

面向综合测绘的空间数据引擎实现

刘付海萍

化州市规划编制与测绘研究信息中心, 广东 化州

收稿日期: 2022年12月1日; 录用日期: 2023年1月10日; 发布日期: 2023年1月18日

摘要

不同底层的GIS地理信息数据格式对空间写入设置屏障, 阻碍了数据的共享和互操作, 为解决这个问题对不同数据格式的解析读取将是关键操作。本文通过对空间数据格式ESRI MDB、超图SDB和数维MDB的读取, 分析了这三种格式的内部组织结构, 提出了存储数据的空间结构。该种存储结构便于对解析大量数据拓扑操作、空间转化和可视化等相关后续处理, 同时分析了转储这几种数据格式的一些关键操作。

关键词

SDB, MDB, 空间数据引擎, 数据结构

Implementation of Spatial Data Engine for Integrated Surveying and Mapping

Fuhaiping Liu

Huazhou Planning and Mapping Research Information Center, Huazhou Guangdong

Received: Dec. 1st, 2022; accepted: Jan. 10th, 2023; published: Jan. 18th, 2023

Abstract

Different underlying GIS geographic information data formats set barriers to spatial writing, which hinders data sharing and interoperability. To solve this problem, parsing and reading different data formats will be the key operation. This paper analyzes the internal organization structure of the three spatial data formats ESRI MDB, Hypergraph SDB and Digital MDB, and proposes the spatial structure for storing data. This kind of storage structure is convenient for the subsequent processing related to the topological operation, spatial transformation and visualization of parsing a large amount of data. At the same time, some key operations of dumping these

data formats are analyzed.

Keywords

SDB, MDB, Spatial Data Engine, Data Structure

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 简述

空间数据引擎(简称 SDE)是基于特定的空间数据模型,在特定的数据存储、管理系统的基础上(典型的是数据库管理系统),提供对空间数据的存储、检索等操作,以提供在此基础上二次开发的程序功能集合。

移动 GIS 在带来众多新的行业应用模式的同时,也遭遇到技术瓶颈,一些问题严重地制约着移动 GIS 平台的进一步发展。如在线数据调度响应滞后,浏览及显示不流畅;空间查询不稳定,影响系统正常运行,数据更新不及时等。这主要是受网络及移动设备软硬件条件的限制,很多传统桌面 GIS 技术无法直接移植到移动 GIS 产品中[1]。因此,设计一个适合于智能移动终端的轻量级空间数据引擎,研究完善的数据组织、管理及检索机制,是提高移动 GIS 服务能力的关键。化州市在工程建设项目涉审测绘领域推行综合测绘,土地、规划、房产等多测合一工作在前期研究的基础上,来到深化落实阶段[2]。长期以来各领域测绘业务已经形成自有体系,成果数据格式主要基于不同底层的 GIS 地理信息处理软件,推行综合测绘首先需要破除厂商间的数据访问壁垒,实现数据互联互通等的业务化处理操作[3]。

为解决现阶段土地、规划、房产测绘成果共建共享,市自然资源和规划局推进深化综合测绘管理平台建设,其中测绘数据成果资源整合是平台建设的奠基石。平台建设以共享开放自主创新为原则,目前市场中土地、规划、房产测绘业务主要依托 AutoDesk、ERIS ArcGIS、超图 SuperMap 及浙江数维 Walk 等平台。体系涵盖主流 GIS 厂商,鉴于数据格式都没有开源分布,同时各家也没有推出针对其他数据格式转换的工具或拓展实现,故基于开源方案从底层实现了面向综合测绘管理平台的空間数据引擎。

2. 数据组织与关键技术分析

数据引擎整合综合测绘成果数据存储格式,实现 SHP、ESRIMDB/GDB、超图 SDB、数维 MDB、POSTGIS 等多种空间数据库的读取与写入,支持点、线、面、注记等对象。数据引擎从底层访问数据库体,无需依赖插件及 SDK,实现对象格式解析并输出成果,下面主要对 ESRI MDB、超图 SDB 与数维 MDB 三种类型结构解析及关键技术介绍。

2.1. ESRI MDB 存取技术

Geodatabase 是 ESRI 公司在 ArcGIS 中引入的空间数据模型,是建立在关系型数据库上的空间型数据库,支持除矢量外的拓扑、网络、栅格等多类型数据存储并统一管理。综合测绘成果多以 Personal/File Geodatabase 存储,其中 MDB 在格式普及程度及访问易用便捷上占据优势。MDB 基于关系型数据库,通过 OLEDB 驱动 32/64 位访问,空间图层本质为关系型表,其中属性信息由结构化表本身存储,空间信息主要由 GDB_Items、GDB_GeomColumns、GDB_SpatialRefs 三者组合维护管理,见图 1 所示。



Figure 1. MDB database spatial entity object table organization definition
图 1. MDB 数据库空间实体对象表组织定义

GDB_SpatialRefs 维护空间参考坐标系统，定义坐标参考系统的最小 X/Y 值坐标与 XY 坐标精度，数据引擎对数据库的写入需要结合空间参考才能准确定位并渲染地理实体的空间位置。GDB_Items 维护管理数据库中包含的所有关系型表，通过 ArcGIS 表达为图层、属性表或注记对象。数据引擎装载 mdb 文件后，初始化图层、属性表等类型、结构、字段定义等信息。若数据组织定义为空间或注记层，则通过表名在 GDB_GeomColumns 中查询该图层的几何字段与参考坐标系定义。

属性信息由数据类型、字段长度、数值精度等规范，空间要素的几何对象由二进制流组织，存放在 Shape 空间字段中，文件流头部 4 字节整型描述要素几何类型。空间点由 2*8 字节浮点型表示 XY 值坐标；线型对象头部分别存储 4*8 字节浮点型描述几何范围，2*4 字节整数描述部件数量及节点数量，其中每个部件由多个折线组成，通过部件及节点数遍历获得每段折线上的所有节点。面几何与线型要素结构一致，额外增加了部件及几何环的构造。

章节简单阐述常用几何对象二进制存储结构概要，空间引擎支持所有满足 ESRI 定义的对象类型，包括 Z 值要素、注记要素图层等。数据引擎针对 Personal Geodatabase 操作的关键技术在对空间数据写入的解译，ESRI 对空间参考系统定义了独特的格网索引方式使得单一写入后无法在 ArcMap 中正常显示，如图 2 表示。数据引擎在插入更新操作时重写对应的几何字段与空间参考信息，使写入数据满足 ESRI 原生规则获得对数据库自由读写的能力。

geoColumn										
TableName	FieldName	ShapeType	ExtentLeft	ExtentBottom	ExtentRight	ExtentTop	IdxOriginX	IdxOriginY	IdxGridSize	SRID
BandLine	Shape	3	75069.845	87495.851	75289.635	87680.683	75064.35025	87491.2302	1.07465E-07	2
BandMark	Shape	1	75069.845	87495.851	75289.635	87680.683	75064.35025	87491.2302	1.07465E-07	2
Building1	Shape	4	75018.1845	87417.48255	75247.09	87766.124	75010.1845	87417.48255	1.66309E-07	3
DLTB	Shape	4	75069.845	87495.851	75289.635	87680.683	75064.35025	87491.2302	1.07465E-07	2
Facility	Shape	3	75125.444	87437.635	75345.435	87727.54	75119.94423	87430.38738	1.41747E-07	4
GDB_Items	Shape	4	-180	-90	180	90	-180	-90	3.72529E-07	1
shapeIndex										
IndexedObjectID	MinGX	MinGY	MaxGX	MaxGY						
3	0	0	966367641	483183820						
4	0	0	966367641	483183820						
5	0	0	966367641	483183820						
6	0	0	966367641	483183820						
7	0	0	966367641	483183820						
spatialRef										
SRID	SRTEXT	FalseX	FalseY	XYUnits	FalseZ	ZUnits	FalseM	MUnits	IsHighPrecision	Tolerance
2	{B286C06B-0879-11D2-AAAC-00C04FA33C20}	75064.35025	87491.2302	3.90295E+13	-100000	10000	-100000	10000	2.56217E-14	
3	{B286C06B-0879-11D2-AAAC-00C04FA33C20}	75010.1845	87417.48255	2.522E+13	-100000	10000	-100000	10000	3.9651E-14	0.001
4	{B286C06B-0879-11D2-AAAC-00C04FA33C20}	75119.94423	87430.38738	2.959E+13	-100000	10000	-100000	10000	3.37952E-14	0.001

Figure 2. ESRI Personal Geodatabase index and grid solution steps
图 2. ESRI 个人地理数据库索引及格网解算步骤

2.2. 超图 SDB 格式解析

房产测绘以超图 SDB 库格式输出，为了读取解析并输出为业务标准成果；数据引擎从底层对 SDB 格式进行字节访问。SDB 是超图自有的基于文件和数据库方式混合存储的空间数据组织方式。SDB 数据工程包括两个文件，扩展名.sdb 文件存储空间数据，采用 OLE 复合文档技术；扩展名.sdd 文件为属性数据库，采用微软 Access 的 MDB 数据库。

其中 sdb 复合文档以扇区构造存储为关键解析对象，提供了在一个工程中存储多个数据集能力，所有要素集与图层在 Datasets 中存储，数据引擎初始化后获取到整体数据集结构与对应节点的数据流。数据集以红黑树节点形式分布在不同扇区上，每个数据节点包含 Info、Index 与 Geometry 三块内容，分别描述：基础信息、位置索引与几何对象，因此数据引擎在要素新增与删除操作中需协同更新基础信息与索引保证数据完整可读。基础信息定义了视图范围、图层类型、触发器、记录数、空间索引等信息，如图 3 所示；Index 位置索引与 Geometry 几何对象互相配合解析，其中对图层类型为空间几何与注册的分析为解析关键技术。

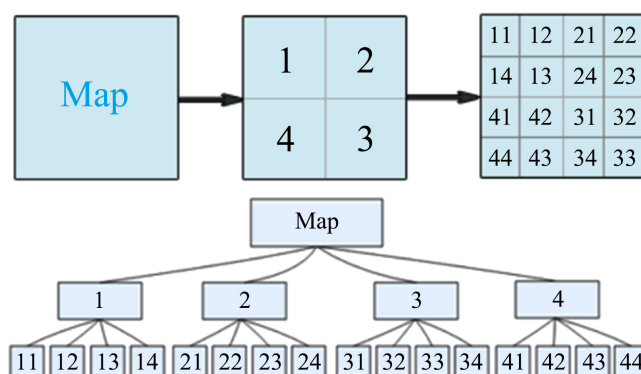


Figure 3. SDB5 uses Grid depth to describe spatial index
图 3. SDB5 采用 Grid 深度描述空间索引

Index 位置索引依次以 4*3 字节长度定义要素的 ID、字节位置、空间索引信息，索引二进制流长度 = 头部长度 + 记录数*12；其中字节位置用于 Geometry 几何对象中定位二进制读取的起始位置。

46349	442b	5fe4	1535	5b40	cf0e	4363	4269	7f40
46350	f7af	45dd	b53b	5b40	506e	75c7	d068	7f40
46351	0100	0000	0600	0000	442b	5fe4	1535	5b40
46352	cf0e	4363	4269	7f40	f7af	45dd	b53b	5b40
46353	506e	75c7	d068	7f40	34e7	4b1e	0d3a	5b40
46354	ece3	f37c	0000	7f40	d8c4	49e3	6438	5b40
46355	f975	86a8	0000	7f40	6ece	2d27	bd36	5b40
46356	c80a	7f4a	2000	7f40	442b	5fe4	1535	5b40
46357	cf0e	4363	4269	7f40	f500	0000	0500	0000
46358	0800	0000	7800	0000	f7af	45dd	b53b	5b40
46359	506e	75c7	d068	7f40	e00f	ae64	5d42	5b40
46360	6986	4298	6668	7f40	0100	0000	0700	0000
46361	f7af	45dd	b53b	5b40	506e	75c7	d068	7f40
46362	97c6	0000	0000	7f40	0000	0000	0000	0000
46363	e00f	ae64	5d42	5b40	0000	0000	0000	0000
46364	f1bf	0000	0000	7f40	0000	0000	0000	0000
46365	1622	0000	0000	7f40	0000	0000	0000	0000
46366	54cf	0000	0000	7f40	0000	0000	0000	0000
46367	f7af	45dd	b53b	5b40	506e	75c7	d068	7f40
46368	f600	0000	0500	0000	0800	0000	6800	0000
46369	e00f	ae64	5d42	5b40	6986	4298	6668	7f40
46370	c6ae	2331	0c49	5b40	1828	3eda	0368	7f40
46371	0100	0000	0600	0000	e00f	ae64	5d42	5b40

Figure 4. Fragment flow of spatial elements in geometry object
图 4. Geometry 对象中关于空间要素的片段流

Geometry 对象中点以 8 位字节浮点型定义 XY 值坐标；线型要素头部含 48 字节数值，分别定义主键编号、几何类型、几何流长度、要素范围等信息，文件头之后定义了线的部件数量及各部件节点数量，总长度 = 节点数量*16 位浮点型数值；面状要素构造与线型要素类似，但是超图对 Polygon 面对象定义通过单一环构成，缺省了环的逻辑顺序从而简化对象结构，缺点是在对象转储中需要计算环的逻辑类型，判断外环或内岛，图 4 表示面对象二进制片段。当数据集类型为注记时，引擎解析对象中的注记范围、文本信息、字体、字体高度、格式、定位旋转等信息，以 Element 元素存储。

2.3. 数维 MDB 存储解析

Walk MDB 是浙江数维定义的空间存储格式，特点是将几何对象与符号化信息以字段形式存储在结构化表中，目前在土地测绘中应用比较广泛，数据读取与写入方式与 ESRI MDB 一致采用 OLEDB 驱动。数据引擎初始化过程从 WalkLayers 中解析数据集，图层符号化与注记分别以数据集名称 + Symbols 与 Annotations 的表结构形式存在。

Symbol 符号化表主要通过 StyleID 与 Memos 属性对外输出样式定义，符号化信息在 Style 字段中内联存储；Annotation 注记层和 Feature 要素层都是空间要素通过符号信息在软件中绘制渲染，关系见图 5 所示，其中注记以 Annotation、Location 表达样式与位置信息，要素以 Geometry 字段表达几何构造。

FeatureID	Geometry	StyleID	CLASS_ID	LOCATION	ZDBH
111	长二进制数据	6	211080	拱墅区临一街1330105007001C	
112	长二进制数据	6	211080	拱墅区临一街1330105007001C	
113	长二进制数据	6	211080	拱墅区临一街1330105007001C	
114	长二进制数据	6	211080	拱墅区临一街1330105007001C	
115	长二进制数据	6	211080	拱墅区半山街1330105007001C	
116	长二进制数据	6	211080	拱墅区半山街1330105007001C	
117	长二进制数据	6	211080	拱墅区半山街1330105007001C	
118	长二进制数据	6	211080	拱墅区半山街1330105007001C	
119	长二进制数据	6	211080	拱墅区半山街1330105007001C	
120	长二进制数据	6	211080	拱墅区半山街1330105007001C	

StyleID	Style	Memos
1	长二进制数据	SolidPen
2	长二进制数据	NullHatch
3	长二进制数据	Symbol_063
4	长二进制数据	黑体
5	长二进制数据	SolidPen
6	长二进制数据	211080 211080 一般房与
7	长二进制数据	212080 212080 简单房与
8	长二进制数据	213080 213080 建筑中房
9	长二进制数据	214080 214080 破坏房与
10	长二进制数据	215080 215080 棚房面

Annotation	Type	Annotation	StyleID	Location
1	10	长二进制数据	73	长二进制数据
2	10	长二进制数据	73	长二进制数据
3	10	长二进制数据	73	长二进制数据
4	10	长二进制数据	73	长二进制数据
5	10	长二进制数据	73	长二进制数据
6	10	长二进制数据	73	长二进制数据
7	10	长二进制数据	73	长二进制数据
8	10	长二进制数据	73	长二进制数据
9	10	长二进制数据	73	长二进制数据
10	10	长二进制数据	73	长二进制数据

要素层Geometry几何对象，StyleID引入样式 符号化表达通过ID与Memos输出样式 注记存储Anno与Location位置，结合Style表达

Figure 5. Expression and association of geometric objects, annotation and symbolic information

图 5. 几何对象、注记与符号信息的表达关联

数据引擎针对 Walk 操作时的关键技术在于几何对象解析与数据库插入更新。构造几何体中点线面细分为多种类型，引擎库定义 7 类对象与之匹配。例如，简单线与多段折线的构造在部件数量及层次存储中存在区别，多环面与多洞面的构造根据部件数据解析几何环类型。数据库插入更新时需要维护表结构主键，确保 ID 唯一性；引擎在插入操作时，同步维护更新符号库及要素表 StyleID 与 Memos 字段，保证插入更新后的数据集在 Walk 中正常渲染。

2.4. 通用型格式转储

为了获得更多数据源支持，提供数据引擎持久兼容且高效解析能力，引擎内部构造了通用模型接收来自各个数据源的几何对象，如图 6 所示，引擎实现几何重建算法，针对例如超图 sdb 中的面状要素的定义为无序混合，重建算法在转储过程中的核心技术为扫描线算法对拓扑逻辑的重新排序，见图 7 所示，即面要素中外环与内岛的拓扑关系。

为满足平台的格式开放、兼容特性，引擎实现针对各厂商 GIS 数据格式的解析与转换能力，实现成果跨平台级访问能力。除支持 WKT、KML 等公共协议格式外，引擎提供对解析成果反序列化并写入数据库的接口，接口底层是通过 SQL 语句对关系型库的增删改查，其上封装成属性、空间几何、注记要素等模块，实现统一体系与标准化输出功能。

Geometry objects

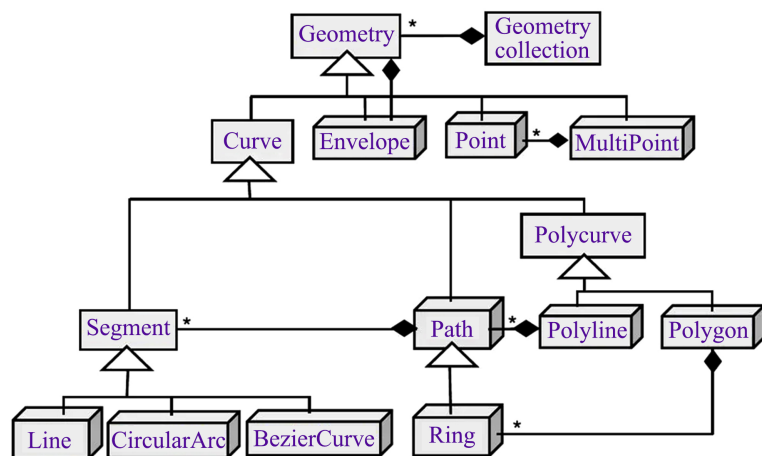


Figure 6. General geometric structure relationship

图 6. 通用型几何结构关系

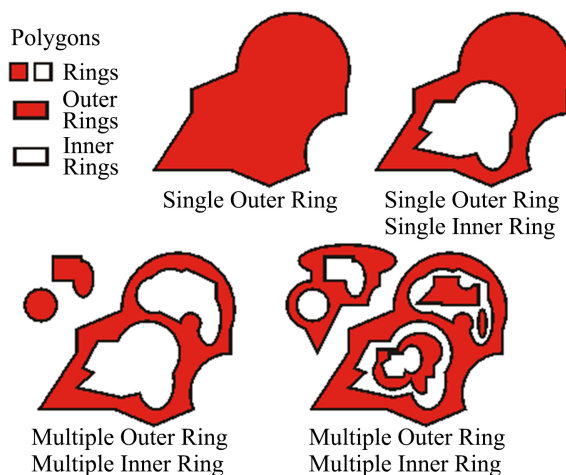


Figure 7. Topological logic of outer ring and inner island

图 7. 外环与内岛的拓扑逻辑

3. 应用

市测绘设计研究院近年来重点开展移动地图服务平台建设,并先后在新区管委会、东部新城规划处、住房和城乡建设委员会等局委办进行推广应用,在轻量级空间数据引擎的技术支撑下,成功搭建移动地图平台、东部新城移动服务平台、现代都市移动服务平台等业务平台,实现基于移动终端设备的空间数据和业务数据的一体化管理,构建一整套从数据呈现和查询分析到后台配置的完整的应用解决方案,改进和优化相关部门的日常工作手段。综合而言,基于轻量级空间数据引擎的移动地图平台具有如下 4 个特点:

1) 支持各类地理数据。平台支持多种地理数据,包括政务电子地图数据、影像数据、三维数据、360°街景数据等,并可在各类地图数据中进行自由切换,方便地图浏览和漫游。

2) 集成多种业务数据。集成多种不同内容、不同类型的业务数据,实现图片、文本、视频等多媒体数据的显示和浏览,并实现与地理底图的无缝集成。

3) 实现离线信息查询。实现各类专题信息的离线综合查询,一键式搜索平台中所有信息(如地名、地址、医疗、卫生等),搜索结果叠加在电子地图、影像数据上,并分类显示缩略信息。

4) 提供完善后台配置。提供数据准入原则,设置用户权限,不同的用户拥有不同级别的数据访问能力和功能模块使用权限。

4. 总结

本文通过对三家 GIS 厂商数据组织的分析,从底层解析二进制流构造并实现反序列化输出,实现数据开源、兼容与跨平台访问能力。数据引擎关键技术解决了转储过程中数据异常显示、渲染问题,保证输出成果在各家平台中正常使用。数据的通用访问为整合测量规范,建立协调统一标准,规范测绘成果管理,推进政府部门间政务信息数据的共建共享奠定基石,实现测绘成果规范化管理、测绘业务便捷化服务及多部门协同办理、测绘成果共享服务、有效提升测绘办事服务效率。

参考文献

- [1] 浙大万维地理信息系统有限公司. WALKScript 开发手册(对象化版本) [EB/OL]. <https://max.book118.com/html/2015/1003/26610242.shtm>, 2019-10-08.
- [2] Brray, Z. FileGDB Geometry 二进制格式分析[EB/OL]. <https://blog.csdn.net/barry114/article/details/6308732?locationNum=6>, 2019-10-08.
- [3] 郑建锋. 基于 VB 和 WalkScript 脚本二次开发实现对地籍数据入库和数据检查[J]. 现代测绘, 2012, 35(3): 60-61.