

文章编号:1002-3682(2012)02-0047-07

面向海岛规划的空间数据仓库构建

张君珏, 丰爱平*, 苏天赞, 吴桑云

(国家海洋局 第一海洋研究所, 山东 青岛 266061)

摘 要:针对海岛规划的具体应用,在了解海岛规划所需信息以及规划重点内容的基础上,对海岛空间数据仓库体系结构进行设计,并以海岛植被分布为例,阐述了维度模型的构建过程,最后对其今后在海岛规划工作中的具体应用进行了展望。

关键词:空间数据仓库;维度模型;海岛规划

中图分类号: P208

文献标识码: A

海岛规划是加强海岛保护,引导海岛资源合理配置,最大限度地推动海岛经济的可持续发展的有效手段。海岛规划信息包括来自不同时期、不同区域、不同数据源、不同尺度的数据,数据量大,数据类型和结构复杂,利用传统的数据库管理难度大,且无法满足管理者和决策者进行多尺度、多角度的复杂查询和综合分析决策的需要。

空间数据仓库兴起于 20 世纪 90 年代,具有高度的数据集成以及对复杂数据分析、高层决策的能力,目前已广泛应用于多个领域。中国南海研究院、南京大学联合研究的“数字南海”系统,采用数据仓库集成管理技术,将南海的水深、地形、地质、地貌、资源、气象信息,以及历史、政治、经济等信息,以虚拟现实的可视化技术通过计算机综合地表现出来,并进行相应的空间分析,为南海资源开发、利用和保护等提供决策和管理所需要的各种空间数据支持^[1]。王永志等^[2]设计了具有数据源、空间 ETL、空间数据存储、基于 SOA 的应用服务和客户应用的 5 层地学空间数据仓库体系结构及物理部署,提出了实现地质领域真正意义上的数据集成与共享、一体化存储的解决方案。

本研究针对海岛规划的具体应用,引入空间数据仓库技术,在了解海岛规划所需信息以及规划重点内容的基础上,研究海岛空间数据仓库的构建技术,以期能为海岛规划和管理部门的分析决策提供有效的空间数据支撑。

1 空间数据仓库概述

Inmon 指出,“数据仓库是一个面向主题的、集成的、时变的、非易失的数据集合,支持

* 收稿日期:2011-11-10

资助项目:国家海洋公益性行业科研专项——基于生态系统的海岛保护与利用规划编制技术研究及应用示例(200905004)

作者简介:张君珏(1987-),女,硕士研究生,主要从事 3S 技术在海岛海岸带中的应用方面研究。

Email:zjj_1234wangyi.com@163.com

* 通讯作者:Email: fengap@fio.org.cn

管理决策制定”^[3]。空间数据仓库(SDW)是在数据仓库的基础上引入空间维数据,增加对空间数据的存储、管理和分析能力,根据决策主题从不同的地理信息系统(GIS)应用系统中截取从瞬态到区段直到全球系统的不同规模时空尺度上的信息,从而为地学研究以及有关环境资源政策的制定提供信息服务。

与空间数据库相比,二者之间的差别在于它们面向的应用不同。空间数据仓库是面向主题的,而传统的空间数据库是面向应用的。空间数据库(源数据库)负责原始数据的日常操作性应用,提供简单的空间查询和分析;空间数据仓库则根据主题通过专业模型对不同源数据库中的原始业务数据进行抽取和聚集,形成一个多维视角,为用户提供一个综合的、面向分析的决策支持环境^[4]。

空间数据仓库主要包括以下几个部分^[5]:

1)元数据。元数据是描述数据内容、质量、表示方式、空间参考、管理方式和其他特征的数据。空间数据仓库的元数据包括数据源的元数据、抽取转换元数据、维表元数据、数据图层元数据。

2)数据源。这是指分布在不同的地理信息系统或其他信息系统中的应用系统数据,它们存储在不同的平台和数据库系统中,为空间数据仓库提供数据源。

3)数据交换工具(ETL)。即是将不同平台和数据库中的源数据交换到空间数据仓库中的数据的工具,包括数据抽取、清洗、转换、加载的过程。

4)空间数据仓库。空间数据仓库用多维数据来组织、显示数据。主题维、空间维、时间维是空间数据仓库最基本的组织方式。

5)客户端分析工具。客户端分析工具主要采用旋转、切片、钻取和高维可视化分析技术,以多维视图的形式展现给用户,使用户能直观地理解、分析数据,进行决策支持。

2 海岛空间数据仓库构建

根据空间数据仓库的技术要求,海岛空间数据仓库构建过程如图1所示。

数据源层:数据源包括海岛综合数据库和海岛空间数据仓库。海岛综合数据库是海岛空间数据仓库系统的基础,包括不同海洋部门应用系统产生的空间数据、属性数据、文档报告等,类型复杂,数据量大,为海岛空间数据仓库提供数据源。海岛空间数据仓库利用一定的数据抽取策略从海岛数据库中抽取数据,经过数据清洗、转换、集成,存储在海岛空间数据仓库中。

驱动层:针对多种异构数据源,采用JDBC建立跨平台、网络数据访问接口,根据用户自定义格式将源数据抽取出来,转换为可扩充标记语言(XML)文件进行处理后,再导入空间数据仓库中,实现了不同网络、不同操作平台、不同数据库之间数据的抽取转换^[6]。ArcSDE是一个能在多种数据库管理系统(DBMS)平台上提供高级的、高性能的GIS数据管理的接口。它在标准的关系数据库之上增加一个空间数据管理层,采用空间索引的方式,可提供对空间、非空间数据的高效率操作的数据库服务。

应用层:应用层是针对海岛规划管理、辅助决策而建立的专业应用系统。它通过JDBC和ArcSDE访问海岛空间数据仓库,提供对海岛信息的访问、查询、发布等服务,并开展面向海岛需求的应用。

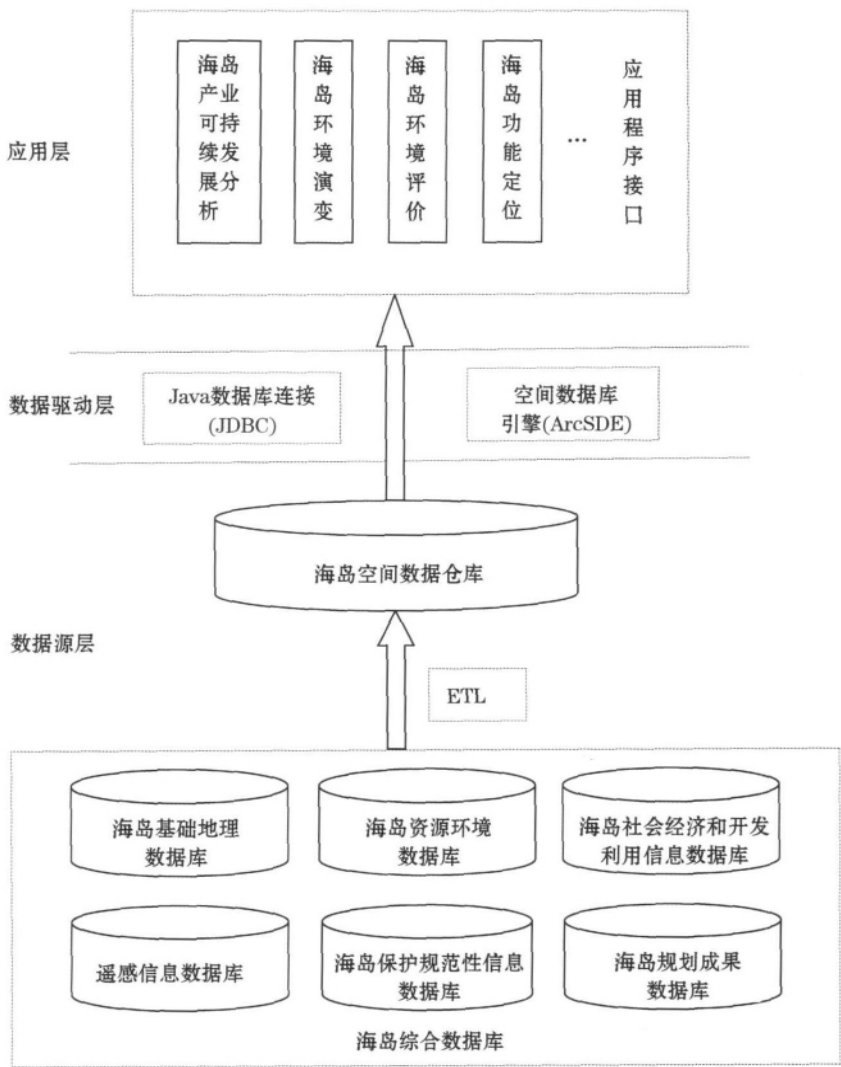


图 1 海岛空间数据仓库体系结构

Fig. 1 The architecture of island spatial data warehouse

2.1 数据来源

海岛规划工作的数据种类繁多,主要包括:

1)海岛调查数据

海岛调查数据包括针对海岛开展的地质、资源、环境、开发利用现状调查,具体包括海岛岸线、海岛地形、岸滩地貌与冲淤、潮间带底质、沉积化学、底栖生物、水文观测、水质、植被资源、土壤环境质量、土地利用情况等。

2)搜集的历史资料

搜集的资料包括海岛地理状况、自然状况、自然资源类型及开发利用现状、社会经济状况以及海岛规划成果五个方面的资料,具体包括不同历史时期的地形图、海图、气象气

候数据、地质资料、自然灾害风险、动植物资源、矿产资源、水产资源、社会经济数据、规划文档、音像资料等。

3) 遥感影像数据

包括不同时期的、不同分辨率的遥感影像。

2.2 海岛综合数据库

针对多源异构的海岛规划数据的要素类别,建立了海岛综合数据库,主要由以下6个分库构成:

1) 海岛基础地理数据库

主要包括地形图、海图、GPS观测数据、海岸线、行政区划图、地理交通图、基础地理设施等。

2) 海岛资源环境数据库

包括滩涂资源、植被资源、水资源、土地资源、矿产资源、生物资源、气象气候数据、地貌数据、空气质量数据、地表水质量、潮间带环境、海岛灾害情况等。

3) 海岛社会经济和开发利用信息数据库

包括人口、综合经济、文化教育、医疗卫生、产业布局、渔业布局、功能区划图、使用现状图等。

4) 遥感信息数据库

包括不同历史时期、不同分辨率、不同来源的遥感影像。

5) 海岛保护规范性信息数据库

包括国家和地方制定的各类海岛保护法律、法规、条例、办法以及各类相关行业标准。

6) 海岛规划成果数据库

包括根据不同专业需要制作的专题图以及音像资料和文档报告等。

2.3 海岛空间数据仓库设计

2.3.1 主题划分

海岛规划的重点内容包括海岛资源的开发利用和保护、海岛生态环境的保护、海岛生态系统健康、海岛功能定位、海岛产业布局规划等。按照海岛规划的应用需求,将海岛空间数据仓库划分为资源环境保护、土地利用规划和社会经济发展三个主题域,主题域下划分为多个子集(图2)。

2.3.2 多维数据模型

空间数据仓库一般采用多维数据模型存储数据,并有维索引及相应的元数据管理文件与数据相对应,避免了对海量空间数据数据的搜索。通过各维的层次比较、向上抽取、向下抽取等操作,可进行快速的复杂检索,优化查询。维度模型主要包括星型模型、雪花模型、事实星座模型等。本研究采用星型维度模型对海岛空间数据仓库进行模型设计。

星型模型一般由事实表、时间维、空间维、主题维组成。事实代表要进行分析的且我们感兴趣的模式或事件,通常包含变化较快的、数值型的事实数据和指向维表的外键。而维表包含数据仓库中相对静态的文本型、描述性的数据。维表是一种高层次的类型划分,是根据用户空间查询和分析需求来决定的。在实际中,维表常常具有多个属性,这些属性可按细节程度组织成层次形式^[7-9]。度量是指向空间对象的指针。这里,事实表以植被分

布为例进行论述(图 3)。度量采用植被覆盖率图和植被分布图。

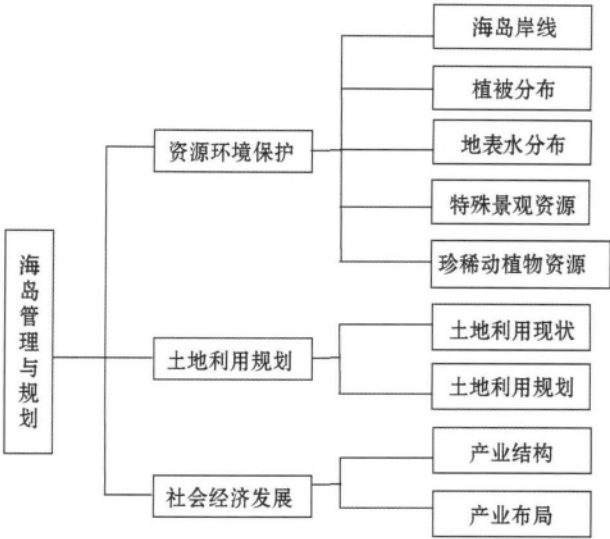


图 2 海岛空间数据仓库主题划分

Fig. 2 The theme division of island spatial data warehouse

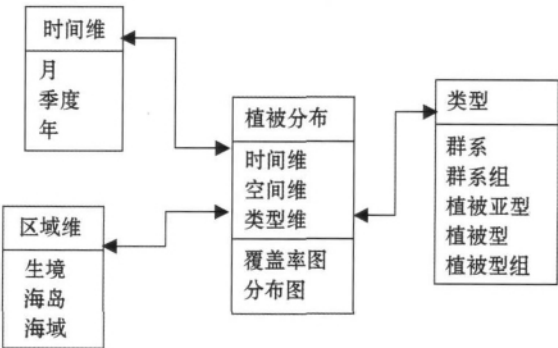


图 3 植被分布星型模型图

Fig. 3 The star model of vegetation distribution

- 1)时间维。根据海岛植被分布应用主题的需求,时间维采用月、季度、年三种时间粒度,以月作为最小时间粒度。
- 2)区域维。区域维包含生境、海岛和海域三种空间粒度。考虑到海岛生态系统的特殊性,海岛生境划分为环岛海域、岛滩、岛基、岛陆子系统。海岛为植被所在海岛名称。海域为渤海、黄海、东海、南海。
- 3)类型维。植被类型分类单位自小到大为亚群系、群系、群系组、植被亚型、植被型和植被型组六个等级。根据海岛植被实地调查情况,类型维采用群系、群系组、植被亚型、植被型、植被型组五种粒度,以群系作为最小粒度。

图 4 为以时间、区域、类型三个维度构建的多维数据集,也称数据立方体,是上述星型模型的物化视图,粒度分别为季度、海岛、群系,度量为植被分布图。基于这种数据立方

体,既可实现植被信息的简单查询,还可进行多维组合查询。比如,可以选择月、季度、年三种时间粒度的任意1个时间段,海岛生境、海岛、海域三种空间粒度,以及群系、群系组、植被亚型、植被型、植被型组五种类型粒度,查询和分析出3个维度任意组合条件下的植被覆盖率图或植被分布图。还可以快速进行统计分析以及趋势分析,如查询某海岛上的植被类型总数,对不同年份植被分布进行比对,分析其时空变化趋势等。

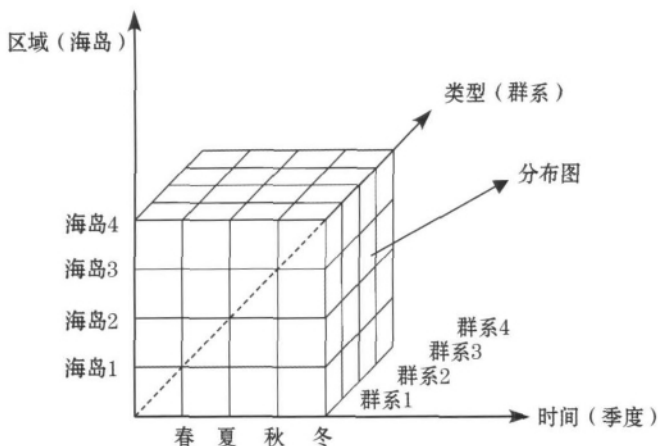


图4 数据立方体

Fig. 4 Data Cube

3 应用展望

在海岛空间数据仓库的基础上,利用数据挖掘算法,可以方便地进行海岛环境演变分析、环境评价分析等,辅助海岛规划决策。主要体现在以下几个方面:

1) 海岛产业可持续发展分析。在海岛空间数据仓库中,可利用空间关联分析算法,通过对海岛产业布局与海岛环境信息的综合分析得出海岛产业布局与海岛环境质量的相互影响与发展趋势,从而为海岛产业布局等提供指导。

2) 海岛环境演变。海岛空间数据仓库对数据库中的基本数据在一定的时间控制机制下进行提取,生成了多年的空间历史数据。利用分类模型分析这些历史数据,可得到有意义的分类模式,还可利用分类模型来进行趋势分析。如利用动态贝叶斯网络对长时间序列的海岛数据进行生态环境变化检测,分析其变化过程与发展规律,分析其驱动力,指导海岛的开发活动。

3) 海岛环境评价。如利用神经网络模型,根据主要的生态问题确定合适的生态因子,并划分敏感性等级,然后通过实验得到的样本,选择特定区域的专题数据,选取合适的网络结构,进行生态敏感性分析,为海岛规划中开发与保护布局提供指导。

4) 海岛功能定位。在海岛空间数据仓库中包含了由海岛综合数据库中的基本数据进行加工形成的专题层数据,通过建立的海岛功能定位多级特征指标体系,选择需要的各专题数据,利用规划分析中建立的数学模型以及各指标体系权重,计算综合指标并确定判定方法。确定海岛主导功能类型,以便更好地保护和利用海岛资源。

4 结 语

本研究引入空间数据仓库,结合海岛规划的应用需求,对海岛空间数据仓库进行了设计和构建,并且对其今后在海岛规划工作中的具体应用进行了展望。空间数据仓库建设的关键问题在于各应用主题的确定、如何根据应用主题的需求进行存储模型的设计以及如何合理高效地进行数据抽取。这需要根据海岛规划具体应用服务的需求进行反复修正,最终建立起满足用户查询、分析、决策等需求的空间数据仓库。空间数据仓库在海岛规划领域中具有广阔的应用前景,本研究今后将结合具体的应用主题,对其中的关键问题作系统、深入的研究与实践。

参考文献:

- [1] 贾培宏,朱大奎,马劲松,等. “数字南海”空间数据仓库构建技术研究[J]. 遥感信息,2007,(4):66-70.
- [2] 王永志,高光大,杨毅恒,等. 地学空间数据仓库的构建技术[J]. 地质通报,2008,27(5):713-718.
- [3] INMON W H. Building the data warehouse[M]. New York:Wiley Press,2005.
- [4] 赵霏生,杨崇俊. 空间数据仓库的技术与实践[J]. 遥感学报,2000,4(2):157-160.
- [5] 夏勇,王丽珍,胥玲芳,等. 澜沧江(湄公河)流域水资源空间数据仓库设计及实现[J]. 计算机工程与应用,2004,19:175-178.
- [6] 田扬戈,边馥苓. 空间数据仓库的 ETL 研究[J]. 武汉大学学报:信息科学版,2007,32(4):363-365.
- [7] 季民,靳奉祥,李婷,等. 海洋多维数据仓库构建研究[J]. 海洋学报,2009,31(6):48-53.
- [8] 纪翠玲,肖永康. 中国城市空气质量数据仓库的研究与实现[J]. 计算机工程,2009,35(17):235-237.
- [9] 阮惠华,韩立国,刘万崧. 空间数据仓库决策支持系统技术研究[J]. 吉林地质,2006,25(1):41-46.

Establishment of Spatial Data Warehouse for Island Planning

ZHANG Jun-jue , FENG Ai-ping , SU Tian-yun , WU Sang-yun

(First Institute of Oceanography , SOA, Qingdao 266061, China)

Abstract: For specific applications for the island planning, the architecture of island spatial data warehouse was designed on the basis of the understanding of the information needed for the island planning and the major planning contents, and the method for building dimensional model was described by taking island vegetation distribution as example. Finally, the applications of island spatial data warehouse to island planning activities were prospected.

Key words: spatial data warehouse; dimensional model; island planning