

# 空间数据仓库元数据系统的研究与设计

董星星<sup>1</sup>, 李景文<sup>1</sup>, 陈大克<sup>1,2</sup>, 陈肖磊<sup>1</sup>

(1. 桂林工学院 土木工程系, 广西 桂林 541004; 2. 广西师范大学, 广西 桂林 541004)

**摘要:** 介绍了空间数据仓库元数据的概念、功能及元数据系统的组成要素, 重点研究了空间数据仓库元数据系统的设计, 提出了邦联式与客户机/服务器、浏览器/服务器相结合的系统结构设计方案, 并进一步设计了空间数据仓库元数据系统的功能和元数据库信息模型。该系统研究与设计的目的是为了较好地解决空间数据仓库中海量数据的管理、共享及互操作等问题。

**关键词:** 空间数据仓库; 元数据; 元数据标准

**中图分类号:** P208

**文献标识码:** B

**文章编号:** 1671-3044(2007)01-0055-04

## 1 引言

随着空间信息获取技术的快速发展, 人类拥有的空间数据量急剧增大, 这些海量数据给人们带来方便的同时也带来了信息过量、难以消化、信息源多元化, 难以充分利用等问题。因此, 如何共享、分析、挖掘和利用这些海量信息资源是信息化建设的关键问题之一。空间数据仓库就为空间数据的挖掘和分析决策提供了一种新的理念, 通过对空间数据的清理、转换、抽取等操作实现对分散的、各自独立的多种地理空间数据库进行统一集成和管理, 为空间决策、空间 OLAP 和空间数据挖掘等提供必要的集成信息和服务。而空间数据仓库元数据是空间数据仓库实现的基础, 它规范了空间数据仓库中的数据模式、来源以及抽取和转换数据的规则等, 对促进空间数据的管理、使用和共享具有重要作用。因此, 元数据是空间数据仓库系统不可缺少的部分, 设计良好的元数据系统是空间数据仓库技术的重要环节。本文通过对空间数据仓库和空间元数据技术的研究, 提出了一种较为全面、系统的空间数据仓库元数据系统的设计方案。

## 2 空间数据仓库的元数据概述

### 2.1 空间数据仓库元数据的概念

空间数据仓库元数据是关于空间数据的数据, 用于描述空间数据仓库中数据的内容、质量、表示方式、空间参照系、管理方式及其他相关的背景信息。

它主要由大地成果数据、数字高程模型数据、栅格地图数据和正射影像数据的元数据内容组成<sup>[1]</sup>。空间数据仓库元数据的使用有利于帮助和促进人们有效地定位、评价、比较和使用多个空间数据库中的空间数据, 实现真正意义上空间信息的共享和互操作。

### 2.2 空间数据仓库元数据的主要作用

空间数据仓库元数据可用来帮助数据提供者和数据使用者解决数据转换、沟通和理解等问题。其功能归纳起来, 主要有以下三个方面<sup>[1-4]</sup>:

(1) 帮助数据生产者有效组织、管理和维护空间数据仓库数据。可保证即便其主要工作人员调离, 也能对过去生产的空间数据仓库数据有较为全面的了解, 这样可以有效、持久地实现对数据的维护、更新, 确保数据生产者对数据的持续投资。

(2) 为用户提供海量数据的目录信息。根据数据目录、数据交换中心等提供的元数据内容, 用户可以很容易从海量空间数据中发现所需数据, 并对数据进行共享和优化等操作。如美国的国家地理空间数据交换中心(NGDC)。

(3) 提供数据转换方面的信息。通过元数据信息, 用户可以接受和理解空间数据仓库中抽象的概括性数据, 并与自己的数据集进行集成, 进行不同方面的分析决策, 使海量的空间信息实现有效共享, 从而发挥空间数据仓库元数据最大潜力。

### 2.3 空间数据仓库元数据系统构成

一个完整的元数据系统通常由三部分组成, 即元数据标准、元数据操作工具和元数据库<sup>[2]</sup>, 如图 1 所示。其中, 元数据标准是由元数据的应用需求和

收稿日期: 2006-08-08; 修回日期: 2006-12-12

基金项目: 国家自然科学基金项目(桂科攻 0330008-7); 广西自然科学基金项目(0448076)。

作者简介: 董星星(1982-), 女, 江苏江都人, 硕士研究生, 主要从事 GIS 技术应用与开发研究。

空间数据仓库数据的特点决定的,是实现空间信息共享的前提条件和保障;元数据操作工具负责对元数据库进行管理和维护,并向用户提供基于元数据的空间信息查询服务;元数据库用于管理数据对象和系统管理的元数据信息等。

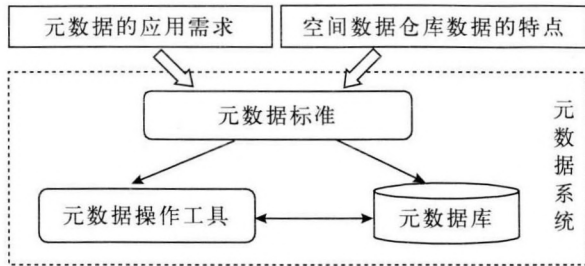


图 1 空间数据仓库元数据系统构成

### 3 空间数据仓库元数据系统设计

空间数据仓库元数据系统设计主要包括:空间数据仓库元数据标准的确定、系统结构设计、功能模块设计及元数据库设计。

#### 3.1 空间数据仓库元数据标准的确定

由空间数据仓库元数据系统组成可知,元数据标准是元数据系统设计的基础。因此,系统设计的首要任务就是确定好元数据的标准化内容。空间数据仓库元数据是由大地成果数据库元数据、数字高程模型数据库元数据、矢量地图数据库元数据、栅格地图数据库元数据和正射影像数据库元数据内容组成的。因此,考虑到空间数据仓库元数据的上述特性及实际操作中的具体情况,参照美国 FGDC 地理

空间数据元数据的内容标准<sup>[5]</sup>和我国《国家基础地理信息系统(NFGIS)元数据标准草案(初稿)》<sup>[6]</sup>,确定空间数据仓库元数据内容为十类信息,它们分别为:标示信息、数据质量信息、数据表示信息、空间参照信息、数据范围信息、数学基础信息、数据发行信息、数据安全信息、数据联系信息、数据时间信息,如图 2 所示。



图 2 空间数据仓库元数据标准化内容

#### 3.2 系统结构设计

随着分布式空间数据库和多源数据库集成技术的迅速发展和广泛应用,传统的单一元数据库集中式管理方法已远远满足不了多元化的空间数据仓库管理理念的需求。因此,本文提出邦联式与客户机/服务器、浏览器/服务器相结合的系统结构设计方案,对多种地理空间数据库(即大地成果数据库、数字高程模型数据、矢量地图数据库、栅格地图数据库及正射影像数据库)的元数据加以管理,并通过应用 UML 语言对不同类型的元数据进行统一的一体化建模,从而有效集成不同类型的空间数据仓库元数据。系统结构如图 3 所示<sup>[1,7,8]</sup>。

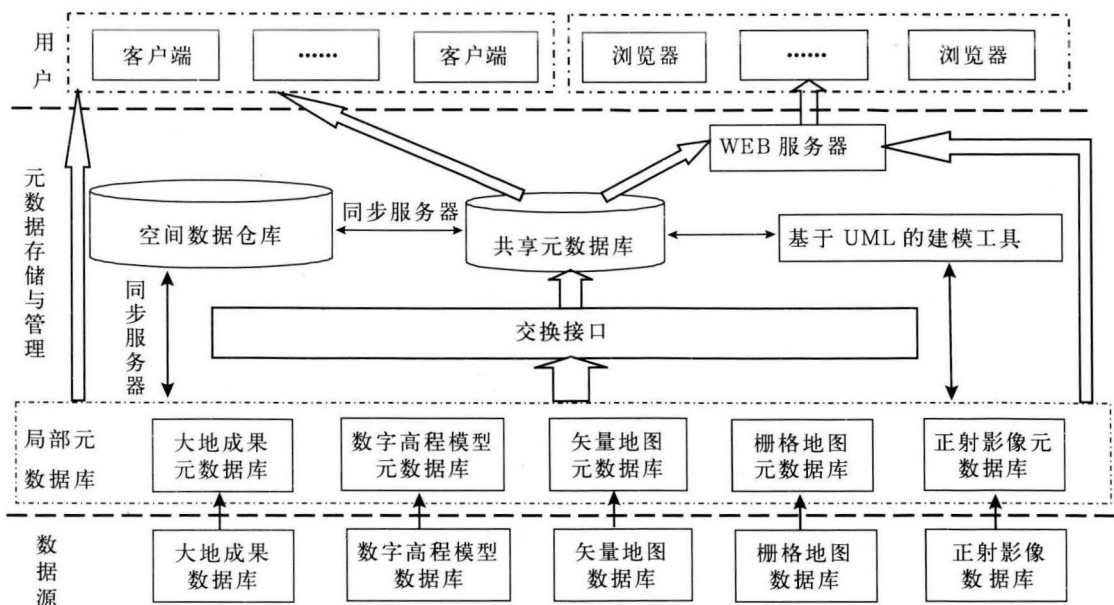


图 3 系统结构图

从系统结构图可以看出,空间数据仓库元数据系统通过服务器从各个源数据库中抽取它们各自的元数据,把需要共享的元数据映射到共享元数据库中,而较低粒度的元数据则存储于各自的元数据库中。在该系统中,局部元数据库与共享元数据库之间需要交换接口,用于获取局部元数据库中的元数据;局部元数据库可以采用异构的表示形式,而共享元数据库必须采用统一的元数据表示形式。用户可以在客户端通过局域网查询共享元数据库和各局部元数据库中的元数据,也可在浏览器上通过 Web 服务器在网上查询元数据库中的元数据。同时用户可通过元数据访问空间数据仓库中的数据,其实现方法是在元数据库中注册空间数据,通过 FTP 或 HTTP 描述的方式进行空间数据地址的描述,用户则根据元数据的这种地址描述最终获得所需数据。需要说明的是,在元数据库服务器和空间数据仓库服务器间有一个同步服务器,以确保空间数据仓库数据与相应元数据的同步性和一致性。

### 3.3 功能设计

空间数据仓库元数据系统包括以下四大功能模块<sup>[1,9,10]</sup>:

#### (1) 元数据输入

该功能是面向普通用户的,包括元数据结构信息输入和元数据内容输入,其中元数据结构信息用于描述元数据在元数据库中的组织方式的信息,它是客户端解码的依据;元数据内容输入是元数据信息录入的核心部分,通过输入界面,数据生产者可以便捷地输入空间数据仓库元数据内容。输入功能还包括格式检查、内容查错等功能,保证输入数据的正确。

#### (2) 元数据查询和数据获取

此功能也是面向普通用户的,是元数据系统功能最重要的组成部分。它提供方便的菜单和按钮,能快捷地浏览自己所关心的元数据信息,同时提供各种关键字(如有关主题词、关键字查询、空间范围查询、时间范围查询、数据集系列查询等)查询,从元数据库中,检索出用户所需要的元数据。获取数据、实现共享,是元数据系统的最终目标。该功能的实现主要是通过元数据中提供一个指向数据存储点的 URL,用户只需在浏览器上点击这个 URL 就可以获得数据。

#### (3) 元数据更新

该功能是面向管理员的,具体包括对空间仓库元数据的增加、删除、修改及提交等操作。其实现方法主要是采用编辑器形式,该编辑器提供一个可视

化、交互式的编辑更新环境;同时整个过程需同步协调技术进行控制,确保元数据更新与空间数据仓库数据更新的一致性。

#### (4) 元数据管理

该功能也是面向管理员的,包括对元数据库表格的建立、删除、结构的修改、记录操作、对用户授权及其他操作。其目的是实现对空间数据仓库系统中的元数据进行安全控制,防止非法用户获取空间数据仓库系统元数据库中的元数据。

### 3.4 元数据库设计

元数据涉及空间数据仓库的各个方面,对元数据库的设计直接影响到整个空间数据仓库系统的性能。因此,本文提出了一个较易实现、结构清晰的元数据库信息模型,如图4所示。该模型由字典表、元数据表、辅助表和系统管理表四个方面的信息构成。

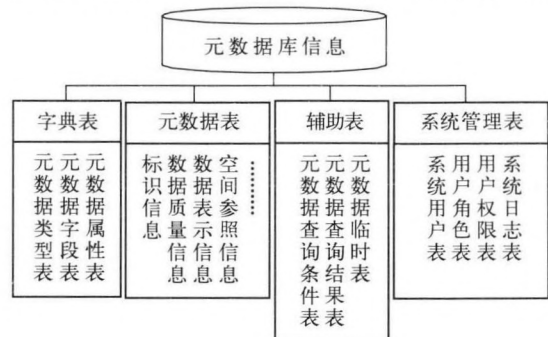


图4 元数据库信息模型

(1)字典表是由元数据类型表、元数据字段表和元数据属性表组成,用来存储元数据的结构特征,并实现对元数据类型、属性、字段的增加、删除、修改等操作。

(2)元数据表是由上述已确定的十类元数据内容组成(图2),它是对空间数据仓库中的 DLG、DRG、DOM、DEM、大地成果数据进行具体的描述,是元数据库的核心内容。该部分涉及的元数据数据量大、结构复杂,应采用层次结构加以存储和管理。

(3)辅助表包括元数据查询条件表、元数据查询结果表和元数据临时表,用于暂时存储相关查询信息,以便快速浏览查询结果及近期的历史查询。

(4)系统管理表由系统用户表、用户角色表、用户权限表和系统日志表组成<sup>[11]</sup>。其中系统用户表是用来记录用户的名字、登录口令、用户所属的角色等信息;用户角色表记录系统中已定义的所有角色、每个角色拥有的权限等;用户权限表用来记录对系统的各种操作权限的名称和编号;系统日志表是用来记录各种系统事件,如用户对元数据的结构和内容的编辑等。

#### 4 结 论

本文系统地研究了空间数据仓库元数据技术,提出了一种联邦式与客户机/服务器、浏览器/服务器相结合的空间数据仓库元数据系统的设计方案。研究表明,本文提出的设计方案优点在于:与传统的地理元数据系统相比,该系统结构灵活,扩展性强,能较好地实现对不同类型元数据有效管理;元数据查询和浏览方便,且根据元数据可快捷地获取所需数据,同时也较好地解决了元数据更新和空间数据仓库数据更新的同步性问题。系统在一定程度上解决了分布式、异构空间数据仓库数据的有效管理和共享等问题,使空间数据仓库技术更好地服务于空间数据挖掘和空间决策分析。同时,它对其他元数据库系统的设计,也具有一定的参考价值。

#### 参考文献:

- [1] 吴金华. 空间数据仓库的元数据研究[D]. 武汉:武汉大学,2004.
- [2] 何建邦, 闫国年, 吴平生. 地理信息共享的原理与方法[M]. 北京: 科学出版社, 2003.
- [3] 祝方雄. 城市地理信息系统中元数据标准的研究[J]. 城市勘测, 2001, (3): 37~41.
- [4] 陈惠荣, 游 雄. 地理空间元数据及其相关技术的探讨[J]. 测绘学院学报, 2003, 20(4): 290~292.
- [5] The FGDC Content standard for Digital Geospatial Metadata [DB/OL]. [http://www.nbi.gov/datainfo/metadata/standards/BRD\\_metadata/version2/](http://www.nbi.gov/datainfo/metadata/standards/BRD_metadata/version2/).
- [6] 国家基础地理信息系统(NFGIS)元数据标准草案(初稿). <http://nfgis.nsd.gov.cn/nfgis/chinese/bz/mt0.htm>.
- [7] 戴超凡, 刘青宝, 黄宏斌, 等. 数据仓库中的元数据管理[J]. 计算机工程与科学, 2003, 25(4): 54~57.
- [8] 张英朝, 邓 苏, 张维明. 数据仓库元数据管理研究[J]. 计算机工程, 2003, 29(1): 8~10.
- [9] 张友水, 阮仁宗, 苏 勤. 空间数据仓库元数据及其实现[J]. 计算机应用, 2003, 23(4): 102~105.
- [10] 王家耀. 空间信息系统原理[M]. 北京: 科学出版社, 2001.
- [11] 祝方雄. 宁波市基础空间信息元数据管理系统的设计与实现[J]. 城市勘测, 2005, (5): 20~23.

## The Research and Design of Metadata System in Spatial Data Warehouse

DONG Xing-xing<sup>1</sup>, LI Jing-wen<sup>1</sup>, CHEN Da-ke<sup>1,2</sup>, CHEN Xiao-lei<sup>1</sup>

(1. Department of Civil Engineering, Guilin University of Technology, Guilin, Guangxi, 541004;

2. Guangxi Normal University, Guilin, Guangxi, 541004)

**Abstract:** With the rapid development of collected spatial data technology, the spatial data which the human has increased sharply. Unfortunately, the magnanimous data is too excessive to be assimilated. How to effectively use and share the collected data is an important problem which is more urgent to resolve. Metadata system in spatial data warehouse provides a good method to resolve the problem. Therefore, this paper studies the technology of metadata in spatial data warehouse. Firstly, it discusses the concept and function of metadata in spatial data warehouse and the components of metadata system in spatial data warehouse. Secondly, it chiefly explores how to design the metadata system in spatial data warehouse, which concludes defining the standard of the metadata in spatial data warehouse, designing the structure and function of the system and the database of metadata.

**Key words:** spatial data warehouse; metadata; standard of metadata