et al.* Flow map layout[C] //Proceedings of the IEEE Symposium on Information Visualization. Los Alamitos: IEEE Computer Society Press, 2005: 219-224
- [23] Zhou Z G, Meng L H, Tang C, *et al.* Visual abstraction of large scale geospatial origin-destination movement data[J]. IEEE Transactions on Visualization and Computer Graphics, 2019, 25(1): 43-53
- [24] Tominski C, Abello J, Van Ham F, *et al.* Fisheye tree views and lenses for graph visualization[C] //Proceedings of the 10th International Conference on Information Visualisation. Los Alamitos: IEEE Computer Society Press, 2006: 17-24