

基于GIS和数据仓库的政务决策支持系统

赵宝林^{1,2}, 夏 斌², 张园园^{1,2}

(1.中国科学院 研究生院, 北京 100049; 2.中国科学院 广州地球化学研究所, 广东 广州 510640)

摘 要: 分析了政务决策支持系统建设的必要性和建设目标, 对关键技术——数据仓库(Data Warehouse)、地理信息系统(GIS)和 Web Service 作了探讨, 提出了基于 GIS和数据仓库的政务决策支持系统的体系架构和设计思路。

关键词: GIS; 数据仓库; 电子政务; 决策支持系统; Web Service

中图分类号: TP317.1

文献标识码: A

文章编号: 1001-7348(2006)04-0148-02

0 前言

经过前些年大规模建设, 电子政务信息资源和社会公共基础信息资源都得到了很好的积累^[1]。但是绝大部分宝贵的信息资源都被封闭在彼此隔离的“孤岛式”信息系统中, 这些系统建立于各自独立的业务数据库系统之上, 长于政务工作的事务处理和政务信息的管理, 但对现有信息的综合分析和知识提取支持不够, 不能为政府管理人员提供必要的决策支持, 致使政务信息资源的社会价值无法得到充分发挥, 政府投入巨资建设的政务资源在重大问题决策过程中却不能

提供相应的决策依据, 电子政务建设走入了一个十分尴尬的境地。近些年来, 政务决策支持系统的研究已经开始得到社会的重视, 相关研究^[2]主要集中在以数据仓库、网络和计算机等技术为支撑, 整合现有的政务信息资源并提供相关主题的分析模型。但考虑到政府管理工作的特殊性(管人、管地、处理人地关系), 而人地关系和人类活动的绝大部分都与空间位置相关, 脱离了空间位置的人类活动是没有意义的。因此, 在政府决策过程中, 必须把特定的人类活动与相应的空间位置联系起来综合考虑。地理信息系统以基础地理空间数据为载体, 挂接各种社会、经

济、资源、环境、生态和人口等信息, 依据空间地理模型, 以文字、图形和多媒体等多种方式提供人类活动和社会发展的决策支持信息。

基于 GIS和数据仓库的政务决策支持系统主要是针对目前电子政务系统缺乏决策支持能力, 缺乏政务信息在地理空间分布方面的分析能力和分析结果在表达形式上的直观性与多样性支持不足等问题而提出的。它是充分挖掘政务资源, 为政府部门重大决策提供科学依据的强有力的工具。本文通过对政务决策系统的关键技术的探讨, 提出整体性的建设思路和系统的体系架构。

如果给定另一个聚类水平 $\lambda=0.992$, 得到:

$$R = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{pmatrix}$$

进行行列调整得到:

$$R = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

所以, 当 $\lambda=0.992$ 时, 得到的分类结果

为: $\{O_1, O_2, O_4, O_5\}, \{O_3\}$ 。在表 3 中我们可以看到, 对象 O_3 的第三项指标“团结协作精神”是所有对象中最高的。如果第三项指标权重重大, 对象 O_3 无疑是最合适的。根据我们的要求选定不同的 λ 值, 就可以得到不同的分类结果, 从而可以确定我们所需要的对象。

5 结束语

数据挖掘能够发现大量数据内部的规律, 结合模糊聚类分析方法, 能够对现有的数据进行聚类分析, 从而发现数据中的类型, 为决策者的决策和规划提供有用的信息。

参考文献:

- [1] 武林, 高学东, M·巴斯蒂安. 数据仓库与数据挖掘[M]. 北京: 冶金工业出版社, 2003.
- [2] 王文德, 巩建闻, 石凤来. 用 EXCEL 进行模糊聚类分析[J]. 聊城师范学院学报, 2000, (6).
- [3] 叶飞跃. 数据挖掘过程中的模糊聚类方法[J]. 计算机与现代化, 2003, (9).
- [4] 李良, 陈钢. 数据挖掘技术——模糊聚类分析在客户关系管理中的应用研究[J]. 工业控制计算机, 2003, (8).
- [5] 郝先臣, 刘小玲等. 模糊聚类挖掘方法在电子商务中的应用[J]. 东北大学学报, 2001, (4).

(责任编辑: 来 扬)

收稿日期: 2005-07-11

作者简介: 赵宝林(1978-), 男, 河北唐山人, 硕士研究生, 主要从事数字城市与3S集成应用方面的研究; 夏斌(1959-), 男, 研究员, 博士生导师, 主要从事数字城市与可持续发展方面研究。

1 总体目标

基于 GIS 和数据仓库的政务决策支持系统的总体目标是:充分利用现代高新信息技术整合电子政务信息资源,在统一的 Web 服务架构之上,以先进的地理信息系统为依托,搭载社会、经济、资源、环境、生态和人口等信息,借助数据仓库强大的数据整合分析能力,建立一个智能化、跨平台的政务决策支持系统,为政府部门在社会发展和建设中重大问题的决策提供科学依据,辅助政府工作,提高政府部门的管理和决策水平,促进社会经济持续、稳定、快速发展。同时为政府提供一个向社会公众发布决策的公共平台,提高政府决策的透明度,加强政府同社会公众的了解与沟通。

2 关键技术

2.1 地理信息系统技术

地理信息系统是一门新兴的交叉学科,是“数字地球”战略的关键技术之一^[3],是一种先进的决策支持系统,它以计算机和网络通讯等技术为支撑,以地理空间数据库(Geospatial Database)为基础,采用地理模型分析方法,适时提供多种空间的和动态的地理信息^[4],为与空间位置、分布相关的研究和应用提供服务。它把客观世界中的空间地理要素和地理现象抽象为模型化的空间数据,利用一个逻辑缩小的、高度信息化的信息系统,从视觉、计量和逻辑上对现实世界在功能方面进行模拟。信息的流动以及信息流动的结果,完全由计算机程序的运行和数据的变换来仿真。用户可以按应用目的观测这个现实世界模型各个方面的内容,取得自然过程的分析和预测的信息,用于管理和决策。

利用地理信息系统,决策人员可以轻松获得自然资源、人口、企事业单位的空间分布规律和空间给定指标的差异,还可以进行某一行业(如医院、社会救助与保障部门和企业、服务业等)空间布局规划决策分析。地理信息系统凭借自身强大的空间分析功能,可以为政府的规划、管理、决策等部门提供及时、明确、形象直观的空间地理信息服务,更好地辅助政府机关决策,提高政府的管理水平。

2.2 数据仓库技术

数据仓库技术是为了克服传统的数据仓库系统难于对决策进行支持的不足,为不同

层次的管理人员提供决策分析的工具,提高管理决策质量与效果。数据仓库之父 W.H. Inmon 把数据仓库定义为一个面向主题的、集成的、随时间变化的非易失性数据的集合,用于支持管理层的决策过程^[5]。数据仓库根据决策分析的要求,把现实世界中的业务和过程抽象为一系列完备且彼此独立的主题,作为问题域决策支持的主要分析对象。然后对分散于各处的源数据进行抽取、转换和综合等集成工作,把对分析对象进行了比较完整的、一致的数据描述的数据集合抽象出来,统一装入数据仓库中。最终由数据仓库前端工具建立决策支持模型^[6]对决策因素进行综合分析,得到需要的决策支持信息。

利用数据仓库技术,不但可以从较高的管理层次抽象和分析政务信息资源,综合管理和分析人口、社会和经济的现状、发展与历史信息,从而发现社会发展、经济建设以及城市化的内在规律,还可以对电子政务自身建设水平进行评估^[7]。

2.3 Web Service 技术

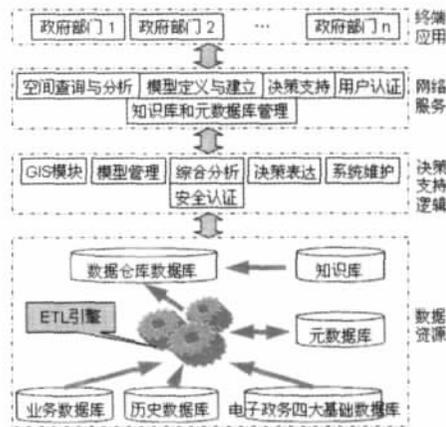
Web Service 技术是基于 XML(eXtensible Markup Language)与 SOAP(Simple Object Access Protocol)标准消息格式和标准通信协议 HTTP(Hyper Text Transfer Protocol),借助 WSDL(Web Service Description Language)进行自我功能描述,通过 UDDI(Universal Description Discovery and Integration)进行服务注册或发现,跨平台、可自由穿越防火墙,便于软件重用与应用集成、可伸缩性强的松散耦合的分布式计算环境。

基于 Web Service 技术,政府部门可以利用现有的政务信息资源和建成的决策支持系统快速定制出适合不同部门具体情况的决策支持系统,为实现电子政务 G2G^[8]提供理想解决方案,共享基础政务信息资源,重用决策分析功能,自由伸缩整个政府决策支持系统。

3 总体架构

基于 GIS 和数据仓库的政务决策支持系统采用松散耦合的 4 层 Web Service^[9]分布式架构(见附图),充分保证了系统的安全性和高度的可扩展性。以数据仓库数据库和包括 GIS 在内的决策支持相关功能组件为基础,对政府各部门提供统一的、直接面向高层管理的决策支持服务项目,保证各部门能够便捷、自由地整合这些服务功能和本部门

内部资源来满足其自身业务和决策需要。



附图 基于 GIS 与数据仓库的政务决策支持系统架构

3.1 数据资源层

这一层是整个系统的基础,而该层又是以各部门的业务数据库和已经存档的历史数据库以及电子政务的 4 大基础数据库^[10](自然人、法人、自然资源与空间地理和宏观经济基础数据库)为基础的。部门业务数据是那些从政府部门目前正在运行的业务处理系统里收集到的,保存在业务系统的数据存储中的数据。历史数据是指政府部门在长期的工作过程中积累下来的数据,这些数据一般经过业务处理系统脱机处理,以磁带或其它存储设备保存,对业务系统的当前运行不起作用。但是这些历史数据对利用数据仓库辅助决策的管理者来说就显得十分重要。电子政务 4 大基础数据库是经过对政府各个职能部门进行详尽的业务调研和数据需求分析之后,概括抽象出的部门公共数据集合,以统一的标准和格式集中存储到数据中心,旨在加强政务基础数据共享并提高政务资源的一致性和权威性。其中的自然资源与空间地理数据库是 GIS 功能组件进行空间分析的数据基础。

ETL(Extract Transfer & Load)是数据抽取、转换和装载的引擎,主要实现以下功能:数据标准化、数据的过滤与匹配、数据净化、标明数据时间戳和元数据的抽取与创建。数据标准化主要是将同名不同内容、同内容不同名和同名同内容但不同结构的数据进行标准化,例如,在不同的数据源里关于地点“北京市”的描述,有的用“北京”,有的用“北京市”,还有的用“京”,但他们的实际含义是一致的,因此需要统一处理以免混乱。数据的过滤与匹配主要是对进入数据仓库的数

据按应用需求进行筛选。数据净化是对准备加载到数据仓库的数据进行正确性判断,对有格式、类型或内容错误的数据进行修正。标明数据时间戳主要是给数据对象添加时间属性,为决策分析过程中基于时间维的聚合分析作准备。元数据的抽取与创建主要是将数据仓库的数据源和数据仓库自身的数据名称、定义和内容作详细描述,同时记录数据的抽取、转换规则,然后将上述描述和规则存入元数据库统一管理,为日后的数据仓库管理工作提供基础。

数据仓库数据库用来存储经过 ETL 引擎处理的、更高层次的、适用于决策支持的数据内容,供决策支持功能组件使用。决策知识库内存储了领域专家的决策经验和与决策相关的基本规则。

3.2 决策支持逻辑层

这一层是整个决策支持系统的功能核心,主要包括以下功能模块:GIS 模块、模型管理模块、综合分析及表达模块、系统维护管理模块、系统安全与用户认证模块。GIS 模块主要完成相关的空间查询、分析与表达;模型管理模块完成模型的定义、创建及编辑更新;综合分析及表达模块包含一组分析、决策支持与表达工具,同时借助 GIS 模块功能,依据特定模型进行决策分析,最终以图形、报表和多媒体等形象直观的方式将分析结果表达出来;系统维护管理模块完成决策知识库和元数据库的维护与更新,同时提供数据备份与恢复工具;系统安全与用户认证模块提供基于 LDAP (Lightweight Directory Access Protocol) 认证服务,在对用户统一认证的基础上,还提供禁止合法用户进行非法操作的功能,确保数据仓库的安全。

3.3 网络服务层

这一层提供跨平台的决策支持 WEB 服务接口,它把决策支持逻辑层的各个组件功能进行封装和整合,使得政府管理部门对决策支持系统的资源与功能可以进行完全透明的访问。借助 Web Service 的强大功能优势,政府部门完全无须考虑政务信息资源和服务提供者的位置,也无须考虑提供服务的系统和平台与本部门系统和平台的兼容性,只要通过 UDDI 中心获得相关服务的注册信息并通过认证,就可以按照 WSDL 的描述使用相应服务功能,甚至还可以把服务集成到本部门的专用系统中协同运作。

具体的 Web 服务项目基本与决策支持逻辑层的功能模块相对应,包括空间查询与分析服务、建模服务、决策支持服务、知识库和元数据库管理服务、用户认证服务等。其中,决策支持服务又包括数据的集成与分解、数据的概况与聚集,数据推算与推导、数据的翻译与格式化和数据的转换与映像。

3.4 终端应用层

这一层是系统价值的直接体现,直接发挥着决策支持的作用。政府部门根据自身管理决策的需要,结合本部门已有的资源,适当选取政务决策支持系统提供的服务,定制出适合本部门的决策支持应用系统,辅助日常的规划、管理和决策。

4 结语

电子政务决策支持系统的建设,是当前公共信息资源开发与应用方面的重点,是电子政务发展新阶段的必然要求,对提高政府部门的规划、管理和决策能力以及我国整体的信息化水平都具有重要的战略意义。基于

GIS和数据仓库的政务决策支持系统借助先进的空间信息技术、数据集成与挖掘技术和系统集成技术,将政府部门“管人、管地和处理人的关系”的日常管理、规划,决策等工作同工作对象的从属空间联系起来,大大提高了管理工作的直观性与准确度;同时将耗资建设的基础政务资源进行了合理的整合,发挥了其应有价值,为政府决策提供了必要的辅助支持。

参考文献:

- [1] 王长胜.中国电子政务发展报告[M].北京:社会科学文献出版社,2003.
- [2] 曹效阳,武立斌,蓝兹贵.基于数据仓库的政府决策支持模型[J].中山大学学报(自然科学版),2004,(6):97-101.
- [3] 李德仁.数字地球与“3S”技术[J].中国测绘,2003,(2):28-31.
- [4] 陈述彭,鲁学军,周成虎.地理信息系统导论[M].北京:科学出版社,2000.240.
- [5] W.H. Inmon. Building The Data Warehouse, 2nd Edition[M]. John Wiley and Sons,1996
- [6] 熊永柱,夏斌.海岸带可持续发展决策支持系统的关键技术探讨[J].科技进步与对策,2005,(3):46-48.
- [7] 陈冬林.电子政务的群体 AHP 评估模型[J].科技进步与对策,2005,(3):128-129.
- [8] 吴小寅,胡运发.上海市电子政务试点示范工程综述[J].计算机工程,2004,(5):42-44.
- [9] 张润莲,周德新,武小年.基于 4 层网络架构的语音电子邮件系统模型[J].计算机工程,2003,(5):114-115.
- [10] 张延松,薛永生.电子政务建设中的基础数据库建设规划研究[J].厦门大学学报(自然科学版),2004,(43):293-299.

(责任编辑:赵贤瑶)

GIS & Data Warehouse Based Decision Support System of E-government

Abstract: It's very hard for current e-government systems to provide decision supporting service to governments. After the analysis of the necessity and goals to establish the GIS & data warehouse based decision support system of e-government, the related key technologies such as Geographical Information System, Data Warehouse & Web Service are discussed. A solution is put forward through the detailed discussion about the architecture of the system.

Key words: geographical information system; e-government; data warehouse; decision support system; web service