

船舶轨迹数据的 Geodatabase 管理方法

陈金海^{1,2}, 陆 锋¹, 彭国均², 柯冉绚²

(1. 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, 北京 100101;

2. 集美大学航海学院 船舶助航技术研究所, 厦门 361021)

摘要: 船舶轨迹的自动观测记录已进入了大数据时代,其呈爆炸式增长的趋势给传统的轨迹数据管理方式带来了巨大挑战。本文针对通用船舶自动识别系统(AIS)岸基网络中船舶轨迹数据上传频率高,数据量大,覆盖范围广的特点,首先,分析了当前常见船舶轨迹数据存储方法存在的缺陷,概括了船舶轨迹数据的特征并对其进行抽象建模,然后,在时空立方体模型的基础上,提出了从抽样时刻、步进时段到每日航次的三层组织框架的建模思想,设计了 Geodatabase 的网格化三级时空立方体模型,实现了海洋运输船舶轨迹观测记录的 Geodatabase 管理方法。通过我国 AIS 岸基网络(温州-汕头)单日观测数据的实例验证,表明该模型存储及时空查询性能良好,且具有轨迹数据存储、查询和空间分析一体化管理的独特优势。

关键词: 船舶轨迹;数据库管理;Geodatabase;时空立方体

DOI: 10.3724/SP.J.1047.2012.00728

1 引言

海洋运输船舶轨迹数据是研究水上运动目标行为特征的重要信息源,可直接应用于海上交通运输工程、航运规划管理和海洋生态环保等领域^[1]。近十年来,随着 GPS 定位和海事监管传感器技术的普及运用,船舶轨迹的获取日趋便捷,国际上不少沿海国家已建立起区域性、全天候、高频度的船舶轨迹数据交换中心,自动接收为静态和动态信息。其中,静态数据是指船舶属性数据,如船舶名称、船台呼号、尺寸、类型,更新频率为 6min。动态信息是指船舶实时位置及航次相关信息,包括船速、航向、载货类型、吃水深度、目的港、预计抵达目的港时间等,更新频率为 2s 至 5min 不等。我国已建有全球最大的通用船舶自动识别系统(AIS)岸基网络,截至 2012 年 7 月全国大约有 13.4 万艘船舶配备有 AIS 设备^[2],后续还将推广到数量更多的小渔船。可以预见,船舶轨迹自动采集数据呈爆炸式增长的趋势将日益凸显。因此,如何有效管理海量船舶运输轨迹数据,以满足数据查询和数据分析需要,是迫切需要解决的问题。

2 船舶轨迹数据管理方式

目前,在海洋运输和海事部门,对于船舶轨迹数据的管理主要采用如下方式:

(1) 文件方式

在早期,海事监管传感器单机作业,覆盖范围较小,采集的轨迹记录相应地也以文件形式存储在本地(包括 Excel^[3]、XML、JSON^[4]等格式)。这种管理方式通过压缩编码存储船舶航行过程中的位置信息,在数据备份和网络传输等场合具有优势,缺点是仅适合小容量系统,难以保证数据一致性和安全性,不支持多用户的并发操作,当数据量存储很大时,由于每条船的上传频率不同,数据存储离散,导致查询效率低下。

(2) 移动函数法

该方法通过构造船舶运动方程来表示船舶的航行轨迹。该方法的优点是对于具有固定航行路线的船舶,仅需存储少量的船舶运动参数就能满足船舶历史轨迹查询的需要,并可应用到船舶位置的短时预测,缺点是不适用于大范围游弋的机动性强的船舶,难以构建理想的移动函数模型。

(3) 传统关系型数据库管理系统(RDBMS)

收稿日期: 2012-11-01; **修回日期:** 2012-12-01.

基金项目: 福建省高校产学研合作重大项目“基于 3G 和 AIS 的船舶引航系统”(2012H6015)。

作者简介: 陈金海(1980-),男,福建漳州人,博士研究生,实验师,从事交通地理学、海上交通信息工程研究。

E-mail: chenjh@lreis.ac.cn

近年来,Access^[5-6]、SQLServer^[7-8]、PostgreSQL^[9]和 Oracle^[10]等 RDBMS 都被用于存储船舶轨迹,在一定程度上保证了数据的一致性和安全性。然而,这些 RDBMS 不支持空间索引,空间查询效率很低,更缺乏基本的空间分析能力,如计算轨迹之间的近邻距离^[11]。

(4) 扩展空间数据模型的 RDBMS

为了支持空间分析,很多 RDBMS 引入了空间扩展功能,包括 Oracle Spatial、DB2 Spatial Extender、MySQL Spatial 和 PostGIS^[12],其优点是对象级的数据存储机制和支持扩展 SQL 的查询,其中开源软件 PostGIS 已被用于管理船舶轨迹数据^[13-14]。这些在原有数据库模型上进行了空间数据模型的扩展,仅是简单的数据空间化,即利用点、线等简单要素来存储、检索船舶轨迹数据,却无法管理平面拓扑、线性拓扑,因此,通过扩展空间功能的关系型数据库,还仅是空间数据的容器,空间数据访问和管理效率与真正的空间数据服务器相比,仍然有一定的差距。

以上 4 种常用的轨迹数据管理方式,在进行大规模船舶轨迹数据的复杂时空分析时存在一定的困难。以我国海事岸基 AIS 台站网络为例,它在东海南部沿海可观测到大约 1 万艘在线船舶(5min 内有更新消息),船舶上传周期为 2s~3min 不等,每天接收到的上传消息高达 4 亿条。由于该观测网络的采集时间粒度小,数量大,因此,第一种方法只能用于存储原始消息报文。我国东南有大量路线不确定的渔船活动,这种多变性不适合采用移动函数法。为了减少冗余和保证数据统一,很多学者采用了第 3、4 种方法来管理轨迹数据^[12-14],后者还可通过空间索引进行空间查询,但这两种方法在空间分析能力上还是比较弱。

理想中的船舶轨迹数据库是一种面向特定移动点对象(船舶)的时空数据库系统,具备描述、存储和处理船舶的位置、航次相关属性等基本功能,并支持对船舶进行现在、过去的空间查询和未来预测等高级应用^[15],同时在船舶轨迹数据的组织管理中,还应考虑到扩展性的需要、海量数据的存储、高效并发访问和安全性。

为达到上述目标,本文引入 ESRI 公司的空间数据管理产品 Geodatabase(下文简称 GDB)。它是构建在空间数据文件、小型 RDBMS,以及“大型 RDBMS+ArcSDE”之上的空间数据访问中间件。

Geodatabase 支持多种 DBMS 结构和多用户访问,伸缩性强,且支持长事务管理和历史数据管理^[12]。因此,用户可以根据船舶轨迹数据规模,灵活采取应用架构。由于采用公共模型框架,GDB 具有较好的扩展性,便于将轨迹外围相关信息进行统一描述,例如,水深点、岸线、船舶定线、地址、地形三维表面、航标、风潮流和能见度等地理数据,进而在后续应用中充分发挥 ArcGIS 的空间分析功能。

3 船舶轨迹数据模型设计

3.1 船舶轨迹数据的独特性

利用 GDB 管理海量空间数据是 GIS 应用中比较成熟的方法,现已由静态空间对象扩展到随时间发生离散变化的空间对象,形成维护某种历史信息的时态数据库。例如,土地利用数据和交通路网数据^[16],其采用的数据模型是将空间对象与时间域(包括有效时间和事务时间两种时态)建立关联,例如,状态关系记录(某个特定时间段成立事实)和事件关系记录(每个特定时刻发生的事件)。

移动点目标是发生连续变化的空间对象,其轨迹是比上述空间对象(如公路网络中的路段)具有更强动态性的线要素,其数据建模方法有两种:位置管理视角和时空数据视角^[17]。位置管理视角关注指定区域内所有移动对象的当前(近期)位置的动态维护和查询,只是一种维护现实场景当前状态的快照数据库,它不保存移动点目标的历史信息^[18];时空数据视角的方法除了描述运动目标的当前状况外,还描述了移动点目标演变的历史,可以回退到任一时刻下的移动点位置。这两种建模方法在车辆轨迹数据管理中都有比较成熟的应用,时空索引的有效性是移动对象轨迹查询及网络分析的重要指标,车辆轨迹数据的索引需与具有拓扑关系的路网数据结合,因此,其轨迹的数据模型相对复杂^[19-20]。

经研究,船舶的运行轨迹具有如下特点:

(1)近乎无空间约束的运动:在海洋中航行的船与受限于公路网络中移动的车辆不同,其移动在空间上几乎是无约束的。

(2)历时长且动态变化平缓:从航次角度看,在远距离载运工具中船舶的历时最长(船速较低)。船舶很少发生骤停或急转,船舶轨迹在时空三维中较少出现阶跃,因此,与车辆相比,其时空索引机制

相对简单。

(3)身份隐私保护较宽松:船舶轨迹数据隐私保护要求较低,船舶档案信息易于获取,特别是商业化的大型船舶,可对各船进行灵活的归类(例如船舶尺寸、总吨位和船舶类型的参数组合),有利于满足多样化的分析需求,但这也对其数据模型的可延伸性提出更高要求。

当前处理船舶实时位置或短时预测的位置管理方法,已出现很多在线查询船舶动态的商业网站上,例如,MarineTraffic,Shipxy,Manyship等,但针对船舶移动历史的建模和查询的相关研究较少。为此,本文以时空数据视角,着眼于解释船舶位置及航次相关信息的变化过程,借鉴移动数据库技术,设计船舶轨迹的时空数据模型。

3.2 船舶轨迹数据的抽象

为了方便讨论,在建模前我们先定义 instant、period 和 periods 等 3 种时间类型。具体而言,时间轴上的一个点记为 instant(瞬间),固定的时间间隔记为 period(时段),不相连的固定时间间隔所构成的集合记为 periods。至于点、线的空间类型,记为 point、line,并以 MultiPoint 表示点要素集合, Polyline 表示相关线要素依次连接起来的组合。

(1)MultiPoint 时态类型

首先,对研究对象—船舶进行抽象,当我们只关心船舶的位置,而不关心船舶的尺寸时,可以将其抽象为在海平面连续移动的点要素(Point), MultiPoint(采集设备接收到的广播点记录集合)的值表示点要素(point)随时间而发生的时态演变,该时态类型是一个连续函数 f

MultiPoint:instant \rightarrow point

(2)Polyline 时态类型

其次,对于航次(voyage)可抽象为折线要素(Polyline),与航次相关的信息包括:航次编号、吃水、目的港口、预订抵达港口时间、起讫时间、船舶识别码、航程长度、货物装载情况等。

(3)时空操作

然后,我们定义该时态类型上的历时长度的、点转折线、两点距离和线要素长度等操作。

历时长度的操作返回 MultiPoint 所定义的时间间隔的集合 periods,即 moving(point) \rightarrow periods,该操作可用于判断消息是否超时(数据缺失);

点转折线操作是指船舶在二维空间范围上的

投影成一条折线(Polyline),即 moving(point) \rightarrow line,该操作用于生成轨迹线;

两点距离、线要素长度的操作是船舶在航行中与时间相关的数值特性,其中,两点距离的函数是 moving(point) \times moving(point) \rightarrow moving(real),折线长度计算函数是 line \rightarrow real。

3.3 网格化管理的时空立体模型

在传统的时空立方体模型(S-T Cube)中,X、Y轴为地面经纬坐标轴,T轴为时间轴,由于所有轨迹数据都属于一个大的立方体。当立方体很大时,例如,在容纳单日温州至汕头的船舶轨迹的时空立方体的范围(640km \times 480km \times 24h),要对如此大规模的数据(4亿条记录,来自1万条船)查询历史轨迹是不现实的,因此,需要进行对传统的时空立方体模型进行改进。

(1)确定采样间隔时段,压缩船舶轨迹数据

海洋运输空间范围大,船舶相对速度较低,可将 10s 的上传周期概化为 1min 的采样间隔时段(sample period),这样单日的轨迹采样在时间轴体现为 1440 个 instant。

(2)按步进时段(step period)切分时空立方体的不同历史版本

Step period 是 sample period 的倍数(如 6min 或 30min)。从时间轴出发将每日形成的一级时空立方体切分为多个高度为 step period 的二级时空立方体历史版本。根据研究需要,还可在二维地理平面上确定出经度方向的距离差 $\angle X$ 、纬度方向的距离差 $\angle Y$,把对应的时空立方体历史版本进一步细分三级时空立方体。三级时空立方体的个数取决于 step period、 $\angle X$ 、 $\angle Y$ 和航路分布等因素。三级时空立方体的属性值将包含在 step period 内访问过该位置的所有船舶和航次的识别号序列。

(3)各步进时段内(step period)的轨迹点组成轨迹线历史版本

在步进时段内将来自同一航次(Voyage ID)的轨迹点组成轨迹线,并标记轨迹历史版本标识(MetaVoyage ID),因此,二级时空立方体可以含有多个轨迹线历史版本;对于某自然日航次的完整轨迹,可将一级时空立方体中来自同一航次(Voyage ID)的轨迹历史版本依时间顺序连接起来。

4 船舶轨迹数据模型的实现

船舶轨迹可按照图 1 所示的空间扩展实体-关

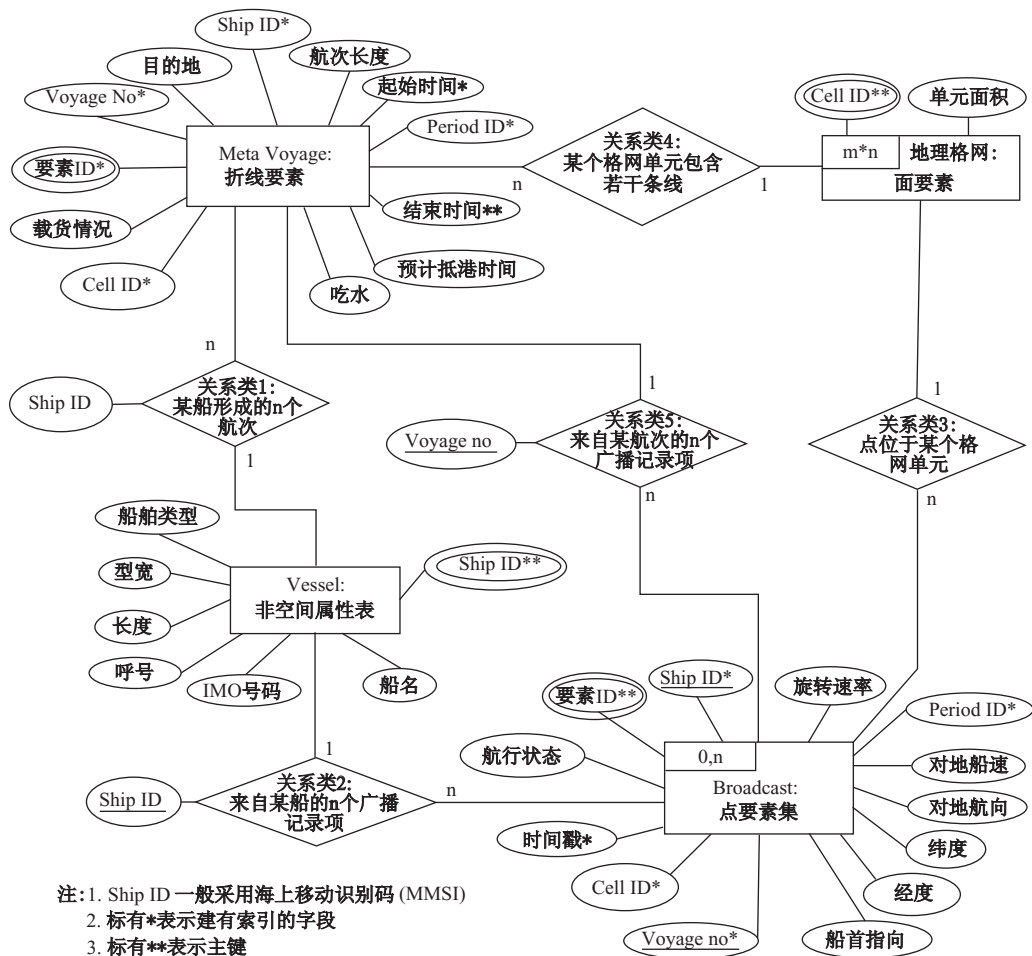


图1 船舶轨迹数据的空间扩展 ER 图

Fig. 1 Spatial extended ER diagram of vessel trajectories

系(ER)进行建模,并采用 GDB 中的点、线、面要素类及非空间属性表来实现。

4.1 广播记录项(Broadcast)

单日所有船接收到的广播记录项(MultiPoint 时态类型)利用带有经纬坐标的点要素类(Multi-Point FeatureClass)来存储其地理描述信息,同时保存各点要素对应的属性信息,包括接收的时间戳、对地船速、对地航向、航行状态、旋转速率、航次编号、船舶的海上移动识别号(MMSI)、Period ID 及其隶属的地理网格编号(Cell ID)等。为提高检索效率,除主键(要素 ID)外,利用 ArcGIS 对航次编号、MMSI、Period ID 和 Cell ID 建立索引。

4.2 船舶静态信息表(Vessel)

MultiPoint 的静态属性信息是非时态数据,可利用非空间属性表(Vessel)来描述,包括 MMSI、船

名、长度、型宽、船舶类型和船台呼号等,其中,MMSI 是主键。

4.3 航次轨迹(Voyage)

对于航次轨迹(Polyline 时态类型)利用折线要素类(Polyline FeatureClass)来存储其地理描述信息,同时与航次相关的信息,包括航次编号、吃水、目的港口、预计抵达港口时间、起始时间、MMSI、Period ID、Cell ID、航程长度、货物装载情况等。除主键(航次编号)外,利用 ArcGIS 对 MMSI、Period ID、Cell ID 和航程长度建立索引。

4.4 规格化的地理网格

利用经度方向的距离差 $\Delta X=1\text{km}$ 、纬度方向的距离差 $\Delta Y=1\text{km}$,将地球平面划分为地理网格,通过一组面要素类(Polygon FeatureClass)来管理,以 Cell ID 为主键并利用 ArcGIS 建立四叉树索引。

4.5 利用关系类保证数据的一致性

为了保证船舶轨迹数据的一致性,我们通过 5 种关系类对上述 3 种要素类,以及船舶静态属性表进行关联。其中,非空间属性表 Vessel 利用主键 MMSI 与线要素类 voyage 和点要素类 Broadcast 的对应字段关联:

关系类 1:某条船行驶过的路线,属于 1:n 的关系;

关系类 2:某条船途经的位置点,属于 1:n 的关系。

此外,线要素类(Voyage)、点要素类(Broadcast)和地理格网面要素类之间的关联有如下 3 种:

关系类 3:n 个广播记录项(Point)位于某个地理格网单元(Polygon)中,属于 n:1 的关系;

关系类 4:n 个线要素(Polyline)位于某个地理格网单元(polygon)中,可以进一步单元时空立方体中的轨迹线集合,属于 n:1 的关系;

关系类 5:n 个广播记录项(Point)构成了某个航次(Polyline),属于 n:1 的关系。

5 原始观测记录到 GDB 的转化过程

船舶原始观测数据往往是以暗码形式(即 VDM,甚高频广播消息)打包的编码电文,将其导入到 GDB 中,需经过以下步骤:

(1) 报文解析:对每天形成的原始 VDM 报文参照相关协议解析后,以设定好的初始步进时长(如 1 分钟)提取位置相关广播消息、船舶静态属性及航次相关信息分别存储为 Broadcast、Vessel 和 Voyage 等 3 个 csv 文件,这 3 种信息的具体解析过程如图 2 所示;

(2) 数据清洁:制定过滤规则对位置报告进行必要的清洁,如 MMSI、IMO 号码的非法识别,指定 ship ID 相邻位置报告的容忍时差或容忍距离,速度或加速度的阈值等;

(3) 数据内插:在 Broadcast.csv 中,针对各船的位置相关消息,根据时间进行排序,按照指定采样时刻(如间隔 6min)进行数据内插,若某船在采样时段内无更新数据时则标记为离线状态;

(4) 创建 GDB:创建空的点要素类、线要素类、Vessel 属性表及其关系类,准备好 GDB。

(5) 时空匹配:遍历各船内插后的采样数据,依次与单元时空立方体进行时空匹配,利用特定的数

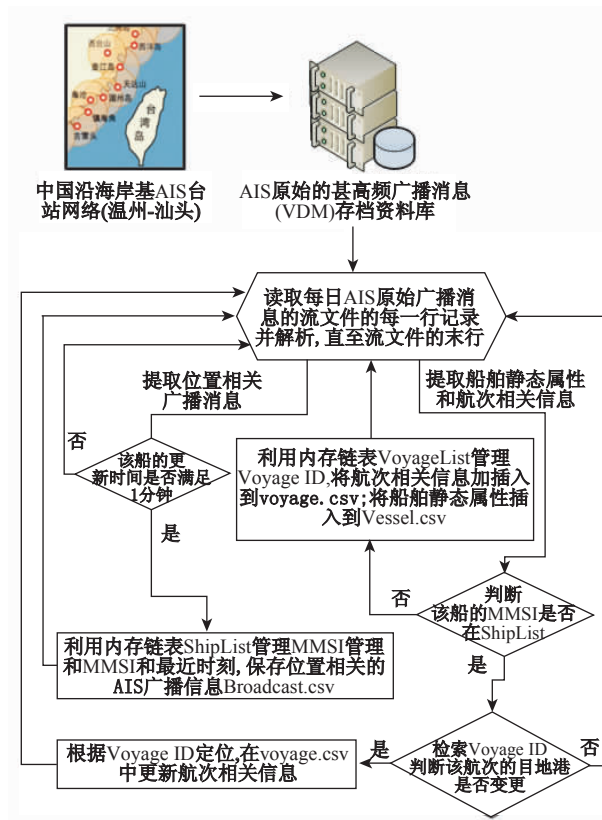


图 2 船舶原始观测数据的解析过程

Fig. 2 Flow chart of original vessel tracking message interpreting

据结构(如关键字为 ship ID,值为 cell ID 和 period ID 的哈希表)管理匹配结果,存储在 GDB 的点要素类 Broadcast 中;

(6) 点转为线:以固定时段(step period)作为时间步进值,将点要素类 Broadcast 中具有相同 Voyage ID 的点进行点转化为折线(单元轨迹版本 MeatVoyage)操作,结果存入 GDB 的 MetaVoyage 折线要素集中,该线要素集可通过 Voyage No 与 Voyage.csv 中航次相关信息相连接;

(7) 保存船舶静态属性:将 Vessel.csv 转换为 GDB 的非空间属性表 Vessel。

6 实例验证

6.1 实验数据准备

以 2011 年 9 月 22 日为例,我国 AIS 岸基网络中温州至汕头沿海 100km 范围内,观测到的船舶近 5000 条,船舶上传位置信息周期为 2s~3min 不等,当日共接收到的广播消息有 4 亿条记录,根据第 5

节所述进行数据预处理和入库操作。

- (1)传统模型:将间隔 1m 的采样数据(即报文解析中的 Broadcast. csv)导入到 Oracle 数据库中;
- (2)GDB 模型:对上述采样数据按 6m 的步进时段将与单元时空立方体(5km×5km×6min)进行时空匹配,并采取“GDB+ArcSDE for Oracle”的管理方式。

6.2 测试环境

后台服务器为 CPU 2. 27G,内存 4G,操作系统 MS Windows Server 2003,数据库是 Oracle 9i。前台客户端为 CPU 2. 27G,内存 4G,操作系统 Windows 7 pro。传统模型的测试平台是客户端软件 Oracle SQL Developer。GDB 模型的测试平台是 ArcGIS 10,基于 ArcObject 与 C#. net 开发的应用程序。

6.3 测试指标

- (1) 存储性能:轨迹数据所占用的存储空间。
- (2) 时态查询性能:根据不同时段大小,设定“时间戳=某时刻”的筛选条件,返回前 50 条记录所耗费的时间。
- (3) 空间查询性能:根据不同的查询几何对象(矩形窗口),设定“被该几何对象包含”的筛选条件,返回前 50 条记录所耗费的时间。

6.4 实验结果分析

- (1) 存储性能比较:LGDB 模型与传统模型的存储性能比较如表 1 所示,记录数和存储空间分别减少了 45%和 69%。
- (2) 时态查询性能比较:GDB 模型与传统模型的时态查询性能比较如表 2 所示。传统模型中的时间戳字段由于没有建立索引,查询的耗费时间与时段大小成正比。GDB 模型中的时间戳建有索引,查询的耗费时间对时段大小不敏感,且效率高,如对于全天候的检索效率提高 5 倍以上。

表 1 存储性能比较		
Tab. 1 Comparison of storage performance		
	传统模型	GDB 模型
记录数	95. 14 万	52. 5 万
存储空间	130MB	40MB

表 2 不同尺度的时态查询比较		
Tab. 2 Comparison of temporal retrieval performance		
时段大小(h)	传统模型(ms)	GDB 模型(ms)
6	547	265
12	948	250
18	1370	265
24	1637	281

表 3 不同尺度的空间查询比较		
Tab. 3 Comparison of spatial retrieval performance		
矩形窗大小(km)	传统模型(s)	GDB 模型(s)
0. 5 * 0. 5	>300	4. 7
1. 0 * 1. 0	>300	4. 3
2. 0 * 2. 0	>300	1. 4
5. 0 * 5. 0	>300	0. 7
10. 0 * 10. 0	>300	0. 5

- (3) 空间查询性能比较:GDB 模型与传统模型的空间查询性能比较如表 3 所示。由于传统模型中对空间查询支持较差,查询时间较难测试,表中记为“大于 300s”。考虑到商船尺寸多数为 200m 以上,表 3 中矩形窗口仅缩至 0. 5km×0. 5km。GDB 模型随着搜索矩形窗口的减小(即空间分辨率的提高),其耗费的时间增加。即便在最坏情况下其空间查询时间耗费小于 5s,优于传统模型。
- (4) 数据的扩展性能:现实中,原始观测 AIS 广播消息提供的信息总是有限的,例如,船舶静态属性表(Vessel)中缺乏载吨重、排水量、最大吃水等信息,而且船型划分相对粗糙(如油轮并无细分为成品油、原油、化学品液货和气槽船等),在更复杂的分析中需要参考其他信息源对 Vessel 表进行扩展操作。GDB 数据模型具有很好的延伸性,用户可以通过域、验证规则和 ArcGIS 软件框架来实现,这些“自定义”的扩展行为可以在不需要编写任何代码的情况下实现,这是采用 GDB 模型的独特优势。
- (5) 空间分析及可视化表达性能:船舶轨迹数据的位置大多是采用 WGS 84 坐标系,由于 ArcGIS 支持动态投影,利用 GDB 管理船舶轨迹数据,更有利于将轨迹数据的查询和分析结果,以人们习惯的不同投影方式展示(例如港区大比例尺下采用高斯投影,海洋小比例尺采用墨卡托投影)。同时,海洋水深面、水深点、海洋岸线、地理标注、锚地、航标等多源数据在 ArcGIS 软件框架中易于集成,例如,图 3 所示叠加了海洋水深等值线图层和 Open-StreetMap 底图,丰富了船舶轨迹相关的背景信息。

传统的轨迹管理方式基本上不考虑其可视化表达。

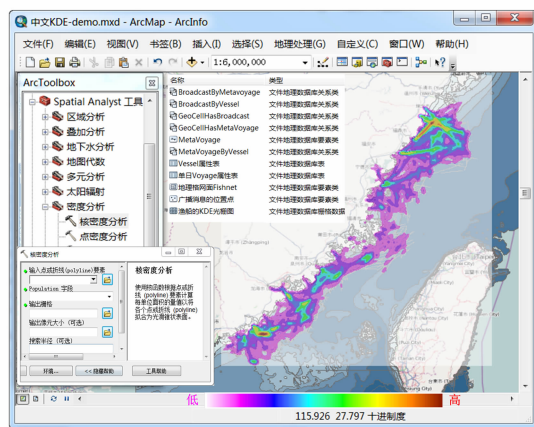


图3 应用 ArcGIS 的地理处理工具展示单日渔船的活动热区

Fig. 3 The kernel density distribution map of fishing vessel activities on western Taiwan Strait

GDB 数据模型将船舶轨迹对象化为要素(点、线和面),是具有地理描述信息、行为和属性的实体,包括哪些位置点要素是某航次线要素的起迄点(点与线的拓扑关系)、哪些位置点要素从属于某船的行为等。这种面向对象的、表达和管理时空数据的模型实现了要素类、栅格数据、网络和地址等地理数据的统一存储仓库和管理模型,从而方便对复杂的空间关系进行建模和分析。典型的空间分析工具会针对某一 GDB 数据集(如要素类、栅格或表)执行操作,并最终生成一个新的数据集,如渔船活动热区分析、海上航路规划等。

7 结论

针对海量船舶轨迹数据管理的现实瓶颈,本文提出用 GDB 管理船舶轨迹数据的模型设计和实现机制,有效保证入库后的轨迹数据一致性、安全性,并支持多用户并发访问,同时与传统模型相比,用 GDB 的船舶轨迹数据管理方式具有以下优势:

(1)应用全局到细部的管理架构,提高了查询效率:在一次时空查询中,首先,定位是哪一天的航次轨迹,然后,再确定到更小时段范围即轨迹历史版本,最后,在单元时空立方体中进行细节查询。这种数据查询范围与传统方式相比缩小了很多,提高了查询效率。从某日、某时段到采样的三层时空立方体管理架构,符合人们探索船舶运动规律时从全局到细部的逻辑习惯,有利于实现船舶、航路和

环境的空间分析过程。

(2)方便融合外围数据,发挥专业 GIS 软件的空间分析优势:由于很多船舶轨迹相关的外围数据都支持 GDB 数据模型,可在 ArcGIS 软件框架中实现多源数据的融合和展示,包括自动导入各种在线底图数据(如谷歌地图、Bing Map 和 OpenStreet-Map 等),并且可直接应用 ArcGIS 软件的各种现成空间分析工具,实现海量船舶轨迹数据的存储、查询、空间分析和可视化表达的一体化管理。

参考文献:

- [1] 陈金海,陆锋,彭国均. 海洋运输船舶轨迹分析研究进展[J]. 中国航海, 2012, 35(3): 53-57.
- [2] 胡晓菁. AIS:构建我国“水上物联网”的信息平台[J]. 中国海事, 2011(7): 21-23.
- [3] 张连东. 基于 AIS 数据的成山头水域船舶交通流研究[D]. 大连:大连海事大学, 2010.
- [4] Calder B R, Schwehr K. Traffic analysis for the calibration of risk assessment methods[C]. Norfolk, VA., 2009.
- [5] 宋扬. 基于 AIS 数据的大型船舶锚泊半径及船间距的研究[D]. 大连:大连海事大学, 2010.
- [6] 李红祥,方逊. 基于 AIS 的船舶交通流量统计方法研究[J]. 武汉理工大学学报(交通科学与工程版), 2011, 35(4): 853-857.
- [7] 林伟珣,邵哲平. 基于嵌入式 Web 服务器的 AIS 数据无线采集系统[J]. 上海海事大学学报, 2011, 32(1): 40-43.
- [8] 李昊. 基于长江干线 AIS 的航运信息平台的开发与研究[D]. 武汉:武汉理工大学, 2011.
- [9] Qinyou. Hu, Jinhai. Chen and Chaojian. Shi. Bringing live AIS information on the web sea charts by using Ajax[C]. // Proceeding of 7th International Conference on ITS Telecommunications, Sophia Antipolis, FRANCE, 2007, 455-459.
- [10] 纪贤标,邵哲平,潘家财,等. AIS 信息分布式采集系统的开发及关键技术[J]. 上海海事大学学报, 2007, 28(1): 28-31.
- [11] 潘家财,姜青山,邵哲平. 船舶会遇的时空数据挖掘算法及应用[J]. 中国航海, 2010, 33(4): 57-60.
- [12] 程昌秀. 空间数据库管理系统概论[M]. 北京:科学出版社, 2012.
- [13] 申慧超,胡勤友,杨春. 基于 AIS 数据的预抵船舶联系信息查询系统[J]. 上海海事大学学报, 2010, 31(4): 13-16.
- [14] Schwehr K D, McGillivray P A. Marine Ship Automatic Identification System (AIS) for Enhanced Coastal Security Capabilities: An Oil Spill Tracking Applica-

- tion[C]. Vancouver, B. C. , 2007.
- [15] 郝忠孝. 移动对象数据库理论基础[M]. 北京:科学出版社, 2012.
- [16] 周亮, 陆锋, 张恒才. 基于动态中介中心性的城市道路网实时分层方法[J]. 地球信息科学学报, 2012, 14(3): 292 - 298.
- [17] Guting R H, Schneider M. Moving Objects Databases [Z]. 金培权, 岳丽华 翻译. 高等教育出版社, 2009.
- [18] 吴建华. 基于 ArcGIS Engine 的车辆监控 GIS 系统开发[J]. 地球信息科学学报, 2011, 13(1): 88 - 94.
- [19] 郭旦怀, 崔伟宏. 面向实时交通信息提取的车辆轨迹数据挖掘[J]. 武汉理工大学学报(交通科学与工程版), 2010, 34(1): 6 - 9.
- [20] 蒋益娟, 李响, 李小杰, 等. 利用车辆轨迹数据提取道路网络的几何特征与精度分析[J]. 地球信息科学学报, 2012, 14(2): 165 - 170.

Study on Vessel Trajectories Database Manage System Based on Geodatabase

CHEN Jinhai^{1,2}, LU Feng¹, PENG Guojun² and KE Ranxuan²

(1. *State Key Laboratory of Resources and Environmental Information System (LERIS), Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;*

2. *Navigation College & Navigation-Aids Technology Research Center, Jimei University, Xiamen 361021, China)*

Abstract: In order to monitor real-time vessel information to improve navigation safety, China's Maritime Safety Administration (MSA) has built the world's biggest Automatic Identification System (AIS) shore-based network, in which data such as ship position, name, purpose, course and speed are automatic collected 24 hours per day primarily in Chinese coastal waters. As a result, China is approaching the era of big data storage of vessel trajectories, which has brought great challenges to traditional moving objects data management systems. Beyond their basic functions of loading and displaying vessels' position records, an ideal vessel trajectories database should bring user more advance functions of analysis of ship tracking records by supporting spatio-temporal query and prediction of vessel movement. In this paper, we start with the character of vessel movement and abstract the data model of vessel trajectories according to state-of-the-art technology of moving objects databases. Due to the characteristics of vessel trajectories data, such as changing frequently, wide cover range and mass datum, it is argued that current methods of trajectories storage still deserve much more research and improvements, especially for spatio-temporal query and geoprocessing support methods. Considering the role of time perspective, vessel trajectories are managed by three kind of time unit (sampling instant, stepping period and 24 hours) so as to built a three-level organizational framework. By compressing the volume of data and matching original vessel tracking message into spatio-temporal cube unit the retrieval efficiency increases significantly. It was also described how to streamline the acquisition, loading, filtering, display and analysis of raw AIS log files. This method is applied in handling daily mass vessel tracking records which are covering western Taiwan Strait. Experiences show that this model satisfied the requirements of application. The storage is reduced and the performance of spatio-temporal query is improved. Using ArcGIS platform of Geodatabase module, vessel trajectories' initial data model is easy to revised, expanded and integrated with various relational base-map data. Furthermore, it is convenient to apply variable ArcGIS geoprocessing tools to obtain customize demand, such as daily hot spot activities of fishing vessel. By generating these synthesized products our solutions would support the ocean planning community to better understand marine transportation patterns and potential use conflicts between vessels and other activities.

Key words: vessel trajectories; database management; Geodatabase; spatio-temporal cube