

doi:10.3799/dqkx.2010.037

无缝空间数据中心概念模型及其技术体系

周顺平^{1,2}, 胡茂胜^{1,2}, 方芳¹

1. 中国地质大学信息工程学院, 湖北武汉 430074

2. 地理信息系统软件及其应用教育部工程中心, 湖北武汉 430074

摘要: 随着空间数据中心技术在国土资源等领域地逐渐成熟, 开拓出一套具有自身特色的空间数据集成管理与开发利用模式, 是一种新型的分布式地理信息系统(geographic information system, GIS)应用模型。然而, 空间数据集成与整合利用、增量更新问题以及不同厂家的功能共享与合作开发仍然是亟待解决的问题。在空间数据中心技术的基础上, 引入分布式异构空间数据无缝集成技术, 建立了数据无缝、功能无缝和应用无缝3个层次的无缝空间数据中心概念模型, 并讨论了其技术体系。

关键词: 空间数据中心; 无缝集成; 地理信息系统应用模型; 数据交换; 增量更新。

中图分类号: TP311

文章编号: 1000-2383(2010)03-0345-06

收稿日期: 2010-01-15

The Concept Model and Technologies of Seamless Spatial Data Center

ZHOU Shun-ping^{1,2}, HU Mao-sheng^{1,2}, FANG Fang¹

1. Faculty of Information Engineering, China University of Geosciences, Wuhan 430074, China

2. Engineering Research Center for GIS Software and Applications, Ministry of Education, Wuhan 430074, China

Abstract: As spatial data center technology is increasingly well-developed and highly sophisticated, the land resource management area is in bad need of a new model of distributed geographic information system (GIS) applications, which is featured by the integrated management and utilization of spatial data. In addition, much has to be worked out soon in the areas such spatial data integration and comprehensive use, incremental updates, as well as software functions sharing and cooperative development of different manufacturers. In this paper, based on the spatial data center technology, and by introducing the distributed heterogeneous spatial data seamless integration technology, a concept model called seamless spatial data center is established, and the technology is discussed.

Key words: spatial data center; seamless integration; geographic information system application model; data exchange; incremental update.

空间数据整合利用力度在不断增加, 内涵也在逐步深化, 涌现出众多具体实现形式。这些不同形式的空间数据整理利用方法分别试图从不同的角度来解决空间数据的共享与集成问题, 空间数据中心(徐世武等, 2006)和空间数据无缝集成(邬伦和张毅, 2002)是其中两种重要形式。空间数据中心模式是对空间数据资源的分布式集成管理模式, 而空间数据无缝集成是对空间数据的较高层次的整合利用, 这两种数据整合利用模式的结合具有重要的理论意义和应用价值。

从空间数据中心的角度看空间数据的无缝集成, 包含更加广泛的内容, 是一种集无缝化、集成、开发和应用为一体的系统观。空间数据中心作为一个分布式集成开发与应用模式包括3个方面(吴信才, 2009), 分别与无缝集成技术相结合, 得到空间数据的无缝集成、地理信息系统(geographic information system, GIS)功能的无缝集成和应用的无缝集成3个方面的研究内容, 而每个方面又包含若干具体的内容。

本文首先简单介绍空间数据中心的一般模式,

然后从数据中心的角度来研究实现空间数据无缝集成的方法与途径。

1 空间数据中心模式

空间数据处理环境可以分为空间数据库系统和 GIS 两个部分。空间数据库系统一般建立在数据库系统的相关技术之上,用于完成对海量空间信息的存储与管理,并提供有效的途径支持空间数据的查询和获取,还要提供用于保证数据安全及完整性的机制。而 GIS 主要关注于空间数据的采集、编辑、制图及分析。

随着空间数据处理环境不断向着分布式网络化的方向发展,空间数据库系统和 GIS 也逐步发展到新的阶段。空间数据库系统向着空间数据仓库的方向发展,以实现分布式异构空间数据的统一管理,它通过空间数据互操作技术、分布式协同操作与并发控制技术来实现数据在各个分布式节点上的同步、共享、互操作,并以透明的方式向外提供与空间数据库相兼容的访问方法。与此同时,GIS 也向着分布式协同计算的方向发展,GIS 分析处理功能转化为比较独立、复用性高、具有分布式协作能力的功能部件;这些功能部件又由统一的 GIS 功能仓库进行管理与维护,以实现对功能的版本控制、共享、发布、查询检索及搭建。这种新型的集成空间数据仓库与 GIS 功能仓库的空间数据分布式处理环境就被称为空间数据中心(吴信才,2009)。

方裕等(2001)提出了第四代 GIS 软件,指出第四代 GIS 软件应该由以系统为中心向以数据为中心发展,实现空间数据共享与服务的转变;空间数据中心正是在第四代 GIS 平台上的新的空间数据集成与 GIS 应用开发方式。

2 空间数据无缝集成相关技术

对空间数据进行集成及无缝化,包含十分广泛的内容,主要分为以下几个方面:

(1)实现数据无缝化组织。无缝 GIS(李爱光等,2005)、无缝空间数据库(朱欣焰等,2002)等都试图通过建立全新的无缝数据组织形式来管理空间数据。

(2)实现空间数据的无缝化拼接,实现几何的拼接和属性的合并是数据无缝集成最基本观点。

(3)地图合并技术。地图数据库合并或者地图合

并是指同一地区两个或两个以上不同来源的地图数据库生成一个新的更好的地图数据库(张桥平等,2001;李德仁等,2004)。其分为水平上的图幅拼接和垂直方向上的地图融合(崔铁军和郭黎,2007)。

(4)多尺度问题。多尺度问题从无缝集成的角度来看就是对同一地理区域上的多份不同比例尺的地图在显示、查询、分析上的合作,能够尽可能以用户透明且平滑的方式实现不同比例尺地图间的过渡、选择以及对比。

(5)多模式复合。交通往往是多样化的,出行者在城市之间、城市内可以选择多种交通方式,这就为出行方案的制定造成了困难。出行路线总是无缝的空间轨迹,从这个角度说实现数据间的无缝集成就是需要解决多模式复合交通网络的建立、分析、处理等问题(吴信才等,2008)。

(6)实现重复信息的冗余去除。人们不希望在一个 GIS 系统中关于相同的地物有多个不同的解释,系统应能屏蔽数据重叠区域的多义性,自动选择或组合成最合适的数据样本。

(7)数据模型间的相互转换。即使是分布式网络节点间实现了共享和互操作,但是数据依然可能以多种数据模型的形式被提供,造成了人们在分析处理这些数据时不得不进行显式的模型转换,不仅不方便还有可能破坏数据的拓扑结构或损失精度。因此在处理异构数据时,数据无缝集成的观点还包括数据模型间无缝的直接互操作。

然而,也有一些学者认为针对空间数据的无缝集成侧重于集成,而无缝是指无差别的对分布式异构空间数据源的访问方式。

3 无缝空间数据中心概念模型

空间数据中心提供对分布式异构空间数据实现集成利用的能力,是一种以数据仓库技术、功能仓库技术、搭建与配置技术为核心技术的新一代 GIS 架构技术和开发模式(吴信才,2009)。

通过对空间数据中心的 3 项核心技术进行无缝化改造,来建立无缝空间数据中心。因此无缝空间数据中心概念模型包括 3 个方面的内容:(1)对空间数据实现集成和无缝化的无缝空间数据仓库;(2)对异构 GIS 平台或应用系统间的功能实现集成与无缝化管理的无缝 GIS 功能仓库;(3)在无缝空间数据仓库及无缝 GIS 功能仓库的基础上的无缝应用建

模与系统搭建平台.

3.1 空间数据的无缝集成

无缝空间数据中心概念模型利用分布式版本机制实现空间数据的无缝化,无缝集成过程转化为无缝初始版本的建立过程以及数据更新过程中的无缝一致性维护过程(图 1).

利用分布式版本机制实现无缝化分布式异构空间数据的无缝化存储与管理,可以最大限度地保持数据本来的存储与管理形式,并通过一次性的无缝化预处理过程避免了每次查询过程中的动态无缝化过程;无缝查询过程转化为版本各分量的合并过程,计算复杂度低,保证了系统的响应效率;同时该机制还提供了对增量信息的存储与管理能力.

空间数据交换是空间数据中心提供的基础机制,在该机制的基础上通过增量识别等技术,可以实现增量信息的提取,以及实体间自动或半自动时序关系的建立;增量信息存储在无缝版本机制的增量数据库中.

管理的无缝集成,包括权限的无缝以及实现资源管理器式的数据无缝集成管理.分布式系统中往往在不同的节点上存储了大量的图幅和图层数据,人们也常常需要能迅速的从这些大量数据集中检索到自己关心的数据.如果这些数据由一个统一的数据库管理,那么就可以直接从该数据库的字典表中检索得到;但是在分布式系统中,用户不希望逐节点地利用查找单数据库的方法检索自己需要的数据.各节点的数据应该被全局无缝的管理起来,由对全局的数据字典的统一查询取代逐网络节点的手工检索.

3.2 GIS 功能的无缝集成

分布式空间数据的无缝集成本质上说是通过软件间的无缝集成来实现的,而软件无缝集成又有多种提法,比如异构系统之间的系统集成,多种同类软件(如多个厂家的 GIS 软件)间的软件集成,系统、功能或服务间的协作等.实现功能无缝集成的过程,

就是不断解决系统之间矛盾和冲突的过程.

互操作的无缝集成体现功能的无缝集成,解决来自多个系统平台的功能在同一体系下融合的问题.组件技术在各大 GIS 平台上的普遍支持,Web Service 技术的日益成熟,可视化建模技术的逐渐兴起,都为 GIS 间功能的互访提供了可能.

中间件技术是在互操作标准尚未成熟和普遍支持情况下的重要方法,一方面可以保证对原有系统的变更最小,另一方面还可以保证系统有较高的响应性能和吞吐能力.

3.3 应用的无缝集成

应用的无缝集成是面向 GIS 应用系统开发过程和用户应用过程而言的.空间数据和 GIS 功能,以及来自其他类型应用系统(例如 MIS)的数据及功能需要灵活的组织在一起,并能根据需求的变化适时地做出调整.当前的开发模式还主要是通过程序开发、软件升级等传统方法来组织、实现、更新、调整这类应用系统,缺乏用户的参与,软件周期漫长,无缝集成程度低;数据中心模式通过可视化建模及系统级搭建机制来解决上述问题,用户得以实时参与系统的建设过程,系统的各种数据资源及功能资源能够通过各种视图明确且形式一致地呈现在用户面前,体现了无缝集成的资源无差别管理及透明访问的思想.

应用的无缝集成的另一个重要方面是多个已建立的 GIS 应用系统间的协同工作,这些系统之间的语义映射是实现该要求的基础.然而在基于相同的开发框架,甚至开发框架提供了在其上建立的应用系统之间的互操作机制的情况下,系统间的协同工作将变得更加容易;空间数据中心模式一个重要方面提供了这样的开发框架.

4 无缝空间数据中心的技术体系

4.1 多层次智能中间件技术

中间件采用一致且对称的设计原则,即不同层次的中间件对上层来说提供的访问接口是相同的,并且能够通过标准的中间件访问接口调用其他的中间件.在这样的设计下,每一个中间件可以看成一种有输入和输出的 agent,每个 agent 有特定功能,这些 agents 从下层中间件的“动作”中进行“感知”,形成多 agents 协同工作的处理链.作为能够进行扩展、组合、协作的对称性中间件体系,可以说它的表

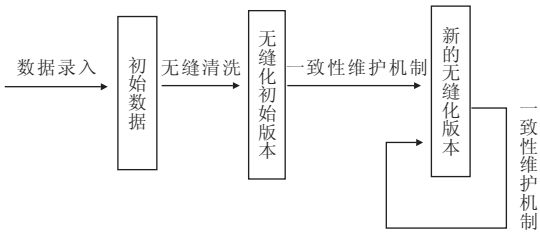


图 1 无缝版本状态演化与维护

Fig. 1 Seamless version states evolvement and maintenance

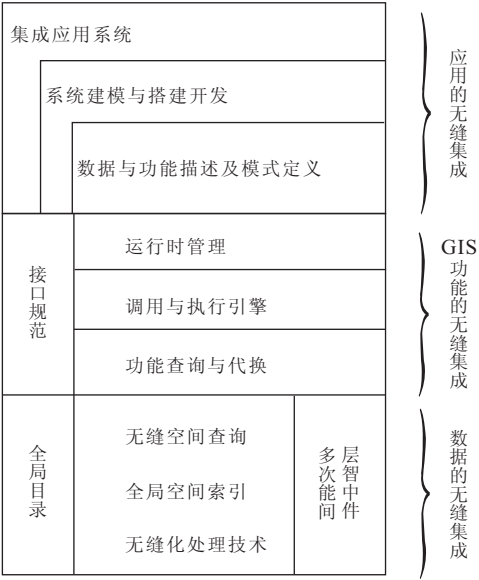


图 2 无缝数据中心技术体系

Fig. 2 Technologies of seamless spatial data center

现是智能的(图 2)。

建立分布式无缝版本机制需要解决数据的互操作、参照系的动态变换、分布式版本机制的实现、无缝查询的分解及查询结果的合成、多尺度的过渡与选择等不同层次的处理过程;多层次智能中间件在每一个层次上提供相同的语义,但在层次内部可以进行不同的处理操作,按照对数据流的处理形式可以把多层次智能中间件中的中间件分为 3 种类型:源类型中间件、功能类型中间件和聚合类型中间件。源类型的中间件对应 GIS 空间数据互操作中间件,这类中间件直接面对分布式异构数据,其他的中间件不能作为它们的输入;功能类型中间件提供对数据特定目标的附加处理,即从其他中间件返回的数据经过该中间件的处理后,形成新的数据并返回到上层;聚合型中间件从多个下层中间件实例中接收数据,施加某种融合操作,形成新的数据并返回到上层。

另一方面,对空间数据的访问和操作可以通过不同层次的抽象模型来完成,多层次智能中间件技术还应解决功能服务上的多层次的互操作,各种数据模型以及处于不同层次的功能服务都可以通过多层次智能中间件接入到系统中。

4.2 开放的空间数据无缝化处理技术

无缝化预处理过程是一个开放的过程,因为根据不同的环境、应用、用户等因素来说无缝空间数据可以有完全不同的定义,例如从图幅拼接到图层融

合、从属性的映射到时序的建立等。分布式版本机制提供了一致的组织和存储无缝空间数据的方法,但采用何种标准和方法生成无缝空间数据原则上来说是开放的。当然无缝空间数据中心做为完整的技术体系应提供一般性的方法以及扩展、重载机制。

4.3 管理集成与应用的集成开发能力

空间数据资源及 GIS 功能资源在空间数据中心中以分门别类的方式进行目录化管理,系统中全部资源都能以资源管理器的方式直观地得以呈现,实现了异构及异类资源的集成管理。

空间数据中心提供的集成开发技术是新一代的 GIS 构架技术与开发模式,其核心思想是以数据仓库技术、功能仓库技术、搭建与配置技术等为基础来解决分布式异构多源空间数据共享,以及不同厂家的功能共享和合作开发等 GIS 开发与应用领域的问题。

4.4 统一的无缝查询及无缝分析处理

对无缝空间数据来说首当其冲的要求就是统一的无缝查询及无缝分析处理能力;在无缝查询的基础上,能够自然地解决无缝分析处理问题;而实现无缝查询有赖于无缝版本机制,并且还要考虑效率的问题。

从技术上说上述问题的解决需要全局目录、全局空间索引以及多层次智能中间件等技术的结合,即通过全局目录和全局空间索引定义、定位资源,进而生成分布式查询任务,最后交由多层次智能中间件来执行并生成无缝查询结果集。

5 无缝空间数据中心的 应用模式

即使基于同样的无缝空间数据中心集成开发技术,在应用上各个 GIS 行业应用系统仍然允许有很大区别,但是这些系统应该遵守空间数据中心所确定的一般使用步骤:

(1)数据模式及应用模式描述与定义。数据模式与应用模式是对空间数据资源与 GIS 功能的资源的描述与建模,通过对已有资源的详细定义,实现并建立整个应用的基础。

(2)空间数据的录入与清洗。数据模式定义了数据的分布特征、命名规范以及元数据信息,通过这些信息可以实现对数据资源的枚举及检查,建立空间数据的元数据库,并由全局目录实现分门别类的管理。数据模式还定义了数据的有效性规则,对空间、

属性、关系等信息进行约束,排查不遵守约束的数据,以自动或交互的方式进行数据校正。通过空间数据录入与清洗过程,分布在各个数据源的数据得到确认与规范化,但是数据的存储形式并不发生变化。

(3)空间数据无缝化预处理。经过录入及清洗的空间数据还必须通过无缝化,才能生成初始的无缝化版本。这一过程首先要定义无缝化规则,并实现对开放式无缝化过程的扩展,最后建立起无缝化流程及实施无缝化过程。

(4)无缝空间数据的集成管理。经过录入、清洗、以及无缝化过程的空间数据已经初步实现了无缝集成。但在集成管理上还需要建立起资源目录,实现资源的资源管理器式的集成管理;建立起全局索引,为无缝查询提供基础。

(5)GIS功能的录入与集成管理。GIS功能在数据中心中也被看成一种资源,同样存在定义描述、录入和集成管理问题。GIS功能经过录入就可以由GIS功能仓库实现查询、定位、重新分类等管理操作,还能由流程设计建立起大粒度的新功能。

(6)应用系统的搭建与集成开发。无缝集成本质上说是为应用服务的,也是对应应用系统的搭建及集成开发过程而言的。这一过程需要综合运用无缝空间数据中心提供的方法及技术,根据具体的实际应用,设计与开发功能以及设计功能与无缝化空间数据的作用过程。

6 总结与展望

空间数据中心技术方兴未艾,仍然存在许多理论与技术问题亟待解决,这也是影响无缝空间数据中心深入发展的重要因素。建立无缝空间数据中心是一项复杂的系统工程,需要利用空间数据无缝化的理论与方法对空间数据中心技术实现全面改造,解决数据无缝集成、功能无缝集成以及系统无缝集成等不同方面、不同层次的问题。空间数据无缝集成必须结合具体的应用,利用无缝化与集成的一般方法,考虑应用中的实际语义,以期达到性能与效果上的最优优化。

随着空间数据中心技术的不断成熟,以及GIS各行业应用中对空间数据集成水平及高级整合利用需求的不断增长,无缝空间数据中心技术也必将得到发展与完善,与机遇同行的更是挑战。

References

- Cui, T. J., Guo, L., 2007. The study for multisource geospatial vector data integration and fusion. *Journal of Zhengzhou Institute of Surveying and Mapping*, 24(1): 1-4 (in Chinese with English abstract).
- Fang, Y., Zhou, C. H., Jing, G. F., et al., 2001. Research of 4th generation GIS software. *Journal of Image and Graphics*, 6(9): 817-823 (in Chinese with English abstract).
- Li, A. G., Wang, H., Guo, J., 2005. The spatial data organization of seamless GIS. *Engineering of Surveying and Mapping*, 14(1): 30-32 (in Chinese with English abstract).
- Li, D. R., Gong, J. Y., Zhang, Q. P., 2004. On the conflation of geographic databases. *Science of Surveying and Mapping*, 29(1): 1-4 (in Chinese with English abstract).
- Wu, L., Zhang, Y., 2002. The integrated framework on distributed multi-spatial database system. *Geography and Territorial Research*, 18(1): 6-10 (in Chinese with English abstract).
- Wu, X. C., 2009. Data center integration development technology: the next generation GIS architecture and development model. *Earth Science—Journal of China University of Geosciences*, 34(3): 540-546 (in Chinese with English abstract).
- Wu, X. C., Yang, L., Zhou, S. P., et al., 2008. Multimodal supported composite transportation network model. *Geomatics and Information Science of Wuhan University*, 33(4): 341-346 (in Chinese with English abstract).
- Xu, S. W., Xie, Z., Huang, Z. C., 2006. Research and design of isomerism distributed multilevel spatial data center. *Earth Science—Journal of China University of Geosciences*, 31(5): 624-630 (in Chinese with English abstract).
- Zhang, Q. P., Li, D. R., Gong, J. Y., 2001. Map conflation. *Bulletin of Surveying and Mapping*, 7: 6-8 (in Chinese with English abstract).
- Zhu, X. Y., Zhang, J. C., Li, D. R., et al., 2002. Concepts, implementation and problems of the seamless spatial database. *Geomatics and Information Science of Wuhan University*, 27(4): 382-386 (in Chinese with English abstract).

附中文参考文献

- 崔铁军, 郭黎, 2007. 多源地理空间矢量数据集成与融合方法探讨. *测绘科学技术学报*, 24(1): 1-4.

方裕,周成虎,景贵飞,等,2001. 第四代 GIS 软件研究. 中国图象图形学报,6(9):817—823.

李爱光,王卉,郭健,2005. 无缝 GIS 的空间数据组织研究. 测绘工程,14(1):30—32.

李德仁,龚健雅,张桥平,2004. 论地图数据库合并技术. 测绘科学,29(1):1—4.

邬伦,张毅,2002. 分布式多空间数据库系统的集成技术. 地理学与国土研究,18(1):6—10.

吴信才,2009. 数据中心集成开发技术:新一代 GIS 架构技术与开发模式. 地球科学——中国地质大学学报,34(3):540—546.

吴信才,杨林,周顺平,等,2008. 支持多模式的复合交通网络模型研究. 武汉大学学报(信息科学版),33(4):341—346.

徐世武,谢忠,黄志超,2006. 分布式异构多级空间数据中心的研发与设计. 地球科学——中国地质大学学报,31(5):624—630.

张桥平,李德仁,龚健雅,2001. 地图合并技术. 测绘通报,7:6—8.

朱欣焰,张建超,李德仁,等,2002. 无缝空间数据库的概念、实现与问题研究. 武汉大学学报(信息科学版),27(4):382—386.

《地球科学——中国地质大学学报》

2010 年 第 35 卷 第 4 期 要目预告

云南罗平中三叠统一海龙类新材料	程 龙等
南海北部晚第四纪高分辨率浮游氧同位素地层学及其古气候信息	葛黄敏等
基于 EMD 分解的甘孜黄土磁化率记录的西南季风演化周期性	刘冬雁等
南海表层沉积物中细粒组分的稀土元素地球化学特征	刘建国等
大别山早白垩世基性脉岩 SHRIMP 锆石 U-Pb 定年、地球化学特征及成因	王世明等
浙东南怀溪铜金矿床与曹门碱性花岗岩体成因关系的年代学制约	李艳军等
川西石棉花岗岩的锆石 U-Pb 年龄和岩石地球化学特征:岩石成因与构造意义	林广春
河南洛宁沙沟 Ag-Pb-Zn 矿床银的赋存状态及成矿机理	李占轲等
南海天然气水合物分解的碳同位素证据	王淑红等
塔里木盆地中央隆起区现今地温场分布特征及其与油气的关系	冯昌格等
利用油包裹体微束荧光光谱判识油气充注期次	李纯泉等