

关系数据库存储空间数据模型与结构分析

毛先成, 彭华熔

(中南大学地学与环境工程学院, 长沙 410083)

摘要: 空间数据的存储成为地理信息系统的核心问题。本文分析了 ERSI 的空间数据库引擎和 MapInfo 公司的 SpatialWare, 得出空间数据和属性数据在数据库体系下, 统一存储管理是必然趋势。另对空间数据的几种存储模型进行了分析, 指出全关系型的数据模型或面向对象的关系数据模型是最佳的存储方式。在此基础上, 提出了一种将空间数据存入关系型数据库的数据模型, 进而对基于 SQL Server 的全关系化空间数据结构进行了分析。

关键词: GIS; 空间数据模型; 关系型数据库; 空间数据结构

中图分类号: P208

1 引言

数据在传统的 GIS 系统中采用的是混合管理模式, 即由文件系统来管理空间数据, 由关系型数据库管理系统来管理属性数据, 通过在空间数据文件和关系型数据库中的属性数据之间建立关联, 构建应用系统。

这种两类数据分别存储的方式, 适用于单机的地学领域应用。随着 GIS 应用向分布式管理系统领域的转移, 空间数据的文件管理模式在实现数据共享、网络通信、并发控制及数据的安全恢复机制等方面出现了难以解决的问题。随着大型关系型数据库技术的日益完善, 其应用也日渐普及。于是现在的 GIS 软件都在向集成结构的空间数据库技术方向发展, 将空间数据和属性数据全部存储在关系型数据库中, 使二者紧密集成, 充分利用关系数据库管理系统处理和分析海量数据的能力, 真正实现将图形与属性融为一体的客户/服务器结构, 并与企业已有的信息系统集成, 使 GIS 融入 IT 主流^[1]。

2 空间数据应用管理现状分析

目前, GIS 软件与大型商用关系型数据库管理系统(RDBMS)的集成, 系采用空间数据引擎来实现。而今代表性的空间数据引擎产品有 ESRI 的 Spatial

Database Engine。它提供了一种标准化的多用户分布式空间数据管理方式, 替代了现有的基于文件的空间数据管理。作为应用服务器的中间件, SDE 在用户和现有的多种空间数据集之间提供了一个独立通用接口, 使来自客户端的基于文件或 DBMS 的空间数据库, 可以被统一存储到关系数据库中^[2](见图 1)。此外, SDE for Coverage 对 Arc/Info 中的 Coverages 数据和 shapefile 格式的空间数据做了大量开发, 并作为向 DBMS 数据管理转化的一个中间件。

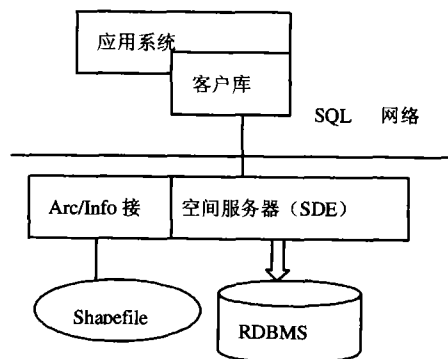


图1 SDE与RDBMS的集成

Fig.1 Integration of SDE and RDBMS

MapInfo 公司的 SpatialWare 是第一个在对象-关系型数据库环境下基于 SQL 进行空间查询和分析的空间信息管理系统。它不仅实现了在数据库中

收稿日期: 2004-04-19.

作者简介: 毛先成(1963-), 男, 湖南长沙人, 教授, 现从事地理信息系统与计算机应用的教学与研究工作;

E-mail: phr_775@163.com 或 phr.775@163.net

存储空间数据类型(如点、折线、区域等)的目标,而且建立了一套基于标准 SQL 的空间运算符,使得空间查询和分析能在服务器端进行。在服务器端, SpatialWare 将地图对象作为一个单独的列,并添加到数据库的表中,使得现有的数据库数据地图化。此外,其他没有地图的数据也可与数据库中的地图数据进行关系型的连接,以实现数据的地图化。

另外,一些主流关系型数据库系统如 Oracle、Informix、Sybase 等都开始提供空间数据的处理功能。其中以 Oracle 公司的数据库系统最为成熟, Oracle 8.04 版本中推出了空间数据管理工具—— Spatial Cartridge(SC)。它是一个真正的 Client/Server 机制的产品,提供了分布式处理能力,全部的 Oracle 安全管理机制,SQL 方式访问空间数据等功能。Oracle 在充分利用现有的关系数据库管理技术上,结合当前面向对象的新技术,提出对象——关系空间数据模型,实现了属性数据和空间数据的一体化存储。

当前已有一些厂商实现了用大型关系数据库来管理 GIS 数据,但国内还处在初步研究阶段。从实用性、普及性以及应用前景等方面的因素考虑,针对微软 SQL Server 数据库平台,我们提出了一种将空间图形数据存入关系型数据库的数据模型。

3 基于 SQL Server 的空间数据模型

SQL Server 是一个基于客户/服务器结构的可伸缩性、高性能的大型关系数据库管理系统(RDBMS)。它的主要优点有:(1)SQL Server 是基于客户/服务器结构的数据库管理系统,其使该系统能够在服务器和客户机之间平衡荷载,在整体性能上与传统的基于文件共享方式数据库管理系统相比有大幅度提高。(2)SQL Server 支持大规模的应用,支持上千个并发用户,并提供自动锁功能,使得并发用户可以安全而高效地访问数据。(3)SQL Server 与 Windows NT 操作系统紧密结合,由于同是微软公司的产品,所以 SQL Server 可以从核心一级利用许多 NT 的技术和功能^[9]。

3.1 空间数据的存储方式

SQL Server 作为一种关系型数据库,而图形数据的关系型数据库存储方案有关系模型和对象关

系模型两种。关系模型具有结构统一、面向记录、字段为原子等特点^[4],在 ESRI 的 SDE 中,就采用了关系模型的解决方法。关系模型的具体存储方式:

(1)数值存储方式^[5] 地理实体以数值存储模型存放时,空间数据存储在几何表中,特征表的几何列存储指向几何表的指针。每一个地理实体的空间数据由一系列坐标对来描述,其坐标值用数值类型来存储(如图 2 所示)。如果坐标对数超过了几何表中的定长坐标对数时,就分行存储。

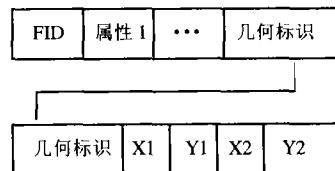


图 2 数值存储方式
Fig.2 The Mode of Number Storage

(2)二进制存储方式 地理实体的空间数据以二进制形式(BLOB)存储在几何表中,每一个地理实体只对应几何表中的一行(如图 3 所示)。

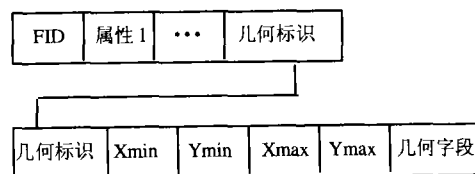


图 3 二进制存储方式
Fig.3 The Mode of Binary System Storage

而对象关系模型是对关系模型的扩展,具有面向对象的更加丰富的类型系统,允许用户定义结构类型的复杂字段类型。它有标准数据类型描述某类(点、线、面)地理实体的属性,由函数描述其方法,从而形成对象,任一地理实体即可通过对象的实例来存储。在 Oracle 8i 以后版本的 Spatial Option 中就使用了对象关系模型这一解决方案(如图 4 所示)。

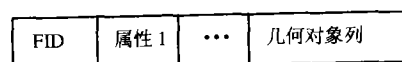


图 4 对象关系模型方式
Fig.4 The Mode of Object Relational Model

虽然对象——关系模型既可充分利用成熟的关系型数据库强大的数据管理功能和可操作性,又兼具面向对象数据库的对象和类的概念及数据模型的扩充,可以很好地实现空间数据的管理,但对 SQL Server 进行这种扩展开发量很大,许多关键技术需要数据库开发商来完成,因而对此不做深入研究。关系模型将点、线、面等空间数据类型作为基本数据类型,建立全关系化的空间数据模型,并在标准 SQL 语言中扩充空间数据和空间数据项的表示方法,在 SQL Server 上采用这种空间数据模型在语言的用户表示上与标准 SQL 保持统一,并提供了数据模型、数据表示和数据完整性约束,而且减少了空间数据库扩展的开发量。因此,我们将采取第一种方案建立基于 SQL Server 的空间数据模型。

3.2 空间数据模型

空间数据模型一般可分为几何元素、空间要素、空间实体、图层 4 层结构,每一种结构都对应于空间数据的一种表达。层是由空间实体组成,而空间实体则由空间要素组成,空间要素最终可以分解为点。

(1)几何元素。点、线、面是空间要素的基本构成单元,它可以构成点串、线串、多个多边形组成的复合多边形。(2)空间要素。表示空间几何特征,它是一个或多个几何元素的有序组成。(3)空间实体。其由空间要素构成。例如:一个带洞的多边形的外部环和内部环被认为是 2 个不同的空间要素构成的复合多边形。(4) 图层是一定空间范围内具有相同属性要素的同类地理空间实体而组成的一个有机集合。针对 SQL Server 数据库的特点以及参考当前一些空间数据模型,建立如图 5 的数据结构模型。

在这种数据模型下,不定长的空间几何体坐标数据以二进制数据块的形式被关系数据库管理。即坐标数据被集成到 RDBMS 中,形成空间数据库,在 SQL Server 数据库中图层以表的形式进行组织和表达,形成表图层。为了方便管理,将不同图层的空间数据分别放在不同数据库的表中,同一层中不同类型的空间数据存放在不同的表中^[9]。

3.3 空间数据结构

空间数据模型是对现实世界的模拟。而空间数据结构是对空间数据模型(空间实体)的一种有效的内部数据表示。在确定了数据模型后,我们就可

以对空间数据结构进行设计,虽然空间库存放在 SQL Server 中,可根据不同的需要设计数据结构,但一些基本的表格还是必不可少的。限于篇幅,这里只给出基本数据结构。

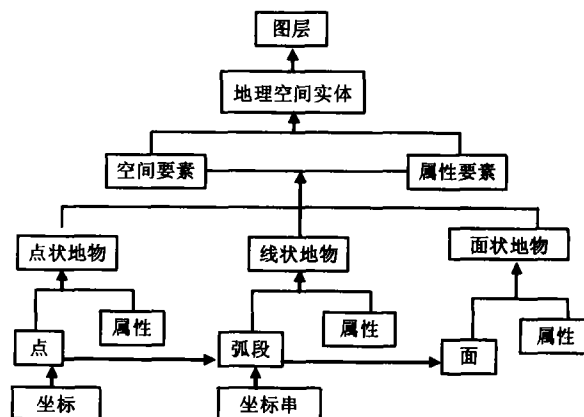


图 5 基于 SQL Server 的空间数据模型

Fig.5 Spatial data model based on SQL server

(1)点状地物基本结构

point

```

{
    int   pointed; /* 内部编码 */
    float pointx; /* 点状地物的 X 坐标 */
    float pointy; /* 点状地物的 Y 坐标 */
    ...
}

```

(2) 线状地物基本结构

line

```

{
    int   lineid; /* 内部编码 */
    int   nodesnum; /* 结点数 */
    image nodeid /* 构成线的结点号序列 */
    float lxmin; /* 坐标范围 */
    float lymin;
    float lxmax;
    float lymin;
    ...
}

```

其中,nodeid 字段是用来存储构成该线状地物的结点坐标序列的。由于构成线状地物的结点数是不固定的,所以该字段也是不定长的,因此这里采用 SQL Server 提供的 image 数据类型,它可以存储每行最大为 2GB 的二进制大型对象 (Binary Large

Objects, BLOB)。

(3) 面状地物基本结构

①面信息

```

polygon
{
    int  polygonid; /* 内部编码 */
    int  arcsnum; /* 该多边形包含的外边界弧
                  段数 */
    int  islandsnum; /* 该多边形包含的岛屿数 */
    image arcid; /* 构成面的弧段号序列 */
    float pxmin; /* 坐标范围 */
    float pymin;
    float pxmax;
    float pymax;
    float xlab; /* 内部标识点坐标 */
    float ylab;
    ...
}

```

其中, arcid 字段是用来存储构成该面的弧段号序列的。由于一个面中包含的弧段数有多有少, 是不固定的, 所以该字段也是不定长的, 因此这里也采用 SQL Server 提供的 image 数据类型, 另外, 弧段号的绝对值是弧段信息表中的 id 字段的值, 有正有负, 负号表示孤岛。

②弧段信息

```

arc
{
    int  arcid /* 组成面状地物的弧段代码 */
    int  startnode /* 起始结点号 */
    int  endnode /* 终止结点号 */
    int  nodesnum /* 结点数 */
    Image nodeid /* 构成弧段的结点序列 */
}

```

其中, nodeid 字段是用来存储构成该弧段的结点坐标序列的。由于构成弧段的结点数有多有少,

是不固定的, 所以该字段也是不定长的, 因此这里也采用 SQL Server 提供的 image 数据类型。

③结点信息

```

node
{
    int  nodeid /* 组成弧段的结点代码 */
    float pointx; /* 结点的 X 坐标 */
    float pointy; /* 结点的 Y 坐标 */
    int  arcsnum /* 与其相连的弧段数 */
}

```

4 结语

利用关系型数据库来存储 GIS 的空间数据是一先进的技术, 它实现了属性数据和空间数据的一体化存储, 从理论上保证了数据的完整性和数据共享。笔者借助现有成熟的空间数据库技术成果, 针对 SQL Server 数据库的特点, 提出基于 SQL Server 的空间数据模型, 对空间数据结构及其存储方式进行了初步的研究, 还有待今后进一步深入。

参考文献

- [1] 江崇礼, 王丽佳. 基于 RDBMS 的地理信息集成数据库系统. 计算机工程, 2002, (11).
- [2] 余倩, 吴国平等. 空间数据全关系型存储的研究——以 MapInfo 空间数据存入 SQL Server 数据库为例. 遥感信息, 2001, (1).
- [3] 王卫安, 王晓艳. 用 Client/Server 数据库存储 GIS 图文数据. 微型电脑应用, 1999, (6).
- [4] 孙红春, 王卫安. 基础地理信息图文一体化数据模型. 测绘通报, 2001, (1).
- [5] 王春晓, 李佳田等. 集成型空间数据库技术探讨与应用实例. 遥感技术与应用, 2003, (2).
- [6] 戴欣明, 章孝灿. 一种关系化的 GIS 空间数据结构. 计算机应用研究, 2003.

下转 P83

The Design and Realization of China Geomagnetic GIS

YANG Zuhu^{1,2}, RONG Kai^{1,2}

(1 Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2 Graduate School of the Chinese Academy of Sciences, Beijing 100081, China)

Abstract: This paper explores the feasibilities of applying GIS to the management of geomagnetic related information, taking the design and realization of the China Geomagnetic Geographic Information System (CGGIS) as an example.

It introduces the brief design process, including the network structure, logical structure and functional structure of the system. It also describes some of the key functions that have been successfully realized, such as the E-card function of the geomagnetic observation station, the GIS aided location design function of the geomagnetic observation stations, and the geomagnetic model visualization function, etc.

The result shows that the database and its application system based on GIS technology not only can manage the related geomagnetic chart data of sea-amount efficiently, but also can be helpful to present the obtained geomagnetic data expediently and directly, integrate the related spatial data efficiently and discover the hidden laws concerned the geomagnetic field. Such a Spatial Enabled Geomagnetism Information System (SEGIS) deserves further research and application.

Key words: geomagnetic; GIS; design

上接 P79

The Research of Database-based Storage for Spatial Data

MAO Xiancheng, PENG Huarong

(Central South University, Changsha Hunan 410083, China)

Abstract: The handling of spatial data is a key issue of Geographic Information System (GIS). The spatial data can be analyzed and applied when stored and operated in GIS. So, it is important for GIS to store and manage spatial data. At first, this paper introduces the present research status, then analyses ESRI's Spatial Database Engine (SDE) and MapInfo's SpatialWare. It is concluded that the development direction that spatial data and attribute data are managed under Database Manage System (DBMS) is a necessary trend for unified storage and management. Then, depended on the practical technology, the paper discusses storage models of spatial data, and points out a good storage model is Relational Database Manage System (RDBMS) or Oriented Relational Database Manage System (ORDBMS). So, a data model of spatial graphical data storage in RDBMS is presented in this paper. Finally, in light of this model, the paper analyses a relational spatial-data structure based on SQL Server.

Key words: GIS; spatial-data model; relational database; spatial-data structure