

实现地理空间数据整合和更新方法的技术研究

赵俊三^{1,2}, 徐涛³, 赵耀龙², 傅晓东³, 李勇⁴

(1. 武汉大学 测绘学院, 湖北 武汉 430079; 2. 昆明理工大学 国土资源工程学院, 云南 昆明 650093;
3. 昆明理工大学 信息工程与自动化学院, 云南 昆明 650051; 4. 昆明金地科技有限公司, 云南 昆明 650106)

摘要: 各种地理空间型信息系统的建立均依赖于多层次、综合性的地理空间信息与属性信息集成的数据库, 这些地理空间数据具有多类型、多尺度、多 GIS 与数据库平台等特点. 为了信息系统开发建设的需要, 必须对这些分散的、异构的和不同格式的“信息孤岛”进行整合, 实现地理空间信息资源的一体化、标准化管理, 从而实现跨部门、跨地区、跨行业、跨应用系统之间的地理空间数据的交换、共享和协同处理. 论述地理空间信息数据整合的原则与方法, 结合具体应用分析信息共享及数据交换的技术路线. 同时, 通过实例说明地理空间数据整合与更新的实现方法.

关键词: 地理空间数据; 数据更新; Web Service; XML/GML

中图分类号: P208; TP309 **文献标识码:** A **文章编号:** 1007-855X(2005)03-0006-05

Methods of Integrating and Updating Geo - Spatial Databases

ZHAO Jun-san^{1,2}, XU Tao³, ZHAO Yao-long², FU Xiao-dong³, LI Yong⁴

(1. Surveying and Mapping College, Wuhan University, Wuhan 430079, China; 2. Faculty of Land Resource Engineering, Kunming University of Science and Technology, Kunming 650093, China; 3. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650051, China; 4. Kunming Yunjindi Geo - Information Co. Ltd., Kunming 650106, China)

Abstract: All the implementations of geo - spatial information systems are dependent on the multi - level, synthetic and integrated databases of spatial and attribute information. These databases have such unique characteristics as multiple classifications, multi - scales, multi - origin and different GIS formats. These separate Information Islands' must be integrated and updated to realize standardized management and data sharing. Principles and methods of integrating and updating geo - spatial databases are firstly discussed. The technique routines to share and exchange information are also analyzed through application. Finally, real examples are cited for elaboration.

Key words: geospatial data; updating data; Web service; XML/GML

0 引言

随着地理信息科学与技术的快速发展和广泛应用, 跨部门、跨地区、跨行业、跨应用系统之间的地理空间信息交换、共享与协同处理成为十分普遍的迫切要求^[5]. 按照信息系统建设的要求, 对多源、多尺度和不同 GIS 平台、不同数据库系统中存储的地理空间数据进行整合, 在统一的信息共享与交换平台支持下, 实现地理空间数据的集中管理和分布式应用, 成为近年来政府、企业和社会组织机构中各种 GIS 型信息系统开发建设所必须解决的关键技术问题^[3].

地理空间数据资源整合的主要目的包括: 首先, 通过数据整合, 协调各个组织机构或部门的关系, 使地理空间数据库能够满足各种应用系统开发建设的需要; 其次, 制定相关的数据政策、标准, 形成完善的地理空间信息共享、交换、数据更新的技术体系和服务环境; 再次, 通过多尺度、多时相、多源、多种类、动态化的

收稿日期: 2004 - 12 - 04. 基金项目: 云南省中青年学术和技术带头人引培项目(项目编号: 2000YP11).

第一作者简介: 赵俊三(1964 ~), 男, 教授. 主要研究方向: GIS/LIS. E - mail: jzhao@yjdgis.com

地理空间信息数据库的整合,为企业信息化建设和电子政务提供强大的地理空间信息支持;最后,提高社会对地理空间信息作用的认知度,推动地理空间信息基础设施的建设,形成良性循环,相互促进,使地理空间库的建设向实用化、规范化、产业化方向发展。

地理空间信息数据整合涉及到 GIS、数据库、计算机信息网络、地图处理等专业技术领域,需要解决主要问题包括:第一,对不同环境下分散存储的地理空间数据进行调查、分析、规范化和标准化处理,并进行信息的分类、抽取和逻辑集中;第二,利用 GIS 技术、网络技术、Web 技术、数据仓库技术、信息安全技术,对地理空间数据进行链接、结构优化、网络互联,建立新的面向应用系统开发建设的信息网络数据库或数据仓库体系;第三,建设统一的地理空间数据库管理机制,使数据种类方便添加、删除、修改,容易扩充和升级;第四,按照应用信息系统的实际开发需要,开发地理空间数据共享应用平台,构建新的应用系统,使地理空间信息在政府或企业的日常办公、内部管理、信息查询和决策支持等方面发挥更大的作用;第五,利用 GIS 空间分析,数据挖掘等技术,建立以地理空间信息为核心的模型预测系统、辅助决策支持等系统;第六,对可以公开的地理空间信息资源通过 Web GIS 等技术进行发布,面向社会公众提供综合信息咨询和信息服务。

1 数据整合的方法与技术路线

1.1 地理空间数据整合的基本原则

地理空间数据以矢量、栅格、多媒体、文本等多种方式存在,在进行多源空间数据整合时通过数据库引擎,把海量空间数据存放在关系数据库中,通过扩展磁盘阵列等方式发挥关系数据库对海量数据的存储和索引的功能,数据无需分块,作为一个整体存放在关系数据库中,对于异构数据利用数据交换实现,异构数据交换平台要建立在要对访问数据格式的充分理解的基础上。在进行地理空间数据整合时,需要制定符合要求的地理空间数据标准、规范与数据编码体系,或遵照国际 GIS 标准化组织制定的规范,如:OGC(Open-Gis Consortium)的 XML、GML3(GeoGraphy Markup Language),W3C 的 SVG 等^[2],通过构建数据共享平台,使地理数据和属性数据能够在不同的系统间自由交换。

1.2 利用数据库技术进行地理空间数据整合

通过利用新一代空间型数据库、中间件技术和数据库间的通讯技术,实现不同部门、组织机构间的异构地理空间信息资源的整合、调用和数据更新。譬如可以采用 Oracle Transparent Gateway 及 iAS 等技术实现对异构数据库的整合。

1.3 基于 Web Service 及其它相关技术的数据整合方法

目前可以采用 Web Service 技术,在 XML、GML 及 SVG 协议的支持下,实现不同机构和政府部门间的网络数据访问共享,其优点是在各组织机构拥有不同数据格式、不同数据源和 GIS 平台情况下,只需要按照 XML、GML 及 SVG 协议对空间数据进行包装,通过 Web Service 技术提供服务,其他用户就可以访问其空间数据;其缺点是,难以支持数据的适时双向交换和复杂的大型分布式 GIS 应用系统。通过分析和实验证明,可以在 XML、GML 协议支持下,采用 GIS 中间件和业务逻辑封装件等,开发具有支撑应用系统开发建设和数据交换等功能的地理空间数据运行引擎,实现对异构地理空间信息资源的整合,以及与专题应用数据库的无缝联接。

1.4 通过数据转换实现数据整合

可以应用 ODBC 数据接口实现与数据库的直接关联,导入数据。不同 GIS 平台所产生的不同数据格式的地理空间数据的转换、存储也可以采用同样的策略实现。

1.5 数据组织方法示例

在很多情况下,地理空间信息及其它相关业务信息分别被相互独立的系统管理和使用,形成了多个“信息孤岛”。为了实现数据的共享,达到综合应用的要求,这些不同的源数据应该进行统一规划、统一整合、统一管理。图 1 为某城市供气企业综合信息系统地理空间数据整合的技术流程。

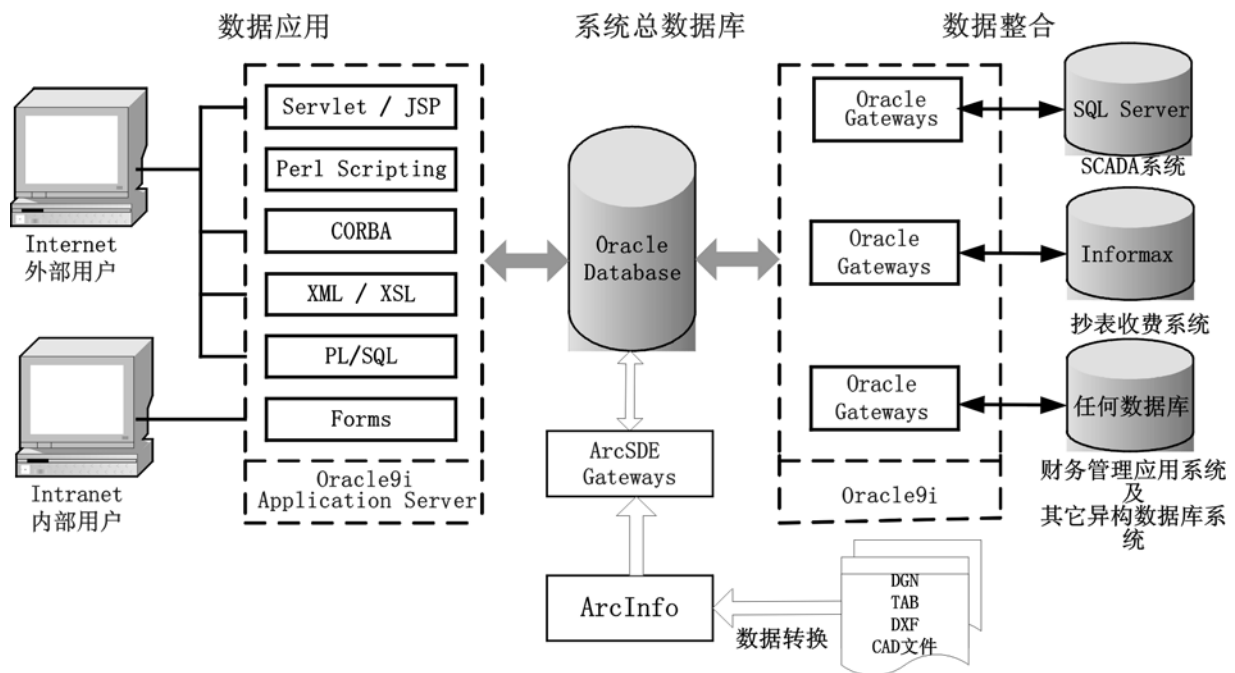


图1 数据整合系统的技术流程示例

Fig.1 Technological flow chart of the geospatial data integration

通过图1的例子可以看出:(1)在数据整合时,通过数据复制、提炼、数据抽取等方式获取这些独立系统的数据,对数据的采集将本着不影响现有系统运行、不增加现有系统负载、保证现有系统运行安全的原则进行;(2)利用数据传输技术、数据库访问技术、数据库管理技术、网络互联技术等实现数据的集中与管理;(3)在实现过程中可以采用商业化数据库管理平台软件,如 Oracle 等管理地理空间数据和属性数据、综合事务数据,异构数据库系统的数据可以通过中间件,如 Oracle Transparent GateWay 等进行数据通讯,并将数据存储到整合后的总数据库中,实现数据共享^[6]; (4)数据组织中,必须考虑元数据(metadata)及元数据库的建立^[7]. 元数据的主要功能是对各个异构的数据库进行各种必要的说明,起到数据库导航、数据文件的调用等数据管理的作用^[1].

2 信息共享技术分析

2.1 信息共享实现

基于地理空间信息资源的各种应用系统是数据库技术、中间件技术、MIS 技术、GIS 技术、网络技术及其它信息技术相集成的综合管理信息系统. 在系统的设计中应充分考虑现有系统的集成与未来发展的要求,软件系统的整合与集成应该建立在三层或多层体系结构理论及技术架构上.

系统的数据层包括集中管理的综合数据库、现有异构数据库地理空间数据库等. 综合数据库可采用 Oracle 等大型数据库管理系统支撑,实现对海量空间信息的管理、对综合数据的管理、对现有系统数据的存储管理. 完成支持中间层业务逻辑访问的服务层,包括数据的检索、数据传输接口、基于数据分析计算的决策支持等.

1) 要实现对地理空间信息资源的访问,需要首先将地理空间数据、属性数据进行格式转换并存储到数据库中,然后通过中间层的空间数据引擎实现访问. 通过分析、整理,选出有价值的数据库后,利用 GIS 平台所支持的 ODBC 接口直接把这些数据存储到 Oracle 数据库中,或利用接口程序、第三方转换工具(如 FME 等)、GIS 平台本身所包含的工具将其转换为通用的标准格式文件,最后通过空间数据引擎(SDE)导入数据库中.

2) 根据数据性质的不同,对各种现有数据库进行分析,存储到 Oracle 数据库系统,实现对信息的统一、

集中管理. 如在图 1 的例子中, 可通过中间层 ArcSDE、iAS 等提供的数据库访问接口实现数据的共享访问和对数据库系统的统一管理.

2.2 数据管理与访问

数据访问可以通过对 GIS 组件或 GIS 平台的二次开发实现. 其主要开发内容包括:

1) 空间数据引擎

建立空间数据引擎实现空间数据高效访问及与 GIS 平台无关的空间数据共享, 空间数据引擎包括以下功能:

- 空间数据存储

采用工业化标准的空间数据引擎进行海量空间数据管理. 如可采用 Oracle Spatial、Informix Spatial DataBlade、SDE 等进行数据组织与管理.

- 空间数据共享中间件

面向行业应用实现空间数据高效访问, 为行业应用提供全面的业务应用模型和管理机制, 实现与 GIS 平台无关的空间数据共享, 需要进行专门的开发.

- 空间数据共享 Web 服务层

利用 Web Service 技术, 基于空间数据共享中间件, 对外提供基本的与空间相关的 Web 服务和与行业相关的应用 Web 服务, 供不同的行业应用系统调用.

2) 应用系统运行引擎

依托应用系统建设, 开发针对地理空间数据访问的中间件, 业务逻辑封装件等, 通过与专题应用知识库的连接, 开发建立支撑应用系统运行引擎.

3) 决策分析引擎

决策分析引擎通过政策法规匹配、知识管理引擎、指标管理引擎、统计分析引擎、GIS 与决策分析系统集成等构建对项目审批、重大决策分析的平台.

数据协同管理应用内容包括: (1) 根据每类地理信息数据的更新机制要求, 利用“远程数据自动更新工具”定期从相关部门更新数据; (2) 根据需要, 从地理空间信息资源库中提取相关数据, 向相关应用部门提供地理信息服务; (3) 与系统应用单位进行日常数据更新和交换, 实现基础数据实时更新, 确保信息同步.

3 分布式地理空间数据更新方法

面向应用系统建立的地理空间信息数据库可分为基础地理数据库、专题业务数据库、综合信息数据库等内容^[4]. 数据库的逻辑层次结构一般划分为五级, 即: 总库——分库——子库——逻辑层——物理层. 首先根据数据的时态性, 可以将地理空间信息数据定义为工作数据库或临时数据库、现状数据库、历史数据库; 然后分析每一类数据库的作用, 地理空间数据整合和集成应用以现状库为基础, 按照满足现行系统运行需要、不破坏原有数据、提升新的应用的原则进行. 数据的更新则按照先专题, 后集中的方式进行. 图 2 说明了异构地理空间更新的技术路线和方法.

3.1 数据的定期更新

对于数据量较大、结构复杂的数据可以采用数据接口方式实现数据的更新^[8]. 对于变化较小的专项(专

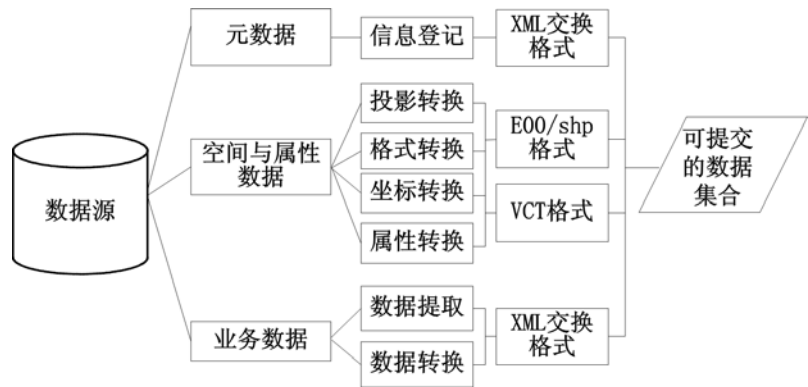


图2 异构地理空间数据的更新方法

Fig.2 The processes of updating the heretical geospatial data

题)数据可以在线编辑与更新方式实现,在数据更新的同时记录数据更新的日志与数据版本的维护。

3.2 数据更新过程

由于空间数据的复杂性,使得系统在集成不同源数据时无法通过通用的接口系统实现对所有数据的无缝连接.因此,数据集成与更新时必须根据不同的数据源开发相应的数据转换与更新接口程序,通过数据的转换形成可转入数据中心基础库的数据集合,并借助于计算机信息网络定期传输到信息中心更新入库。

3.2.1 接口结构设计与实现

需要更新到中心基础库的数据包括:地理空间数据、与之相关联的属性数据、描述空间数据的元数据、业务办理或统计结果等非空间数据等,需要采用不同的更新技术与方法。

元数据:根据元数据标准进行登记,形成 XML 交换格式。

地理空间与属性数据:通过转换程序依据数据交换标准,将属性转换、坐标转换、各式转换、投影转换等不同格式的空间数据转换为 E00/shp/VCT 等统一格式。

业务数据:通过数据提取、数据转换,导出为 XML 格式数据。

3.2.2 数据的上报

从具体的业务处理部门到中心的数据上报,可以通过 FTP 等方式上传到中心服务器存储区域. FTP 上传模块可以与数据转换与接口程序集成发放,上报时记录相应的数据集名称、更新日期、更新人员等信息。

3.2.3 数据更新

通过对更新信息的监控,获取上报待更新数据信息,如果按照规定时间段数据没有上报,则通过消息发送机制通知数据源管理部门提醒其上报.如果数据已经提交,可以根据数据集结构复杂程度,选择自动入库或是手工入库,进行数据的更新。

3.3 在线编辑与更新

对于空间数据编辑较少的情况,可以采用在线编辑与更新方式.授权对某一特定数据集或图层的可编辑,系统提供空间数据和属性数据的编辑功能,通过事务控制,在编辑后一次性提交数据,同时登记修改后的元数据信息并导入相应的业务数据。

无论通过数据接口还是在线编辑,在数据提交时都要进行记录数据变更的相关信息、保存元数据、变更数据版本等工作。

4 地理空间信息资源整合与更新的总体思路

根据地理空间信息资源整合与更新的特点和要求,其技术体系的核心思想是:通过由静态到动态的交换机制逐步实现数据的全面共享,地理空间数据资源中心建立或数据交换体系应用初期,以静态数据报送实现数据的整合和更新,此方式适用于应用系统建设初期数据交换量大的情况.当应用系统数据资源中心建立并运行一段时间后,数据交换量相对较小,可以通过网络采用动态交换的机制,实现真正意义上的信息共享和整合.整合后的地理空间数据的标准、数据格式、范围、坐标系等,可以通过对数据来源和新的应用系统的分析来确定.地理空间数据整合系统应具有数据质量检查、控制、安全认证、数据交换监视等功能.在进行数据整合和更新时,可通过商业数据库系统提供的数据集成中间件,开发的专用空间数据引擎实现下级部门或机构与数据交换中心的同步和异步数据复制;通过 XML 传输格式应用实现数据集的交换;通过消息机制实现数据的更新和实时发布。

5 结语

本文对地理空间信息资源整合与数据更新的方法及技术实现等内容进行了研究,主要目的是为地理空间信息资源的广泛应用提出一些探索性的解决方案.由于问题的复杂性,这里提出的方案只是初步的、概要性的,结合具体应用所提供的解决具有一定的参考性.随着 GIS 技术、数据库技术、软件开发技术和信息技术的发展,将为地理空间数据整合与更新提供更多的方法和实现途径,该领域的研究将会取得更加先进实用的成果。

(下转第 14 页)

3 总结

在实现两个不同基准下的坐标系坐标成果转换过程中,如果已知转换参数,将它们代入相应的转换模型,便可实现相互转换;不同参心基准下的坐标与 GPS 坐标如能进行合理转换,即借 GPS 技术这一中介,利用 GPS 在全球地心框架下的高精度测量成果,实际上便实现了各种坐标系中坐标成果间的转换.

在进行国家(地方)坐标系与 WGS-84 坐标系利用公共点进行相互转换时,通过上面介绍的数学模型及应用实例,可以发现:按已知点至它们所围区域的中心点距离平方倒数定初权,并用稳健估计理论进行数据处理,相对过去仅按各点等权的数据处理方法,求解出的转换参数性能会更优.另外,通过所得参数计算转化后已知点坐标的变化量,可以发现部分变化量较大的点,按已知点兼容性方法^[3]分析,可以发现它们正好属于内部符合精度较差的点.即利用该方法,亦可间接实现两网联合平差时起始基准点的选择,并赋予它们更合理的定权方法.当然,随着 GPS 技术的进一步发展,观测数据精度的进一步提高,一定会有更加完善的数据处理方法出现.

参考文献:

- [1] 孔祥元,郭际明,刘宗泉.大地测量学基础[M].武汉:武汉大学出版社,2001.171~173.
- [2] 刘大杰,施一民,过静珺.全球定位系统(GPS)的原理与数据处理[M].上海:同济大学出版社,2001.189~192.
- [3] 董世清,李志诚.GPS 测量控制点起算点的兼容性分析[J].四川测绘,2004,26(4):169~172.
- [4] 武汉测绘科技大学测量平差教研室.测量平差基础·第三版[M].北京:测绘出版社,1996.200~202.
- [5] 刘斌,王忠.面向对象程序设计 Visual C++[M].北京:清华大学出版社,2003.2~200.

(上接第 10 页)

参考文献:

- [1] 蒋景瞳,王启明,曾澜,等.中国地理信息元数据标准研究[M].北京:科学出版社,1999.
- [2] [美]Steven Holzner.XML 使用详解[M].北京:机械工业出版社,1999.
- [3] 张清浦,刘纪平.政府地理信息系统[M].北京:科学出版社,2003.
- [4] 国土资源部信息化领导小组办公室.全国国土资源政务管理信息系统与信息服务系统建设总体方案[J].国土资源信息化建设,2001,(2,3):2~13;21~28.
- [5] 赵俊三.GIS 发展的最新趋势及其应用前景[J].测绘工程,2000,9(2):21~25.
- [6] Ted Burroughs. Oracle 9i Advanced Replication, Release 2 (9.2) [M]. Redwood City: Oracle Corporation, 2002. 38~183.
- [7] Barbara Bicking, Russell East. Towards Dynamically Insetting Spatial Data And Its Metadata[A], First IEEE Metadata Conference[C]. NOAA Auditorium, Silver Spring, Maryland, 1996, April 16~18.
- [8] Junsan Zhao, Barry Wellar. Development of GIS/LIS for Cadastral Survey and Land Registration in China: A Status Report [J]. Geomatica, 1999,53(1):56~62.