

# SIG 中的元数据共享服务体系研究与实践

张明刚

(上海交通大学 图像处理与模式识别研究所, 上海 200030)

[摘要]元数据是数字地球建设的关键技术之一,是建设信息化城市和实施空间信息网格基础设施的重要内容。本文基于 Web 服务对象及其标准,建立了分布、异构的网格环境下的元数据服务管理模型,包括基于 Web 服务的空间元数据共享服务体系、空间元数据服务与发布流程及其实现。最后,探讨了如何使用目前 JAVA 技术中比较流行的 JDBC 和懒加载技术开发空间元数据服务管理模型。

[关键字]空间信息网格; 空间元数据; web 服务; 元数据服务体系

[中图分类号]P208 [文献标识码] A [文章编号]1001-8379(2005)01-0007-05

## RESEARCHING AND IMPLEMENTING METADATA SHARING SERVICES ARCHITECTURE OF SPATIAL INFORMATION GRID

ZHANG Ming-gang

(Institute of Image Processing & Pattern Recognition, Shanghai Jiao Tong University, Shanghai 200030, China)

**Abstract:** Metadata is one of the key technologies of Digital Earth, which is also an important part in building the Information City and implementing the SIG infrastructure. Based on the researching in the standards and objects of web services, the authors came up with a service framework of metadata service and management in the distributed heterogeneous network environment. It includes the architecture of spatial metadata sharing service system on the basis of Web service, spatial metadata service and process of promulgation. In the end, we discuss how to use JDBC and lazy loaded, which are popular methods to develop the mode of spatial metadata service and management system and related to the JAVA technology.

**Key words:** Spatial Information Grid (SIG); Spatial Metadata; Web Services; Metadata Service Architecture

### 1 前言

随着城市信息化的不断发展,城市成为最重要的信息资源获取、处理和利用的载体,与此同时,和城市有关的各种空间信息资源也不断涌现,如市政规划、城市建设、城市管理、交通导航、生态环境数字化生存等,建立了众多的信息系统。但由于这些系统分散在不同部门、不同地点,而且基础数据不一致,信息资源的描述结构不规范,使得这些资源没有一种统一的信息发布和使用格式,无法共享,为了将这些空间信息资源充分利用,避免重复建设,空间信息网格(SIG)便应运而生。

空间信息网格(Spatial Information Grid, SIG)是一种基于 Grid 架构,以 Web Services 为基础,汇集和共享空间信息资源,进行一体化组织与处理,

具有按需服务能力的空间信息基础设施[5]。SIG 提供了一体化的空间信息获取、处理与应用服务的基本技术框架,以及智能化的空间信息处理平台和基本应用环境,为实现海量空间信息在线分析处理和服务提供了基础[5]。

由于空间信息资源具有分布性、异构性等特点,并由不同行业、不同单位与组织机构所拥有,因此要实现分布式环境下空间信息资源的汇集与共享,组建空间信息网格,方便海量空间数据管理,同时为用户提供方便的检索和共享使用,就要用到元数据共享服务。

### 2 元数据及其标准概述

#### 2.1 元数据有关概念

随着城市信息化进程的加快,地理空间数据生产者

和用户数量的增加,从空间数据海洋中快速获取目标数据,越来越依赖于元数据,利用适当的元数据文献来描述数据,使用空间元数据技术平台合理的管理空间元数据成为数据生产、存储、更新、和在线利用的必然趋势。通过元数据,用户可对空间信息数据库进行浏览、检索和分析等。

元数据(Metadata):通常来说,是描述数据的数据,或者说是用于提供某种类型资源的有关信息、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。FGDC[2]里“元数据是关于数据内容、质量、条件以及其他特征的数据”。

地理空间元数据(Geographical spatial Metadata):是地理空间相关数据和信息资源的描述信息,是对于地理数据特征的概括和抽取。它通过对地理空间数据的内容、质量、条件和其他特征的描述和说明,帮助人们有效的定位、评价、获取和使用地理信息。

地理空间元数据可提供空间数据集的特征资料,使数据用户可据此来确定数据的名称、来源、组织结构、适用范围等。特别是在基于 Web 的 GIS 解决方案中,地理空间元数据已经扩展到包括数据发现、数据转换、数据管理和数据使用的整个网络信息过程,是数字地球和空间信息网格基础设施的核心内容之一。

## 2.2 地理空间元数据的特点

在构建 SIG 的过程中,使用地理空间元数据标识海量数据,既有助于数据生产者对自己的产品进行管理、发布,又有助于数据用户更快捷、有效的查询、访问、获取和使用自己需要的、具有现实性强、精度高、便于访问等特点的地理空间海量数据。具体来说,地理空间元数据的主要功能有:1)是实现空间信息共享的前提和保障,是组织和管理空间信息、并挖掘空间信息资源的有效途径。通过它可以在 Internet 上准确的识别、定位和访问空间海量数据;2)可以帮助生产者和管理者有效组织、管理及维护空间数据,同时组织和维护一个机构对数据的投资,同时确保数据的一致性;3)是海量空间数据的索引,集成了大量零散数据的基本信息,用来建立空间海量数据信息的数据目录和数据交换中心,为空间数据的共享提供了可能;4)提供通过网络对数据进行查询检索的方法和途径,为用户处理

和转换数据提供必要的参考。如按照不同的地理区间,指定的语言以及具体的时间段,来查找空间信息资源。

## 2.3 元数据标准研究比较

空间信息元数据标准在构建空间信息网格及数字地球等项目中有重要意义。利用适当的元数据文献来描述数据,用空间元数据技术平台合理地管理空间元数据,已经成为数据生产、存储、更新和再利用的必然趋势。目前地理空间元数据已成为国际地理信息领域的研究热点,其主要的国内外标准化组织及其成果如表 1。

表 1 国内外几种主要的地理空间元数据标准

地理空间元数据标准	制定标准的组织或机构
CSDGM 数字地理空间元数据内容标准	美国联邦地理数据委员会 FGDC
CEN 地理信息-数据描述-元数据	CEN/TC287 欧洲标准化组织
ISO 地理信息元数据	国际标准化组织 ISO/TC211
中国可持续发展信息共享元数据标准	国家地理空间数据交换中心 NGDC

在本文所研究和讨论的过程中,拟采用了中国可持续发展信息共享元数据标准,作为地理空间元数据信息内容的表示标准,这一标准十分适合我国现阶段城市建设对规范空间元数据内容的实际需要,其结构如图 1,其中 M 表示此元数据元素集为必选,C 表示为可选。

## 3 元数据服务管理模型

### 3.1 WebGIS 到 SIG(GridGIS)的发展

当前,Grid 和 Web Services 技术的涌现,为 GIS 系统的发展提供了前所未有的机遇。Web 服务技术是为了解决当前在 Internet 环境下松散耦合的 Web 对象之间进行相互协作、相互集成而设计的技术框架,其本质是一种支持互操作的协议,以 XML/SOAP/WSDL/UDDI 为主干,Web 服务技术赋予了 Web 服务对象一个与传统对象调用技术相似但又不大相同的体系结构。

以前的 WebGIS 空间信息系统,基本上都是基于 CORBA、DCOM 等技术,以客户机/服务器为模式,协

议复杂，系统之间不能实现互操作，协作性差。因为 CORBA 和 DCOM 所采用的 ORPC(Object RPC)协议中，Obj Ref 代表了一个正在运行对象的引用，在 CORBA/IIOP(Internet Inter-Orb Protocol)中，用可交换互操作对象引用 IOR(Interoperable Object Reference)代表一个服务器的对象引用，但 IOR 和 Obj Ref 不能相互关联，可扩展性低；同时，CORBA 和 DCOM 表示了传送信息的语义，对参数和返回值，使用二进制编码，而对诸如参数名称或类型的任何

元信息都没有编码，使中介很难处理消息，而且各个系统使用不同的二进制编码，使得互操作更难以实现。现在，这些平台开发方式正逐渐被以 Web 对象为构件的 SIG 开发系统所取代，SIG 使用 Web 服务概念，综合运用 Java 强大的跨平台性，实现在垂直应用层面上 CORBA，DCOM 等技术功能的水平整合，能够更好的集成各个空间服务框架为一个体系松散、功能集合性强的整体，实现空间元数据信息的共享。

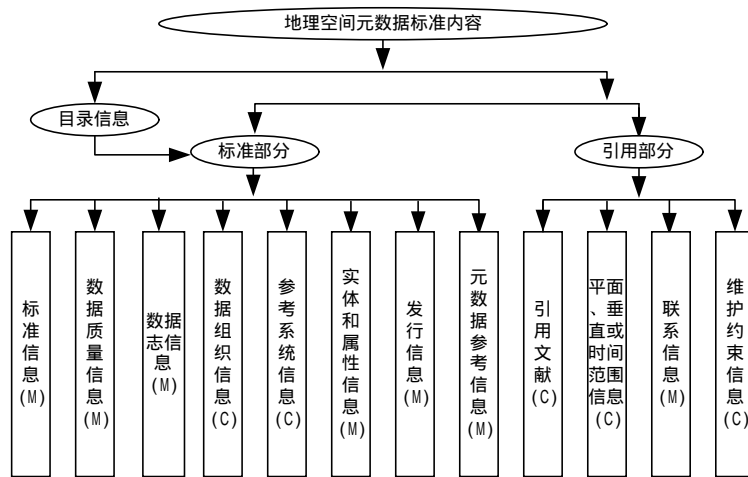


图 1 元数据标准内容图

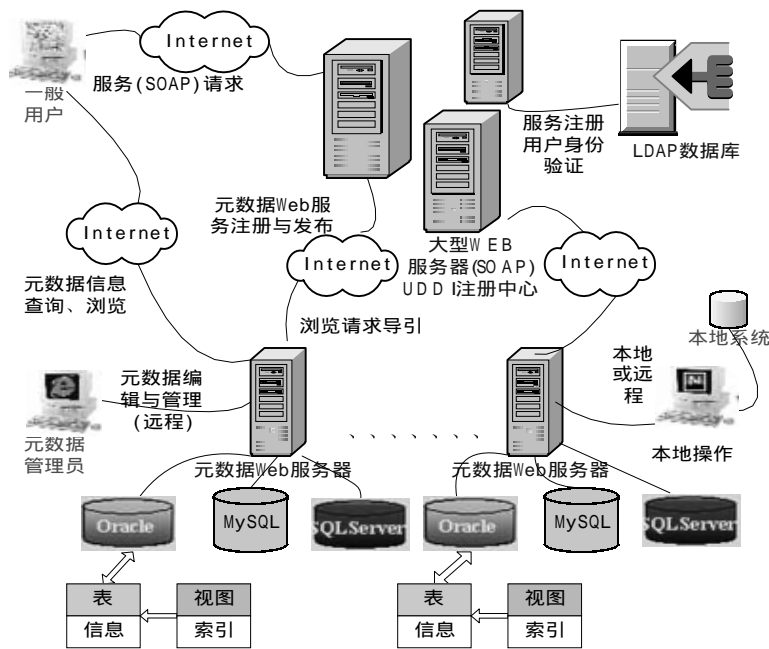


图 2 空间元数据共享服务体系

### 3.2 基于 Web 服务的空间元数据共享服务体系

根据 Web Services 的实现理论，SIG 的空间元

数据共享服务体系由 3 个角色部分组成，见图 2：元数据服务提供者、服务管理注册中心和元数据服务

请求者。元数据服务提供者是提供最终元数据 Web 服务的供应商, 即图 2 中的元数据 Web 服务器, 它实现了元数据共享的 Web 服务, 并放在在线服务器上供用户使用; 服务管理注册中心是一个 Web 服务的注册地, 汇集了众多服务提供者提供的 Web 服务, 供服务请求者查询利用, 同时对注册服务进行管理, 为服务注册者提供统一的模式标准, 即图 2 中的 UDDI 注册中心; 对一般用户即服务请求者来说, 通常, 他首先去查询注册中心, 尝试在注册中心找到所需要的 Web 服务, 当发现合适的服务之后, 将在注册中心获取这些 Web 服务的技术信息引用, 通过这些信息找到真正的 Web 服务及其相关的技术信息, 从而完成服务请求和服务提供者之间的绑定, 进行服务调用。

### 3.3 服务流程及实现

按照上述体系结构和 Grid Web Services 的基本理论, SIG 下的空间元数据共享服务系统以 JAVA 开发为基础, 采用分布式技术、中间件技术、Web 技

术实现, 各个对等的元数据 Web 服务器及元数据共享交换基于 Web 服务体系架构, 同时元数据的存储管理与发布采用标准的 XML 格式, 保证元数据在不同的操作系统和应用系统间能顺利的交互, 系统的跨平台性好。

#### 3.3.1 元数据服务建立与发布流程

对空间数据进行采集后, 遵从中国可持续发展信息共享元数据标准, 利用开发的元数据编辑管理器进行空间信息的元数据抽取, 以 XML 格式对元数据进行表示, 结合编辑器的树形表结构, 对元数据进行友好的管理和表示, 将元数据存储在不同的数据库中; 对元数据服务器, 采用 Java 中的 Servlet 和 JSP 技术, 遵循 EJB 架构, 用 tomcat + IIS 等服务器实现可扩展性和良好的访问性; 对建成的服务, 通过 SOAP 技术对服务信息进行封装, 注册发布到 UDDI (SOAP 服务器)注册中心, 同时, 注册中心根据用户的权限等级, 可授予服务注册用户不同的对服务管理的权限, 启动服务, 如图 3 所示。

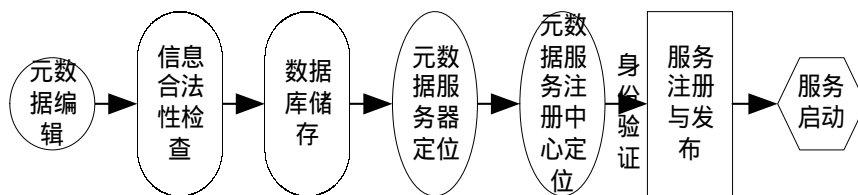


图 3 空间元数据服务建立与发布流程

元数据管理人员利用元数据编辑管理器, 可对信息进行添加、删除、修改、展开和层叠节点等操作, 如图 5 所示——元数据编辑管理器, 编辑管理器支持本地和远程数据库服务器的存取操作, 以 XML 格式表示数据形式, 黄色区域是对选中节点的信息说明, 各节点有三个信息项, 节点项、具体信息值、节点属性(条件/性质)。元数据经过采集整理后, 存储到元数据库中, 利用 Web 服务器发布系统提供服务的在线共享应用, 元数据管理员对服务系统有直接的管理权限, 是元数据服务的控制者。

#### 3.3.2 元数据服务流程

元数据服务启动后, 用户可以通过 Internet, 在服务发布中心即 UDDI 中心, 找到他所需要的元数据服务信息, 利用 SOAP 消息封装其请求, UDDI 中心直接将服务调用转到提供此元数据服务的元数据 Web 服务器, 绑定服务, 服务器解析其 SOAP 请求, 调用服务, 运用 Servlet 和 JavaBean 查询数据库信

息, 提交用户请求的元数据信息检索, 并转换基础 XML 格式为 HTML 格式, 通过 JSP 以 IE 或其他浏览器的形式显示在用户端, 其显示结果如图 6 所示——发布元数据浏览。结果遵从 XML 数据显示格式, 信息节点缩进, 便于用户浏览及参考节点之间的相关性, 流程如图 4。

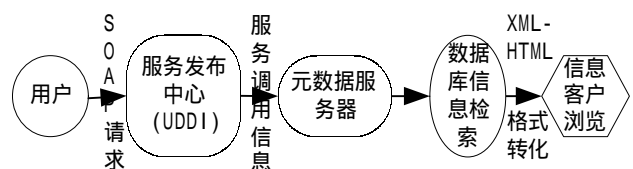


图 4 空间元数据服务流程

元数据服务为空间数据信息的有效利用提供了保障和方便, 用户可以从元数据信息中提取有用信息, 对所检索的空间信息进行精简提取, 节省了从海量空间数据中查询信息的时间, 并简化了操作。



图5 元数据编辑管理器

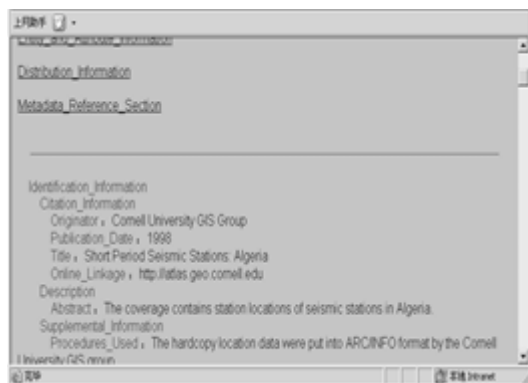


图6 发布元数据浏览

#### 4 空间元数据共享服务体系的相关技术探讨

在实现元数据共享服务体系过程中, 在服务器与数据库的交互上, 使用了目前 JAVA 技术中比较流行的 JDBC 方式。JDBC 是 SUN 公司为了简化 Java 应用程序与各个异构数据库的访问而开发的程序包, JDBC API 是数据库访问的理想接口, 在开发自己的 JDBC 应用的同时, 可以将其部署在与任意关系数据库通信的任何应用服务器的任何操作系统上。同时, 在数据库信息存取过程中, 执行了懒加载, 他首先从 JDBC 结果中搜索返回通过一次查询没有完全填充的对象, 然后, 根据需要在后台线程中, 各个对象再添加所缺少的数据, 这样, 不管需要多少数据库调用都不会影响返回搜索结果的速度。

在接下来的开发中, 可以运用负载均衡技术和服务器端高速缓存机制, 来减小服务器的负荷和扩展系统的性能。使用负载均衡技术, 可以使局部服务器超负荷运行状态减少到最小, 改善局部网络带宽拥塞现象出现的概率。使用服务器端高速缓存机制, 在服务器上通过键/值对查询, 缓存用户请求与

数据库信息, 减少了数据库操作, 提高了响应性能。

#### 5 结论

目前, Grid 与 Web Services 的兴起, 为空间信息基础设施的建设提供了良好的契机, 本文中基于 Web 服务的元数据共享服务体系的构建还处在一种探索和研究的阶段, 虽然完成了体系结构的关键技术的研究和实现, 但在项目的实际运用中还有很长的路要走。无疑, 元数据共享服务体系的构建, 将支持空间信息网格 SIG 及数字城市等项目的海量空间信息资源的管理、综合开发和利用, 对城市多元化及异构、分布的空间信息资源的整合、交换和共享有着重要作用。

#### 参考文献

- [1] 中国可持续发展信息共享元数据标准方案(修改稿)  
<http://nfgis.nsd.gov.cn/sdinfo/download.asp>
- [2] 陈惠荣, 游雄. 地理空间元数据及其相关技术的探讨. 测绘学院学报, 2003, 20(4).
- [3] 王浒, 李琦, 承继成. 数字城市元数据服务体系的研究和实践. 北京大学学报, 2004, 40(1).
- [4] 骆成凤, 吴国平, 余倩. 地理信息共享与 GIS 互操作的实现初探. 计算机应用研究, 2001, 18(8): 20-22.
- [5] 袁智德主编. 空间信息产业化现状与趋势. 北京: 科学出版社, 2004.
- [6] Wolfgang Hoeschek, Javier Jaen-Martinez, Asad Switzerland, Heinz Stockinger. Data Management in an International Data Grid Project. <http://www-igt.lbl.gov/~hoschek/publications/grid2000.pdf>
- [7] Savas Parastatidis. Working with Grid Services and Resource Metadata using Existing Specifications and Tools. [http://www.ics.forth.gr/isl/ist\\_workshop/presentations/Parastatidis.pdf](http://www.ics.forth.gr/isl/ist_workshop/presentations/Parastatidis.pdf)
- [8] Ozgur Balsoy, Ying Jin, Galip Aydin, Marlon Pierce, Geoffrey Fox. Automating Metadata Web Service Deployment for Problem Solving Environments. <http://grids.ucs.indiana.edu/ptliupages/publications/bal-soy-iccs03.pdf>

[收稿日期] 2004-03-29

[作者简介] 张明刚(1980—), 男, 四川广安人, 硕士研究生, 研究方向: 空间网格 J2EE 架构 Web 服务