

# 分布式网络环境下地理信息元数据框架研究

梅 琨<sup>1</sup> 边馥苓<sup>1</sup>

(1 武汉大学空间信息与数字工程研究中心, 武汉市珞喻路 129 号, 430079)

**摘要:** 在研究现有的地理信息元数据标准定义的元数据框架及其管理模式的基础上, 提出了一种可操作的、面向对象的元数据内容框架和分级的、基于地图的元数据组织方式, 方便了分布式网络环境下地理信息服务的检索、查找和获取。

**关键词:** 元数据; GIS; Web 地图服务; 分级元数据仓库

**中图法分类号:** P208; TP311.12

随着计算机和 GIS 技术特别是网络通信技术的发展, 地理空间数据量呈几何级数增长, 如何有效地管理和应用存储着这些海量数据的 GIS 数据库, 成为 GIS 专家面临的一大挑战。分布式的地理信息服务(GIServices)是解决这一问题的有效途径。但由于有些地理空间数据未直接公开、格式不统一以及服务的异构性等原因, 很难访问到远程的 GIS 数据库和 Web 地图服务。在这种情况下, 描述地理空间数据和服务的元数据就变得尤为重要。尽管地理信息元数据系统的建设越来越普及, 现有的地理信息元数据的应用仍然存在许多问题。本文在研究了现有的地理信息元数据标准定义的元数据框架及其管理模式的基础上, 提出了一种面向对象的、可操作的地理信息元数据内容框架。

## 1 地理信息元数据及相关标准

通常, 元数据可以理解为关于数据的数据。一些研究者认为, 元数据是描述某种类型资源(或对象)的属性, 并对这种资源进行定位和管理, 同时有助于数据检索的数据。在本文中, 笔者将描述分布式地理信息服务的元数据定义为一种帮助用户访问、存储和操纵集中式或分布式 GIService 的信息。

国际上用于地理空间数据的元数据标准已有很多, 比较典型的有 CSDGM 元数据标准<sup>[1]</sup>和

ISO/TC211 19115 元数据标准<sup>[2]</sup>。此外, 我国的相关部门也以 ISO/TC211 所提出的草案为基础, 参考美国 FGDC 的 CSDGM, 并结合我国国情, 于 2005 年 10 月 1 日正式发布了《地理信息元数据》国家标准(GB/T 19710-2005)。无论是哪种标准, 都含有两个层次的元数据信息: 目录信息和详细信息。同时, 标准将元数据内容以必选、一定条件下的必选和可选的方式组织, 使其更容易在分布式的网络环境中采用。

然而, 元数据标准只是系统实现数据维护、数据共享、数据查询和数据传输等功能所涉及数据的元数据的描述。通常情况下, 元数据与数据本身没有建立直接的联系, 有时还会出现地理空间数据对象已经被修改或者被删除了, 而元数据却依然存在的情况。在分布式网络环境下, 地理空间数据对象和元数据记录还经常会从一个位置被移动或拷贝到另一个位置, 因此, 实现地理信息元数据和地理信息服务的联系, 增强地理信息元数据在实际应用中的可操作性就是本文研究的目的。

## 2 现有的元数据管理模式

目前, 元数据管理主要有两种方式: 一种是创建一个专门的元数据库或目录, 这种方式被称为“数字图书馆”; 另一种是建立分布式的元数据仓库, 并用网关服务器将它们连接起来, 这种方式被称为“数据交换中心”。这两种方式都能够帮助用

户发布和查找分布式的 GIS 服务。

和数字图书馆相比, 数据交换中心为地理空间元数据提供了更为灵活的管理方式。此外, 数据交换中心的一个最大优势是用户可以通过 Z39.50 协议<sup>[3]</sup> 同时查询分布在各地的多个空间数据交换中心的元数据仓库。然而, FGDC 的数据交换方式对于分布式 GIS 服务来说, 还存有隐患: FGDC 的空间数据交换节点是同级部署的, 未经分类; 无法在查询结果集中进行多次查询; 所有的空间数据交换中心都必须通过 FGDC 的网关服务器注册它们的节点。

### 3 元数据框架设计

针对现有的地理信息元数据标准对元数据内容的定义和元数据管理模式存在的问题, 需要重新设计地理信息元数据的基础框架, 为此, 本文提出了一种可操作的元数据内容框架, 并且改进了元数据的组织管理方式。

#### 3.1 可操作的元数据内容框架

为了实现元数据与数据对象的链接, 本文采用面向对象的建模技术将元数据重新封装。这种面向对象的元数据框架包括两部分内容: 描述元数据和操作元数据。描述元数据包括传统标准定义的元数据信息, 如地理空间数据标识信息、空间参照系信息、内容信息、发布信息等。操作元数据包括一些可以自动应用于某些特定的 GIS 操作的机读信息, 如地图显示、空间分析、GIS 建模等。此外, 元数据中还记录了 GIS 服务的定位信息。

增加操作元数据部分, 有助于地理空间数据对象服务、Web 地图服务和 GIS 组件服务的互操作和集成。下面具体介绍关于这三种地理信息服务元数据的内容框架的可行性设想。

地理空间数据操作元数据的设计应当考虑到地理空间数据的用途和将要进行的操作。这里列举其中 4 种操作元数据: 地图显示元数据、空间查询元数据、空间操作元数据和数据链接元数据。地图显示元数据描述了要素类型、属性字段、制图符号、屏幕分辨率、颜色质量和比例尺等, 用户可以利用这些元数据定制显示图层。空间查询元数据描述了能使用的查询语言、查询语法、查询接口和结果显示方式等。空间操作元数据记录了数据对象可以执行的空间操作和操作日志。数据链接元数据描述的是在线访问和下载数据的机制, 包括本地访问接口、远程访问接口、数据注册情况、数据压缩方式等。

Web 地图服务操作元数据的设计使得访问、发布和集成 Web 地图服务变得更加容易。它包括访问方式元数据、图形工具元数据和图层信息元数据。访问方式元数据描述了远程访问 Web 地图服务的通信协议、通信语言、浏览器需求、网络带宽需求等信息。图形工具元数据描述了地图服务能实现的各项查询和分析功能以及用户要实现这些功能所需要的设备。图层信息元数据记录了图层列表、每幅图层的地图范围以及数据对象链接等信息。

GIS 组件操作元数据的设计主要是反映 GIS 组件所能实现的功能, 具体包括功能列表、数据输入、数据输出、运行环境需求和组件注册元数据。功能列表元数据记录了 GIS 组件所能实现的功能。数据输入元数据描述了输入数据的格式、投影方式、坐标系统、拓扑关系等信息。数据输出元数据描述了输出数据的格式、输出精度、操作结果等信息。运行环境需求元数据描述了计算机的硬件需求、虚拟机需求和组件大小、类型等信息。组件注册元数据是在组件通过 UDDI 节点注册后自动生成的, 包括惟一组件标识 ID、功能所属类别、涉及 GIS 应用等信息。

可操作的元数据内容框架的提出将有助于 GIS 的分布式处理、地图的精确显示和数据的网上交换。同时, 将元数据封装在数据对象中能减少元数据的丢失, 且一旦数据被更新, 可及时更新元数据描述的内容。

#### 3.2 元数据管理框架

##### 3.2.1 分级组织方式

目前, FGDC 地理空间数据交换中心的组织方式是平行的, 同一级别上有成千上万的数据交换中心节点, 不利于 GIS 用户找到自己想要的数据所在的交换中心。因此, 本文认为应采取分级的方式组织节点。用户向数据交换中心注册地理信息服务时, 可以按照地理信息服务所在的地理位置或所表达的主题等进行分级, 注册的同时封装在地理信息服务中的元数据被提交到数据交换中心, 从而帮助用户在网络上更容易地检索、查询和发布各种 GIS 服务。如果这些被更新, 则需要重新注册。

图 1 给出了地理空间数据库集的分级元数据组织方式。该例中的数据对象——城市土地利用类型和城市道路数据等, 连同用来描述它们的元数据被注册到市级元数据交换中心的分节点上, 市级元数据交换中心的分节点再被注册到省级元数据交换中心的节点上, 依此类推。

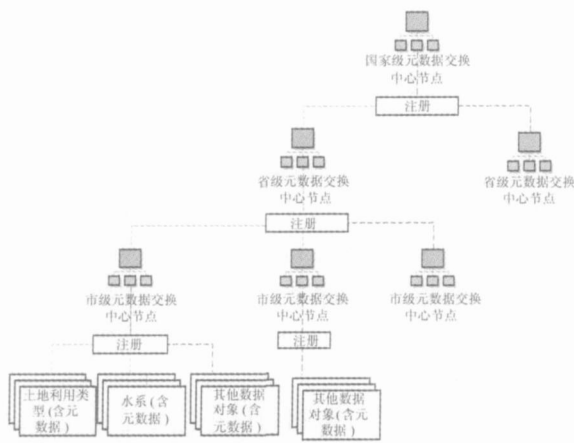


图1 地理空间数据库集分级组织框架

Fig. 1 Hierarchical Metadata Repository Framework for Geospatial Datasets

这种树状分级结构的优点是可以提供意思表达更明确的元数据目录结构,从而改进系统的检索机制。由于父节点存储了下属子节点的所有注册信息,系统能够查询到其相应的绑定信息,然后就可以调用子节点查询相应服务,获取相关元数据信息。此外,同一个数据对象可以同时被注册到不同的节点上。如城市道路层可以被注册到市级数据交换中心的节点上,也可以被注册到省级交通运输系统数据交换中心的节点上。这种分级的元数据索引框架同样可以用于GIS组件和Web地图服务的管理,它们也可以进行多重注册。地理空间数据的所有者和GIS软件程序员将他们的产品注册到多个元数据交换中心的节点上,用户和应用程序就可以利用分级元数据仓库,通过特定的目录查询他们所需要的数据或程序。

### 3.2.2 地图组织方式

由于元数据所涉及的内容十分广泛或GIS用户出于某种特定的目的,在对元数据进行查询时,常常会用到地名查询,如查询某个省内的元数据,或者是某条河流流域范围内的元数据,甚至是某地理要素附近的元数据,通过添加地名字典表,并将其与元数据的唯一标识建立关联就可以实现地名查询。但是,地名字典表的建立和维护工作十分困难,本文认为可以通过地图组织元数据,通过元数据中提供的定位信息建立表示元数据覆盖范围的格网,再在格网层上加入背景数据,以便于从空间位置上显示索引范围,进行更为直观、方便的图形化的元数据检索与查询。背景数据可以是大比例尺的境界、点状居民地、水系、国省道骨干交通网、铁路等,也可以是某一特定的地理空间要素。利用这些背景数据,就可以对各种地理空间

要素作缓冲区分析,实现元数据的地名查询。

## 4 应用实例

本文以客运港选址为例,具体说明可操作元数据框架是如何应用的,从而证明其可行性。

假设规划部门想在长江边上新建一座客运港,规划部门必须先拿到相关的地图数据,然后作GIS叠加分析。客运港选址至少要遵循以下原则:选址地点必须在长江边;选址地点的土地利用类型必须是城镇居民地;选址地点必须是近50a来未发生过洪水的地带。

为了完成这项客运港选址的任务,首先,规划部门必须收集到长江流域的数字线划图(主要是水系)、土地利用类型专题图和洪水专题图。因为元数据的地图注册方式已经实现,所以规划部门可以通过地名查询找到长江流域的数据,再通过主题分级的组织方式直接在水系图、土地利用类型专题图和洪水专题图中搜索,就能找到满足要求的元数据,根据数据链接元数据所提供的信息就可以下载到所需的数据对象。然后,规划部门通过GIS组件元数据找到能执行“合并”操作的工具,根据它的数据输入需求元数据的内容确定水系数据、土地利用类型数据和洪水数据是否满足该工具的要求。如果满足,就执行叠加操作,生成一个新的数据对象,即为适合修建客运港的地点。新的数据对象的属性项包括叠加前的原始数据的全部属性项,即包括坐标系统、地图范围以及相关操作。最后,在新的数据对象的操作元数据中自动保存操作日志(见图2)。

## 5 结语

本文研究了传统的地理信息元数据标准和元数据管理框架,分析了现行机制在分布式的网络环境下存在的问题。然而,仍有一些问题有待解决:目前采用的查询机制仅仅是着眼于元数据内容文本中的关键字,而不是采用基于语义的查询机制,查询准确性上有所欠缺。因此,在分布式的地理信息服务中实现基于语义的查询是下一步研究工作所面临的问题。一种可行性解决方案就是引入“元数据本体论”或“元数据先验知识”的概念<sup>[4]</sup>。非专业用户要想从大量的查询结果中精确地找到所需要的地理空间数据,并保证数据的质量,就必须制定合理的查询机制,否则不是得不到,就是得到的查询结果过多。目前,有人研究

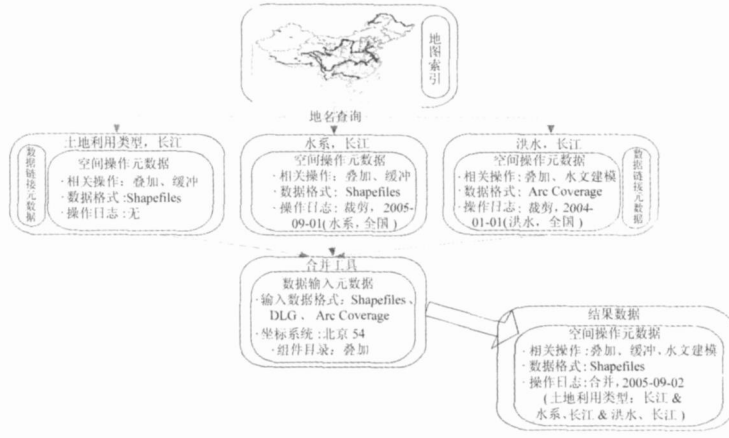


图 2 元数据框架应用实例

Fig. 2 An Example of Metadata Framework

信息可视化技术就是为了解决数据的直观选择问题<sup>[5]</sup>。此外, 本文只是从概念层面上给出了元数据的内容和管理组织结构的框架, 这还有待在实际的 GIS 中得到进一步的验证。

参 考 文 献

[ 1 ] MetaData Ad Hoc Working Group. Content Standard for Digital Geospatial Metadata Workbook[ S]. 2nd ed. Washington D C: Federal Geographic Data Committee, 2000

[ 2 ] ISO/ TC211 Secretariat. ISO/ T C211 Geographic Information/ Geomatics [ S]. Norway: Norwegian Technology Center, 2001

[ 3 ] Z39. 50 Maintenance Agency. Information Retrieval ( Z39. 50): Application Service Definition and Proto-

col Specification[ S]. Ray Denenberg: American National Standards Institute ( ANSI ), 1995

[ 4 ] Hubner S, Spittel R, Visse U, et al. Ontology-based Search for Interactive Digital Maps[ J]. IEEE Intelligent Systems, 2004, 19( 3): 80- 86

[ 5 ] Albertoni R, Bertone A, De Martino M. Visual Analysis of Geographic Metadata in a Spatial Data Infrastructure[ J]. Computer Society, 2004: 864- 865

[ 6 ] 王 浒, 李 琦, 承 继 成. 数字城市元数据服务体系的研究和实践[ J]. 北京大学学报( 自然科学版), 2004, 4( 1): 107- 117

第一作者简介: 梅琨, 博士生。现从事 GIS 研究。  
E-mail: kunjuly@gmail.com

Geospatial Metadata Framework in Distributed Network Environments

MEI Kun<sup>1</sup> BIAN Fuling<sup>1</sup>

(1 Research Center of Spatial Information & Digital Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

**Abstract:** After examining the existing metadata standards and their implementation frameworks, this paper presents an operational, object-oriented metadata content framework and a hierarchical, map-based metadata repository architecture as an alternative for searching, indexing, and retrieving distributed GIServices on the Internet, and gives an instance for implementation.

**Key words:** metadata; GIS; WMS; hierarchical metadata repository

About the first author: MEI Kun, Ph. D candidate, majors in GIS.  
E-mail: kunjuly@gmail.com