

基于数据仓库的政府决策支持模型^{*}

曹效阳, 武立斌, 蓝兹贵

(中山大学信息科学与技术学院, 广东 广州 510275)

摘 要: 电子政务正在成为政府进行决策和管理的重要手段。基于数据库的决策支持系统无法有效辅助政府决策活动。提出的一个基于数据仓库的政府决策支持模型, 通过构建 3 层结构来实现原始数据到统一数据的转换、存储和开发, 进而实现辅助决策支持的功能。

关键词: 电子政务; 数据仓库; 决策支持系统

中图分类号: TP311.13 **文献标识码:** A **文章编号:** 0529-6579 (2004) 06-0097-05

决策是管理的核心。现代社会, 无处不存在决策行为。人们通过各种决策活动, 不断地提升自身的发展。决策支持系统 (decision supported system, DSS, 图 1) 是人们在信息技术支持下进行决策活动的重要工具, 在为人们提供各种决策信息和解决方案同时, 大大减轻了管理者从事低层次信息处理和负担, 使得他们专注于最需要决策智慧和经验的工作, 因此提高了决策的质量和效率。可以说, 决策支持系统的进一步发展, 将会给人们的决策活动带来深刻的变化。

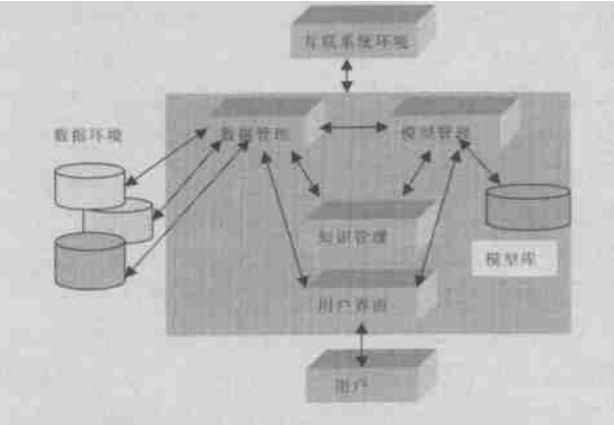


图 1 决策支持系统 (DSS) 基本结构图
Fig. 1 The framework of the DSS

1 基于 DB 的 DSS

DSS 是在 MIS (management information system) 发展基础上产生的^[1]。它通过结合个人的智力资源和计算机的能力来改进决策的质量, 是一个基于计算机的支持系统, 服务于处理半结构化问题的管理

决策制定者。因此一个好的 DSS 应具有如下功能: ①辅助决策者解决半结构化或非结构化的问题; ②允许用户试探几种不同的决策方案; ③必须具备决策支持模型的管理功能; ④把数学模型或分析技术与数据存储和检索功能结合起来; ⑤系统必须具备良好的适应能力, 可以满足不同环境和用户的需求; ⑥系统必须具备友好的人机交互界面。

然而, 由于 MIS 在数据挖掘、信息分析、预测等功能的先天性欠缺, DSS 的产生并没有体现它的应有作用, 反而暴露出种种问题, 应用性大大降低。这些问题体现在: ①基于 DB (database) 的 DSS, 只能对原始数据进行一般的加工和汇总, 不能有效地利用各种异构数据, 没有可靠的基于时间的较完整的数据存储, 从而大大降低了决策的可靠性; ②决策行为与决策模型的交互不足, 导致 DSS 的决策结果无法获得经验决策的认可; ③数据存储日渐海量, 大量有价值的信息隐藏在海量存储环境里, 而基于 DB 的决策分析无法有效地提取这些信息。

2 基于数据仓库的 DSS

数据仓库及其相关技术的兴起和发展使 DSS 重新进入人们的视野。数据仓库的主要倡导者是美国著名信息工程学家 William Inmon。他对数据仓库的解释是: 一个数据仓库通常是一个面向主题的、综合的、不同时间的、稳定的数据的集合。所谓主题, 是指用户使用数据仓库进行决策时所关心的重点方面, 如人事分析、贷款分析、地区产业选择、技术投资分析、异常情况分析等。所谓面向主题,

* 收稿日期: 2004-04-13

基金项目: 广东电信优惠管理系统数学模型开发资助项目

作者简介: 曹效阳 (1963 年生), 男, 讲师, E-mail: caoxxy@21cn.com

是指数据仓库中的信息不是直接从各个业务部门的处理系统中简单地提取出来，而是根据需求，经过系统的预处理即加工、汇总和整理而成，以确保数据仓库内的信息并不是现行数据，而是记录了从过去某一时间开始的直至目前各个阶段的信息。所谓信息的相对稳定，是指一旦数据被提取到数据仓库以后，一般都被长期保留^[2]。

数据仓库的兴起还伴随着一系列相关技术的研究和应用。这些技术主要有 OLAP（在线联机分析处理）和数据挖掘（datamining）。OLAP 是针对某个特定的主题进行联机数据访问、处理和分析，通过直观方式从多个维度、多种数据综合程度将系统的运营情况展现给用户，它的特征性技术有上卷（drill up）、下钻（drill down）、切片（slices）、切块（dices）、转轴（pivot）、钻过（drill across）、钻透（drill-through）等^[3]。

数据挖掘是指从存放在数据库、数据仓库或其他信息库中的大量数据中发现有趣的知识的过程。数据挖掘的特征性技术有：数据总结、数据泛化、数据分类和预测、数据聚类、数据关联规则等。

我们知道，决策所需的数据总是与一些维数（每一维代表对数据的一个特定的观察视角，如地区、时间等）和不同级别（如部门、单位、地区和国家）的统计和计算有关。以多维数据为核心的多维数据分析是决策的主要内容。但数据库在提供决策所需的多维数据上是力不从心的，因为它不具有多维特征。然而，要构建数据仓库，却必须以数据库为基础。MIS 的广泛应用已经积累大量数据，而且目前数据库的数量和规模还在迅速增加和扩大。建立数据仓库进行各种相关应用成为可能。数据仓库的多维特征可以满足 DSS 对数据的分析和利用要求，可以较好解决 DSS 在面向数据库的应用中所遇到数据粒度、数据分区、数据可信度、数据集成、数据的时间基准等问题，使 DSS 的应用更接近现实。目前较为可行的 DSS 开发方案是以数据仓库技术为基础，综合 OLAP 和数据挖掘的一整套解决方案。

3 电子政务的决策特征描述

电子政务是政府在国民经济和社会信息化的背景下，以提高政府办公效率，改善决策和投资环境为目标，将政府的信息发布、管理、服务、沟通功能向国际互联网上迁移的系统解决方案。同时也提供了结合政府管理流程再造，构建和优化政府内部管理系统、决策支持系统、办公自动化系统，为政府信息管理、服务水平的提高提供强大的技术和咨

询支持^[4]。

实施电子政务，意味着要将更多的政府信息资源信息化。相对于企业而言，政府一大特点是拥有海量的信息资源，而且这些资源往往都分散在不同部门、不同机构。对多年所积累的业务数据的整理是一项非常繁杂的工作，此外，由于各级政府部门实行信息化建设的规模和程度各不相同，各个数据库管理信息系统对数据的存储格式也不尽相同，如何整理和归并各种形式的数据，有效地为政府决策服务，是关系到国计民生的问题（图 2）。

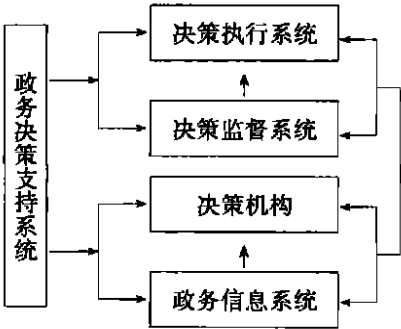


图 2 政务决策支持系统基本结构
Fig 2 The basic structure of DSS for E-government

电子政务要为政府决策提供有效支持。如何解决去除文山会海以后，不同部门和机构的信息交流、决策经验、专家意见等信息的有效获得和正确取舍就成了一个必须面对的问题。电子政务的决策支持功能就是要整合地区经济、科技、人文、社会、环境、港口、海洋等信息资源，加强和促进各部门之间的信息共享和交流，并通过对地区各种信息进行实时、动态的综合处理和分析，为政府制定城市整体发展战略、优化城市规划等全局性宏观决策提供科学依据，为领导决策提供服务。

电子政务决策支持系统应该具备足够的数据分析能力。电子政务的决策支持过程是一个从非结构化数据中抽取结构化信息，再提供非结构化决策分析结果的过程。在这个过程中最重要的是中间环节——抽取结构化信息，即数据分析环节。政府决策不能是一个纯粹的看数字下结论的过程，而是一个运用了多个合理的专家模型和经验模型的理性分析过程。为此，电子政务系统要能够为政府提供其应有功能首先就是要为决策者提供一个良好的决策环境，要为决策者提供进行科学决策所必须的工具，能够利用电子计算机进行信息收集、存储、加工和处理，模拟不同的决策方案，提供给决策者进行优化和选择。

政务决策支持过程是一个分布式过程。政务决策工作的特征是跨区域、跨行业，各个部门领导在

5 基于数据仓库的政府决策支持模型

如图 4 所示，系统整体结构分 3 层：原始数据层、数据仓库层和前端用户层^{[6 7]①}。

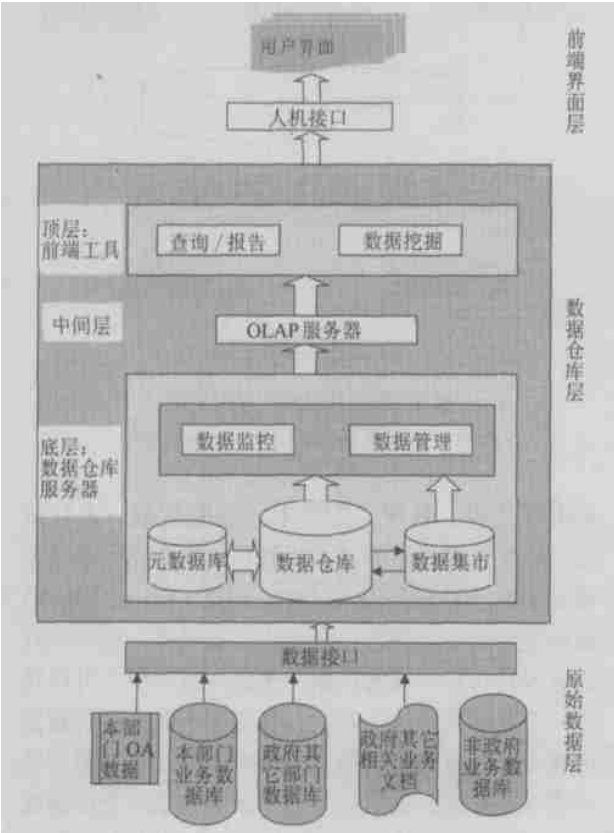


图 4 基于数据仓库的政府决策支持系统模型

Fig. 4 A DSS model based DW for the E-government

5.1 原始数据层

原始数据层主要是完成各种异构数据的处理，使其以合理的格式向数据仓库装载。这一层分为两个层次，即原始数据层和数据接口层。

原始数据层表明了政府决策支持系统的各种复杂的数据来源，主要有①政府本部门数据，包括政府本部门办公文档，本部门业务数据库，与部门的其它相关业务文档；②政府其它部门数据库；③非政府部门的外部数据库。这些数据通过数据接口层所提供的数据粒度规则，按照数据库高级复制、数据库远程访问、数据文件导入或手工录入等不同的方式进入到系统的数据仓库中。数据接口层主要为原始数据进入系统提供维护手段。主要功能包括：数据的提取、清洗、转换、集成、导入、刷新；内部业务子系统与本系统的接口。

原始数据层是一个物理分布结构，它基于一个地区或更大范围的应用，由分布在众多地理位置上的政府部门数据系统组成。在电子政务里，这些部门系统都具有独立运作业务数据的能力，即节点自

治性。但同时，各部门的职能数据系统也要通过网络进行全局应用。这种全局应用是在电子政府的统一控制与管理下，保证部门数据系统的自治性和全局范围内的透明性，并通过两阶段提交协议（2PC）保证系统中的数据一致性。各部门数据系统没有必要也不应该将数据全部对外开放，而应该按照统一规定将需要开放的数据通过协同服务平台和数据接口存入在数据仓库，辅助决策和提供其它部门使用。部门共享数据采用主副本存储机制。共享数据主本归发布单位所有，副本进入数据仓库。其它部门根据规定权限从数据仓库获得所需数据，并在本地存放该数据的另一个副本。由数据仓库保存部门共享数据的完整数据，各部门数据系统与数据仓库定期更新数据（图 5）。

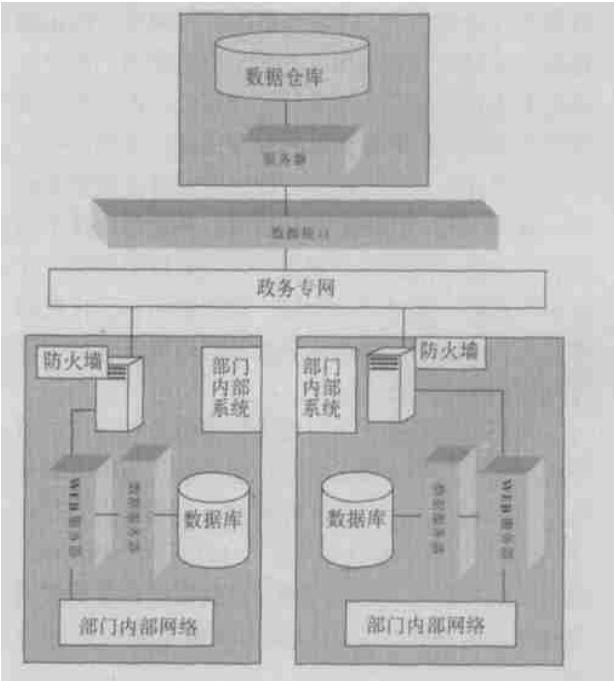


图 5 政务分布式数据系统结构图

Fig. 5 The structure of distributed data system for the E-government

5.2 数据仓库层

数据仓库层是政府决策支持系统的核心部份，由 3 个子层次构成。首先是底层的数据仓库服务器。数据仓库服务器统一管理着数据仓库、数据集市、元数据库以及数据监控和数据管理模块，将各种来源的数据按照系统设计的规范组织和管理起来，为后续的数据分析提供完整的数据源。数据仓

① The Working Group on E-government in the Developing World. Roadmap for E-Government in the Developing World > . Pacific Council on International Policy, 2002.

库主要存入经过清洗、转换和处理后的数据；元数据库主要是存放数据源、数据转换规则、数据维、事实表的定义等关于数据的数据，是维护、监控、管理数据仓库的根据。数据集市是根据数据仓库提供的数据，根据主题需要集成某一领域的数据，为联机分析和数据挖掘提供所需的所有模型元素；同时针对政府决策支持的业务需求建立起业务分析模型，将数据仓库提供的数据分析、组织形成各种适用于联机分析的模型元素。

其次是中间层。中间层主要联机分析处理(OLAP)。OLAP 服务器使用为用户预定义的多维数据视图对数据仓库（或数据集市）的信息进行统计分析处理，为具有明确分析范围和分析要求的用户提供高性能的决策支持。OLAP 将分析结果存储在信息库中，便于决策者通过对比多种分析结果作出更好的决策。此外，信息库中还存放决策准则、管理经验、常识。

最后是顶层。顶层包括查询/报告和数据挖掘。数据挖掘为辅助决策提供最完整的工具。与 OLAP 着重于对过去的已成事实的数据进行分类、分析、统计和集成相比较，数据挖掘注重对未来的可能发生的数据进行估算、预测，以发现新的关系。数据挖掘提供的功能有关联分析、分类和预测、聚类分析、孤立点分析、演变分析等。一些有代表性的预测模型有增长曲线模型、概率预测模型、确定性决策模型、风险决策模型、不确定性决策模型、回归模型、时间序列模型、层次分析决策模型等(图 6)。

5.3 前端界面层

前端界面层主要是人机接口和用户界面。人机对话接口是实现用户和系统之间的对话，通过对话以各种形式输入有关信息，包括数据，模型，公式，经验，判断等，通过推理和运算充分发挥决策者的智慧和创造力，充分利用系统提供的定量算法，做出正确的决策的。

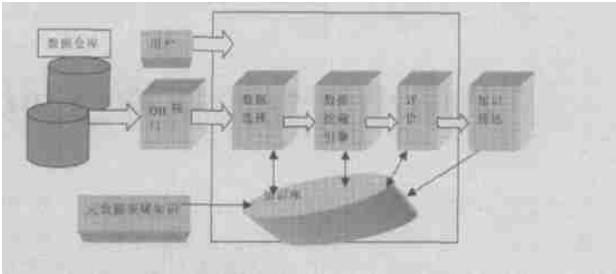


图 6 基于数据仓库的政务数据挖掘过程

Fig 6 The process of DM based DW for the E-government

用户界面支持用户按照不同的显示需要，形成不同内容和风格的个性化页面，不同的用户可以选择自己喜好的显示方式。提供了多种形式的模板，用于定制不同内容和风格的页面，可以快速的将用户关注的内容以用户最欣赏的方式提供给用户。不同的用户可以根据自己的需要定制不同内容及风格的显示页面。

参考文献:

[1] 国家信息安全工程技术研究中心, 国家信息安全基础设施研究中心. 电子政务决体技术设计与技术实现 [M] . 北京: 电子工业出版社, 2003.

[2] [美] INMON W H. 数据仓库[M] . 王志海, 林友芳, 等译. 北京: 机械工业出版社, 2003.

[3] [加] JIAWEI Han, MICHELINE K. 数据挖掘概念与技术 [M] . 范明, 孟小峰, 等译. 北京: 机械工业出版社, 2002.

[4] 苏新宁. 电子政务技术[M] . 北京: 国防工业出版社, 2003.

[5] 陈燕. 数据仓库技术及其应用[M] . 大连: 大连海事大学出版社. 2002.

[6] 刘海峰. 数据仓库在政府决策支持中的作用[EB/OL] . <http://219.239.110.3/articals1.php?id=503&gname=应用集成>《解决方案》数据仓库在政府决策支持中的作用. 2004.

[7] 泛微协同政务系统 enation 解决方案[EB/OL] . <http://www.grp.com.cn/fazx/fazx-view.asp?id=3300&product=&counter=1>. 2004.

A Decision Support Model Based on Data Warehouse for E-government in China

CAO Xiao yang, WU Li bin, LAN Zi gui

(School of Information Science and technology, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: E government is being an important tool for China government to make decision and management, but the DSS based on database does not support these actions efficiently. A decision support model based on data warehouse for E-government in China is designed to make the transformation, storage and exploitation by a three layer architecture, and then support decisions efficiently.

Key words: E government; data warehouse; DSS