

文章编号: 100F-1498(2002)06-0637-07

林业资源环境网络在线决策支持系统研究

张怀清, 鞠洪波, 陈永富

(中国林业科学研究院资源信息研究所, 北京 100091)

摘要: 为了通过 Internet 网络环境, 为决策者提供快速、及时、准确的在线数据分析和辅助决策, 本研究通过建立以部级水平的林业资源环境网络数据库, 并在此基础上设计和建立支持多维数据的数据仓库; 利用面向对象的思想, 设计和开发了模块化的方法库、模型库和知识库; 采用公共网关接口 (CGI) 技术, 进行用户交互界面、数据库访问和分析接口的设计, 建立在 Internet 上运行的林业资源环境网络在线决策支持系统。

关键词: 林业资源; 决策支持系统; 网络在线决策支持系统; 数据仓库

中图分类号: S757.2

文献标识码: A

决策支持系统 (Decision Support System, 简称 DSS) 产生于 20 世纪 70 年代, 是传统的管理信息系统理论上发展起来的一门适用于不同领域的、概念和技术全新的信息系统发展分支, 也是目前发展最为迅速的一个分支^[1, 2]。

20 世纪 80 年代以来, 决策支持系统广泛地应用于林业的信息系统中, 用来对森林资源预测和监测, 森林病虫害防治, 森林防火、生态保护等领域, 取得了丰硕的成果。这些系统大都是基于单机或基于局域网 (客户/服务器) 的结构, 以满足本单位、本部门的辅助决策需求。随着信息技术的发展, 基于 Internet 技术的网络在线决策支持系统得以迅速发展。本研究将决策支持系统通过 Internet 实现基于 Internet 的网络在线决策支持, 为不同的用户提供在线数据分析, 辅助决策者进行规划和决策。

1 系统结构设计

1.1 系统基本结构

林业资源环境网络决策系统采用基于数据仓库的决策支持系统, 在系统基本结构中与众不同的“三库”和“四库”基本相同。其系统基本结构由数据管理子系统、模型管理子系统、方法管理子系统、知识管理子系统和人机交互子系统 5 部分构成, 各部分之间互相协作, 共同完成辅助决策功能, 基本结构见图 1。

数据管理子系统主要负责管理林业资源和环境信息数据。在数据共享子系统的基础上提供多种数据查询, 以及提供数据的提取、净化、综合、组装并进行数据挖掘和在线分析。模型管理子系统负责各种决策分析方法的组合应用、模型的选择。方法管理子系统主要负责方法的选优和应用。知识管理子系统主要负责知识的综合、推理和应用。人机交互子系统主要负责

收稿日期: 2001-06-15

基金项目: 国家九·五攻关“部级资源与环境信息服务系统”部分内容 (96B020106)

作者简介: 张怀清 (1973), 男, 湖南宁乡人, 博士。

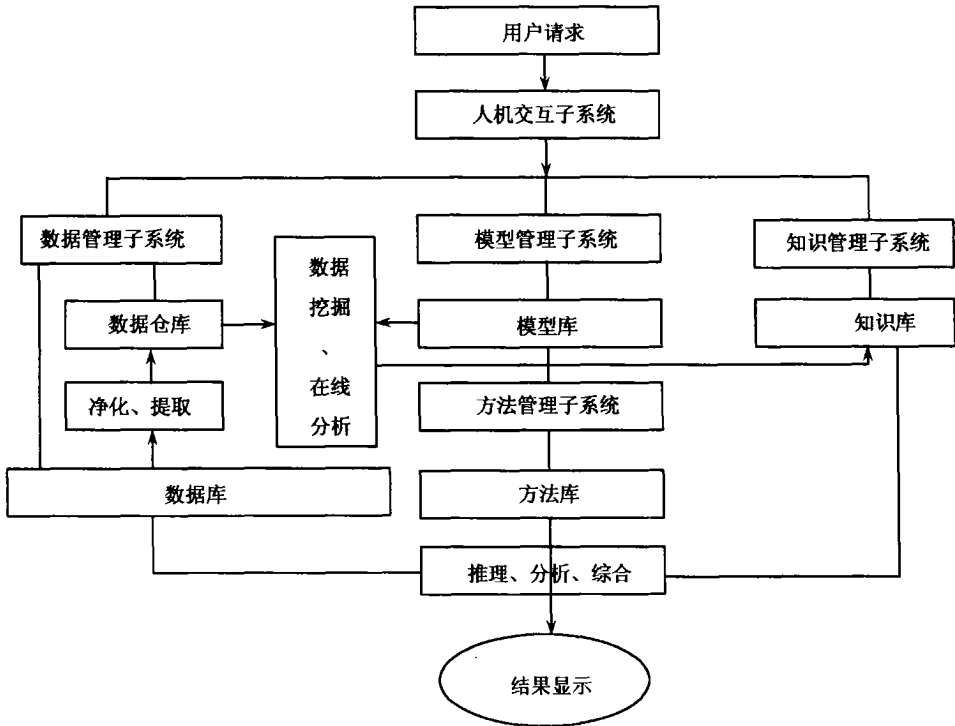


图 1 系统基本结构

系统的网络运行、用户和系统的交互及结果的输出和显示。

1.2 运行模式结构

林业资源环境网络决策支持系统运行采用基于 Internet 的 3 层结构模式,即(用户/WEB 服务器/数据库服务器)运行模式。

(1) 用户层: 用户的软硬件要求低, 只要求上网, 通过普通的浏览器即可与系统进行交互, 完成系统的查询、统计、分析和决策功能。

(2) 服务器层: 包括 WEB 服务器和应用服务器。用户交互的数据, 通过 WEB 服务器进行处理, 从数据库中提取用户所需要的数据, 应用服务器对数据进行必要的分析和计算, 最后经 WEB 服务器格式化显示于用户的浏览器端。其中模型、方法和知识库的分析、综合和应用发生在这一层。

(3) 数据服务层: 指用于分析的网络数据库、数据仓库, 在线分析一般采用分布式的网络数据库和数据仓库。

2 数据管理子系统

数据管理子系统以建立的网络数据库为基础, 并在此基础上根据决策的主题和决策目标建立数据仓库, 开展数据挖掘和在线分析, 数据管理子系统的结构如图 2。

(1) 网络数据库: 基于 Internet 的决策支持系统需要建立支持在线分析的网络数据库, 林业资源环境在线决策支持系统采用 MS SQL Server 2000 企业版作为数据库服务器平台, 它能支持

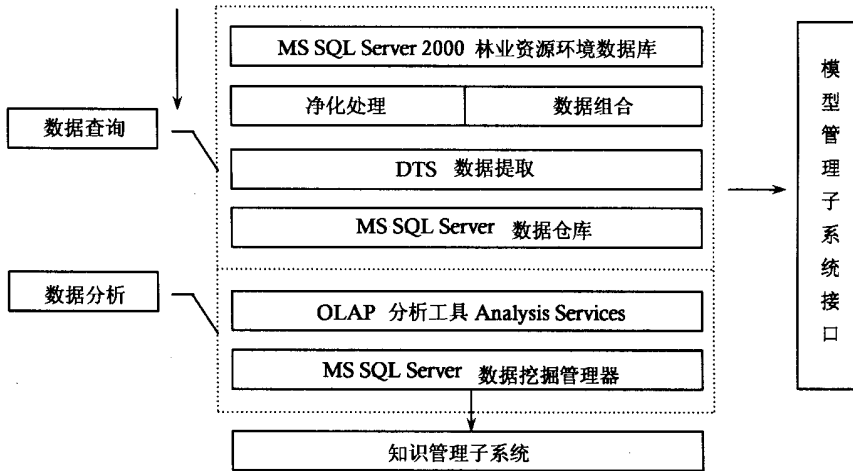


图2 数据管理子系统结构

林业资源环境数据的海量存储和多用户的并发快速访问。林业资源环境网络数据库的内容包括:历年森林资源状况、历年林业生产情况、历年林业社会情况、历年林业经济情况、历年林业工程建设情况、历年林业营林情况和林业自然资源等数据库。

(2) 数据仓库: 由于林业资源环境信息的数据量大而分散, 在进行决策分析时需要涉及大量的历史数据和半结构化的数据, 在传统的数据库管理系统基础上建立的决策支持系统只能提供辅助决策过程中的数据级支持, 难以求解复杂的半结构化的决策问题。面向决策的数据的集成化和规范化程度要求较高, 传统的数据库难以满足这种需求。同时, 面向决策的数据要求数据的综合和重组能力强, 数据灵活, 面向决策的主题。建立林业资源环境数据仓库能避免这些缺陷。数据仓库是面向主题的、集成的、非易失的且随时间变化的数据集, 用来支持管理人员的决策^[3]。数据仓库与传统的数据库相比较, 具有如下特点: 数据面向主题、集成和不可更新且随时间变化; 数据的多重粒度级, 容量大; 数据的综合和概括。系统采用 MS SQL Server 2000 建立和管理林业资源和环境数据仓库, 建立数据仓库的步骤如下:

设计和建立林业资源环境多维数据概念模型, 确定多维数据的组织方式, 进行数据仓库建模。系统采用 Analysis Service 进行多维数据的设计, 以多维数据的组织方式管理数据仓库, 以便进行接下来的数据在线分析和数据挖掘。多维数据模型和多维数据组织方式如图 3。

数据的净化、重组、提取、转换。根据多维数据的组织形式和数据仓库的内容, 对数据库中数据进行净化处理, 包括单位、编码的统一; 时间标记的加入; 清除非法数据等操作。按照多维模型进行数据重组, 包括重定义数据结构, 综合粒度结构。最后, 利用 SQL Server 的 DTS 服务进行数据的提取和转换。

数据在线分析。以建立的数据仓库为数据源, 通过 SQL Server 的 Analysis 管理系统建立数据立方体(多维数据集), 提供 OLAP 在线查询和分析服务。系统建立的数据立方体有全国森林资源数据立方体、全国林业社会经济统计数据立方体、全国林业工程建设立方体、海南省社会经济立方体。

对 OLAP 中数据立方体进行数据挖掘, 利用 MS SQL Server 2000 中的数据挖掘模型编辑器

维度 \longrightarrow

时间	地区	森林资源	社会	经济	自然资源	工程
年	全国	权属	人口	投入	土壤	造林
月	省	土地类型	劳力	产出	水份	治理
日	市	面积	...	工资	坡度
...	...	蓄积	坡向
				

图3 林业资源环境多维数据示意图

和向导,根据决策树算法和群集算法,建立数据挖掘模型,对多维数据进行分析,发现隐藏在数据背后的规律、趋势,并将这些数据挖掘的规律、趋势加入到知识管理子系统中。

3 模型管理子系统

如果说管理信息系统是数据驱动的,那么决策支持系统一般都是模型驱动的。模型库能为决策者提供推理、比较选择和分析的工具。同时模型也是整个决策支持系统的最重要的,也是最有特色的部分之一。根据林业的特点和系统涉及的内容,系统设计的用于决策活动的模型分为:

(1) 规划模型:包括线性规划、非线性规划、目标规划等模型。可以用来对森林资源的趋势分析,森林调查规划设计和林业生产计划的编制等方面。

(2) 推理模型:包括归纳推理、类比推理、演绎推理等模型。可以用来对营林、林业生产和林业工程的评价、森林病虫害、森林火灾发生的分析和预警等方面。

(3) 分析模型:包括聚类与判别分析、趋势分析、经济活动分析、投入产出分析等模型。可以用来对林业经营活动的分析,林业投入产出分析等方面。

(4) 预测模型:包括灰色预测模型($GM(1,1)$ 、 $GM(2,1)$ 、 $GM(1,N)$)、灰色关联分析、回归预测模型(线性回归、非线性回归预测)等模型。可以用来对森林资源、森林病虫害和森林火灾等的预警和预测。

(5) 优化模型:包括动态规划(最优路径、投资最优等)、决策树(不确定型决策、风险型决策等)等模型。可以用来辅助林业的规划设计、林网和道路选择以及林业生产投资等方面。

(6) 评价模型:包括投入产出分析、模糊判别、主成分分析等模型。可以用来对林业投资效益分析,林业经济投入产出分析,经营采伐方案的对比分析等方面。

4 方法管理子系统

方法库主要为模型的实现提供计算和分析的方法,方法库作为系统模型的共享资源,一种模型可以选择一个或多个共享方法用于实现。系统设计的方法库主要包括:

(1) 数据统计方法:曲线回归、方差分析、卡平方检验、参数假设检验、多元回归和逐步回归等方法。

(2) 预测方法:灰色预测(各种时间序列法)、线性预测、非线性预测等方法。

(3) 优化方法:线性规划、决策树等各种方法。

林业资源环境在线决策支持系统的方法库和模型库采用模块化进行开发和管理。为了便于方法、模型的复用和扩展,同时考虑程序的网络运行,程序的开发采用面向对象的 Perl 语言进行模块化封装,每一个方法都是一个对象包(Perl Package Model),由方法组建各种模型,每个模型同时也成为一个对外封装的对象,这样既保证了方法库的复用,也可以实现模型库之间的复用(方法和模型库见图4)。

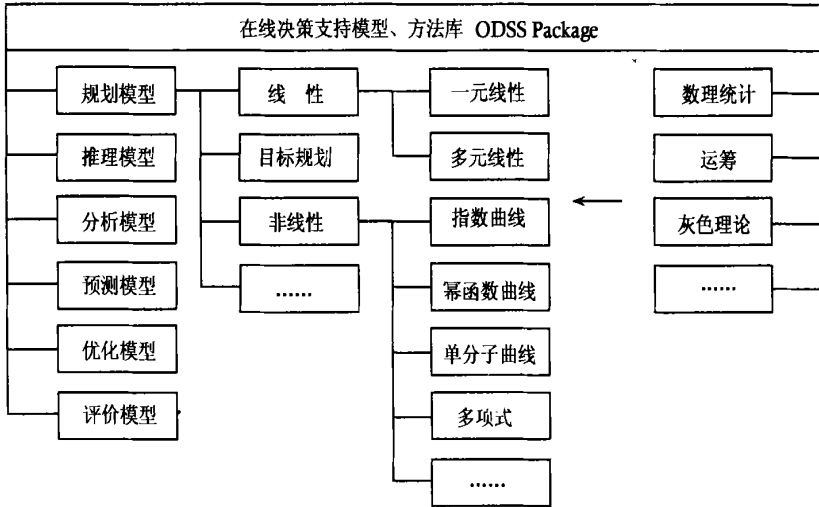


图4 模型和方法库结构示意图

5 知识管理子系统

知识库的内容包括林业专家的知识体系,还包括数据挖掘的知识,表现为概念、规则、规律、趋势、模式和可视化等形式。这些知识经过解释后可以直接在实际系统中应用,用以辅助决策过程。如造林季节、营林方式、抚育期限、间伐力度、采伐期、森林成熟期等专业知识,以及通过数据挖掘得出的森林资源发展、森林病虫害趋势、投入产出规律等知识。由于知识发现、知识推理是不断完善和修正的过程,知识库采用开放的管理结构,运用 MS SQL Server 数据库进行存储、添加和更新。

6 人机交互子系统

林业资源环境网络在线决策支持系统运行于网络环境,因此不仅涉及系统的用户界面接口(即人机交互界面),而且系统的功能实现依赖于网络的接口设计及其运行环境。

6.1 运行环境

系统运行于 Internet 网络环境,采用性能稳定的 Windows 2000 高级服务器系统。数据库和数据仓库采用大型数据库 MS SQL Server 2000 企业版进行管理,能保证林业资源环境的海量数据存储和多用户的并发访问。WWW 服务器采用支持模块化的 Apache Server,既能保证模块化的网络接口,同时能支持大量用户的点击。

6.2 网络接口

网络接口是指 WEB 服务器和数据库之间, 以及 WEB 服务器和决策模型运行之间的机制和接口。目前主要流行的接口程序采用 JAVA/ JSP/ Applet, ASP, ISAPI, PHP, CGI 等进行设计。该系统采用 CGI(Common Gateway Interfact ——公共网关接口), 它的特点是性能稳定, 安全性最好, 支持模块化的编程。CGI 在系统中的运行机制为: 第一, 收集从 WEB 浏览器发送 WEB 服务器的信息, 并使这些信息供辅助决策程序使用; 第二, 辅助决策程序根据选择的决策模型, 调用相应的数据库/ 数据仓库中的数据进行分析和处理。第三, 分析处理的结果交给 WEB 服务器, 最后 WEB 服务器将结果格式化输出于用户的浏览器端。

6.3 人机交互

林业资源环境在线决策系统通过普通网络浏览器进行人机交互, 用户通过网络浏览器提交数据或发送请求, 服务器接受请求, 并根据提交的用户信息, 将数据库或数据仓库中的数据进行查询、分析和处理, 最后将数据分析或决策结果返回给用户的浏览器。具体的人机交互包括:

(1) 用户根据分析和决策目的, 对数据库或数据仓库进行查询, 将查询的结果作为数据分析源; 然后从模型库中选择合适的模型对数据进行分析。分析决策结果显示在浏览器端, 同时也在服务器中保存, 以供下载存储和进一步处理。

(2) 用户对自己的现有数据进行分析 and 辅助决策: 用户选择合适的分析模型, 然后提交数据进行分析。用户提交数据有 3 种途径: 按照系统提供的表单依次输入数据; 将现有数据表中的数据拷入表单中; 直接打开本地数据。

(3) 本地或服务器的非空间数据通过 CGI 程序进行分析和处理, 全过程均为网络在线即时完成, 数据结果直接存储于服务器端, 供用户下载和浏览。

7 结论和讨论

(1) 基于单机和局域网模式的林业决策支持系统经过多年的开发和探索, 已经取得了很大的成果, 得到普遍的应用。但基于 Internet 的网络在线决策支持系统, 还处于新生的探索阶段。由于所受的网络环境和网络技术影响较大, 因此在系统的设计和实现上与传统的决策支持系统存在很大的不同, 两系统之间的直接转换和过渡很困难。但是由于网络的高速发展和基于网络的应用存在着庞大数量的用户, 信息的共享和系统的开放性巨大优势, 而成为新的研究趋势, 并有逐级取代传统系统的可能^[4]。该系统采用了浏览器/ 服务器(B/S) 的多层结构, 系统所有功能均通过网络运行和实现, 与其它信息系统相比具有的优点: 支持多用户并发访问、跨用户平台运行、用户的操作简单、用户的软硬件要求低、实时更新等特点。同时系统采用分布式数据库、实现了多数据源的融合。

(2) 该系统在 MS SQL Server 2000 大型数据库的基础上建立了历年森林资源状况、历年林业生产情况、历年林业社会情况、历年林业经济情况、历年林业工程建设情况、历年林业营林情况和林业自然资源等多类数据库。提供了数据的网络查询、管理和维护功能, 为分析和决策提供了较全面的数据源, 提高了林业信息资源的利用率。

(3) 本系统利用 DTS、OLAP Analysis Services 等工具建立了林业资源环境数据仓库, 开展多维数据的在线分析和数据挖掘, 便于发现数据之间的规律和潜在的关系, 提高了数据分析和处

理能力,使数据直接面向决策分析。建立了统计、分析方法库,并在此基础上建立了6类适用于网络运行的模型库,实现了林业资源环境在线决策支持的网络运行。为数据的网络在线分析和处理开辟了新的内容。

(4) 系统使用面向对象的Perl编程思想设计CGI程序,开发了易于扩展、能直接运行多个系统平台方法库,在此基础上建立6类决策模型(规划、推理、分析、预测、优化、评价)程序库,每一模型均以一对象模块封装。程序的开发采用开放式源代码,使功能的优化和模型的扩展变得非常简单。

(5) 本系统主要针对林业资源环境的统计数据 and 属性数据进行分析,为决策者提供有效的辅助决策支持。在Internet的试运行中,性能稳定,能支持大量用户的并发访问,决策结果和图形的输出正常。本研究收集的数据内容主要针对于部级水平,在保证数据更新的前提下,系统也可以运行于具有网络条件的省市水平和地区级水平。

(6) 林业资源环境数据存在着大量的空间数据,运用空间信息进行网络在线辅助决策应该是另外一个值得研究的重要内容。但是由于涉及到空间数据的网络表达和网络数据的传输瓶颈等问题,所受的网络环境影响更大。国内外在近几年来开始了这方面的探索和研究,如Autodesk, Esri, Mapinfo 等公司开发了成型的空间信息表达产品,但要开发与专业领域相结合、成熟的空间信息决策支持系统,是有待进一步研究的课题。

参考文献:

- [1] 陈晓红. 决策支持系统理论和应用[M]. 北京:清华大学出版社,2000
- [2] 高洪深. 决策支持系统(DSS)理论、方法、案例(第二版)[M]. 北京:清华大学出版社,2000
- [3] Inmon W H. 数据仓库[M]. 王志海等译. 北京:机械工业出版社,2000
- [4] Power D J. Web based decision support systems[J]. The On Line Executive Journal for Data Intensive Decision Support, 1998, (2): 33-34

Study on Online Decision Support System of Forestry Resources and Environment Information

ZHANG Huai-qing, JU Hong-bo, CHEN Yong-fu

(Research Institute of Resource and Information Technique, CAF, Beijing 100091, China)

Abstract: In order to provide quick, in time and exact online data analysis and decision support for decision-maker by means of networks, a network database of forestry resource and environment information on ministry level was built. Based on this, the data warehouse which support multi-dimension data was designed and set up. At the same time, the modularized method, model and knowledge database were developed using the object-oriented ideal. An online decision support system of forestry resource and environment information was built by using the technique of Common Gateway Interface (CGI) to realize the interface of users, databases and the system.

Key words: forestry resources; decision support system; network online decision support system; data warehouse