

商业智能在科技服务决策支持系统中的应用

李心科, 林 昕

(合肥工业大学 计算机与信息学院, 安徽 合肥 230009)

摘 要:在介绍了商业智能的核心技术和体系结构包括数据仓库、联机分析处理、数据挖掘等的基础上,通过分析科技信息服务这个特定领域的决策需求,给出了基于商业智能的科技服务决策支持系统的解决方案。实际应用表明该系统提高了科技信息供需双方对现有信息数据的利用效率,还提高了科技信息服务机构的决策分析能力。

关键词:商业智能;决策支持系统;科技服务

中图分类号: TP311.5

文献标识码: A

文章编号: 1673-629X(2007)12-0200-04

Application of Business Intelligence in Technological Service DSS

LI Xin-ke, LIN Xin

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

Abstract: Introduces the main technology and system structure of business intelligence including data warehouse (DW), OLAP and data mining. By analyzing the decision demand in the specific field of technological service, put forward the technological service DSS solutions based on business intelligence. The result in the practical application indicates that the using efficiency of the existed technological information is enhanced, and the analytic ability of the technological service organization is improved.

Key words: business intelligence; DSS; technological service

0 引 言

目前在科技服务体系中存在一些问题,如科技成果的生产力转化率低、转化成本高,科技投入存在地域和行业的不平衡等。一个重要的原因就是科技成果转化的双方所获取的科技信息有限,沟通成本较高。就科技中介服务机构来说,虽然占有大量的科技信息资源,但缺乏对这些信息资源的科学分析手段和工具,不能有效地从这些信息资源库中提取有用的信息,缺乏对供需双方相应的技术咨询和决策支持。如何把已有的海量的科技信息数据转换成更有价值的信息,以便用于决策支持呢?商业智能被广泛认为是最好的解决方案之一。

把商业智能(BI, Business Intelligence)技术用于科技服务决策支持系统,对于充分利用已有大量科技服务信息,提高信息供需双方辅助决策能力和科技中介机构的服务能力有重要的理论价值和现实意义。

1 商业智能相关概念

商业智能的概念最早是 Gartner Group 于 1996 年提出来的。当时将商业智能定义为一类由数据仓库(或数据集市)、查询报表、数据分析、数据挖掘、数据备份和恢复等部分组成的、以帮助企业决策为目标的技术及其应用。BI 将数据仓库、联机分析处理(OLAP)和数据挖掘等结合起来应用于商业活动中,从不同的数据源收集的数据中提取有用的数据,对数据进行清洗以保证数据的正确性,将数据经转换、重构后载入数据仓库或数据集市;然后利用合适的工具对数据进行处理,这时信息变为辅助决策的知识;最后将知识呈现于用户面前,为管理决策提供参考。可见 BI 并不是基础技术或者产品技术,而是一种解决方案^[1]。

商业智能的核心技术主要包括三个关键部分:数据仓库和数据集市、联机分析处理以及数据挖掘,其中数据仓库和数据集市是商业智能的基础,而 OLAP 和数据挖掘是对数据仓库中的数据进行分析和处理的重要工具。

数据仓库是一种支持企业或组织管理决策过程的、面向主题的、集成的、稳定的、随时间不断变化的管理技术^[2]。主题是一个抽象的概念,在逻辑意义上,它是对应某一宏观分析领域所涉及的分析对象;集成是

收稿日期:2007-03-06

基金项目:安徽省科技攻关重点项目(0012021A)

作者简介:李心科(1963-),男,江苏徐州人,副教授,博士,研究方向为软件工程、神经网络等。

指原始数据库的数据进入数据仓库前,首先经过清理、抽取、转换等加工,将原始数据的组织方式从面向应用转换到面向主题;稳定是指数据经集成进入数据仓库后就极少或根本就不改变,随时间不断变化是指数据仓库被不断追加数据,而其中已有的数据逐渐成为历史数据,数据具有时间属性。数据仓库以构建新的分析处理环境,为决策者提供信息,支持决策分析处理为目标。

联机分析处理的核心就是多维分析,它支持用户从多角度、多个层面观察和处理数据,其用于分析的基本动作主要有切片、切块、旋转、钻取等^[3]。按照数据组织方式的不同,OLAP 的实现方式目前主要有三种:基于多维数据库模型的 OLAP——MOLAP,基于关系数据库模型的 OLAP——ROLAP,混合型 OLAP——HOLAP。具体选择何种实现方式,要依据用于存储的数据量大小、要求的数据存取速度等因素来决定。

数据挖掘就是从大量的、不完全的、有噪声的、模糊的、随机的实际应用数据中,提取隐含在其中的、人们事先不知道的但又是潜在有用的信息和知识的过程^[4]。数据挖掘实质是一种深层次的数据分析方法,它能够在数据库、数据仓库中自己寻找模型,它在本质上是一个归纳的过程。

BI 体系结构是指通过识别和理解数据在系统中的流动过程和数据在企业的应用过程来提供 BI 系统应用的主框架。BI 的体系结构主要包括数据预处理、数据仓库、数据分析以及数据展现等几部分,BI 基本体系结构如图 1 所示。

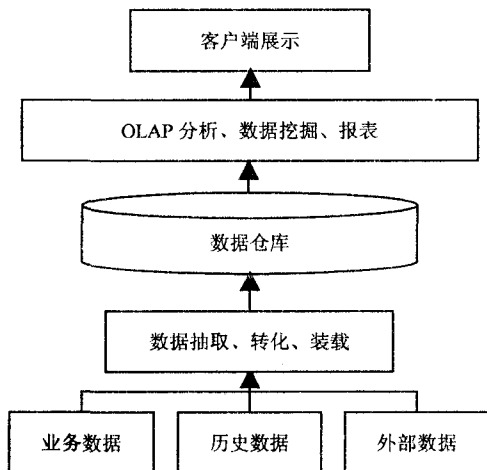


图 1 BI 基本体系

2 科技服务 DSS 的系统设计

2.1 系统总体结构

科技服务 DSS 是为科技信息供需双方和科技信息服务机构提供决策支持信息的。该系统采用 B/S

架构,服务器端提供的服务包括数据逻辑、数据服务、性能监控以及元数据存储、数据挖掘和数据聚集等。客户端执行的功能包括用户界面、查询分析、报表格式化以及数据访问等功能。系统结构如图 2 所示。

数据源层:也可称作操作型数据层,是整个数据仓库的基础,提供了整个系统最原始的数据。通常为业务数据库和其它外部数据,本系统数据库为科技厅科技信息历史数据库以及本年度各个地上市上报的科技项目信息。

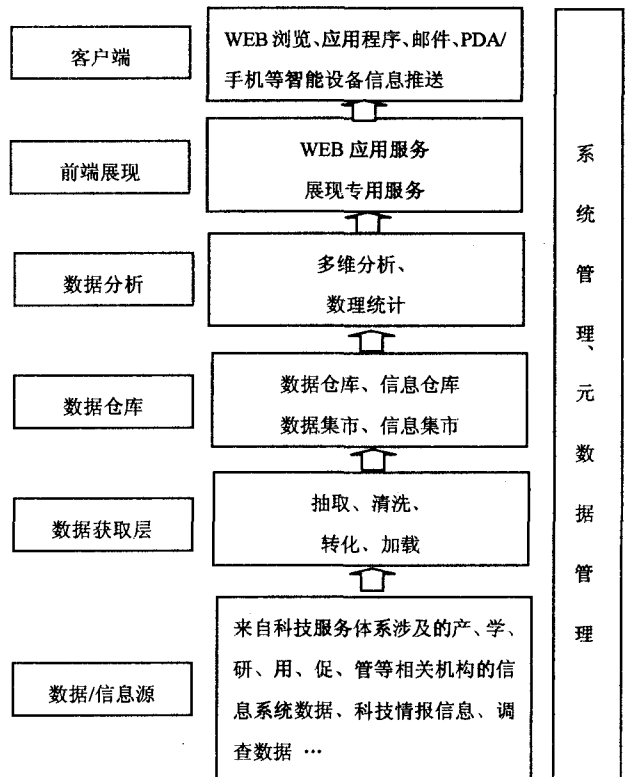


图 2 系统总体结构

数据获取层:也可称作数据转换层,主要是把数据源层的数据通过 ETCL 过程转换到数据仓库中,包括数据的抽取(Extract)、转换(Transform)、清洗(Cleaning)和加载(Load)4 个部分,这一层在一定程度上决定着数据仓库中数据的质量。ETCL 的其目的是从数据源中抽取该系统所需要的数据,对其进行检验和清洗,并根据数据仓库设计要求对数据进行重新组织和加工,并装载到数据仓库的目标数据库中。

数据仓库层:该层是按主题进行分析和对相关的数据进行挖掘的数据源,包括每一个按主题进行分类的数据集市和专门用于数据挖掘的数据挖掘库。

数据分析服务层:该层是数据存储和前端分析工具的桥梁,它包括 OLAP 分析引擎、安全控制机制等等,能按照用户的要求设计、生成具有多维分析功能的分析主题,予以组织,以便进行多角度、多层次的分析,

并发现趋势。它们响应前端用户的分析请求,将多维数据传送给前端的分析工具显示。

前端展现层:在用户眼中,用户界面的显示才是最重要的,本系统中采用 Hyperion Performance Suite 来设计基于 Web 的数据展现和图形展现,并提供给用户多种查询方式,能根据用户要求钻取到相关层,获得相关的明细数据。

2.2 数据仓库的设计

数据仓库的指导思想是多维数据模型围绕中心主题进行组织。该系统把历年来的科技信息和各个地市本年度上报的科技信息按照主题进行重新组织。同一主题的相关数据作为一个数据立方存储在一起,立方中存储的数据称为度量,它是决策人员分析的对象,放在概念模型的中心位置,概念模型的四周为分析问题所考虑的维。进行数据分析时,有时需从宏观上把握,有时需了解具体细节,所以同一维的数据要具有多个层次,从而实现数据的上卷、下钻、旋转、统计等分析操作。

2.2.1 科技服务信息数据仓库的主题

数据仓库中的每个主题对应一个宏观分析领域所涉及的分析对象。数据仓库从不同的数据源抽取数据,面向不同的主题分别进行存储,并根据分析所需对数据进行预聚集。经过与科技服务中介机构 and 科技厅管理人员多次交流,认真分析,发现科技信息供需双方进行的决策主要有以下几个方面:科技项目立项时的审批、成果推广时的导向、科技经费的分配等。与此对应,本模型中分别用立项主题、成果主题和经费主题来表示。

2.2.2 科技服务信息多维概念模型中的维

决策人员在分析问题时,针对不同的主题,考虑问题的角度也往往不同,这里的角度在数据仓库技术中称为维^[5]。经过与科技厅管理人员的反复交流,发现该厅进行项目立项审批时所考虑的因素主要有项目负责人的情况、项目本身的情况、项目负责人所在单位的情况等几个方面,在多维概念模型中分别用负责人维、项目维、单位维表示。在成果推广时主要考虑成果的市场需求、利润情况、对环境的影响等几个方面,分别用需求维、利润维、地区维表示。在考虑经费的分配时,主要考虑项目进展情况、项目负责人情况等,分别用进展情况维和负责人维表示。当然,数据仓库中的所有信息都是与时间密切相关的,科技服务信息多维数据仓库中的数据立方均含有时间维。

2.2.3 科技服务信息多维概念模型中维层次的划分

在科技服务信息多维概念模型中,维层次的划分方法根据维性质的不同而不同。时间维、地区维、专业

维、职称维、学历维、年龄维等需要进行上卷及下钻操作,需进行层次的划分;负责人维、项目维等不需上卷及下钻操作,就不需要划分层次。在需要划分层次的维中,有的维具有多个层次,上下层之间具有包含与被包含关系,如时间维具有年、月、日、时、分、秒等层次,且高层次包含低层次;而职称维、学历维、进展情况维等的层次之间不具有包含关系,它们只是将全集按某一标准划分为不同的类,划分后的所有成员都属于全集;年龄维和利润维的值为数值型,取值范围广,本研究将其划分为几段。在多维概念模型中,任何一个维都具有默认的“全部”层次,它属于所有层次的最高层。

2.2.4 科技管理信息多维概念模型中的度量值

建立数据仓库的目的是供决策人员分析时使用,仓库中存储的数据应视分析的需要而定。根据科技服务信息决策中分析的需要,确定科技服务信息多维概念模型的度量如下:立项主题的度量有课题数量和资助金额两个,成果主题的度量是成果应用的次数和成果所创价值,经费主题中则有总经费、已用经费和本期计划经费三个度量。

2.3 OLAP 设计

在本系统中,采用 ROLAP。ROLAP 尽管在数据存取速度上、多维计算能力上不如 MOLAP,但它也具有自己特有的优势:数据存储容量大,对软硬件平台的适应性比较好,最重要的是,它可以很好地适应维度变化^[6]。ROLAP 最常用的数据模型是星形模型,图 3 就是笔者为实施 ROLAP 所建的数据仓库的星形模型。对于 OLAP 的前端展现,我们采用 Web 方式来代替传统的 C/S 方式,如图 3 所示。Web 具有很好的跨平台性,客户只需利用浏览器而无需安装其他的终端软件就可以浏览丰富多彩的信息,这有利于我们对数据的集中管理,简化用户的安装。这种方式非常适合机关和企事业单位使用,他们无需进行任何培训就可以非常方便地通过网络查看他们关心的信息。

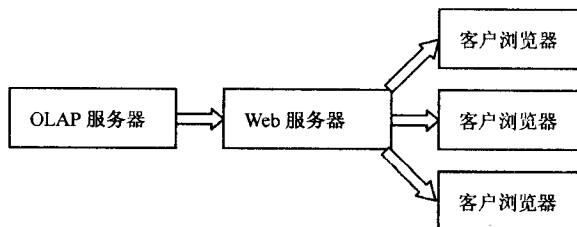


图 3 OLAP 服务器 Web 展现方式

2.4 智能分析引擎和数据挖掘

OLAP 是一种强大的数据分析技术,但是它有自己的局限性。它只能通过对历史数据和当前数据的复杂分析来告诉用户所关心主题的去和现状,并不能预测将来的状况和事物发展趋势。

智能分析引擎是实现“数据→信息→知识→智慧”提升的关键。智能分析引擎融合多维分析、数理统计、数据挖掘与知识发现、模式识别、专家系统等智能分析和处理技术构建而成,并与前端展现层无缝结合,通过丰富、直观的展现形式将分析结果呈送给相关用户,实现联机分析和决策支持功能。

为了得到科技信息更加科学的预测模型,最好的方法就是从数据仓库中存储的当前数据和时间跨度较长的历史数据中找出科技信息随时间变化的规律,从而形成预测模式,这就是笔者使用数据挖掘的目标。数据挖掘将其挖掘结果——数据模式作为预测模型添加到预测模型库中,从而不断丰富和修正现有的预测模型。

3 结束语

在当今信息时代,科技信息供需双方往往被淹没在来源于多个渠道的、庞大的、丰富的海量数据中,只有及时地将数据有机地组合在一起,及时地将信息转化为知识和智能,才能更好地指导企业进行决策和行动。商业智能的作用就在于此,它帮助管理者做出科学的计划、判断、决策,避免主观、片面等因素引起的重大失误。

(上接第 196 页)

参考文献:

- [1] Negus C. Red Hat Linux 9 宝典[M]. 王 勇译. 北京:电子工业出版社,2004.
- [2] 毛曙福. Linux C 高级程序员指南[M]. 北京:国防工业出版社,2001.

(上接第 199 页)

主程序读入 SDRAM 并开始执行。当 A 机通话时,A 方的音频和视频数据通过 VI、AI 端口输入被主程序处理后经过以太网发送到网络,传送到对方 B,B 机对数据进行处理并把视频数据通过 VO 端口输出到图像显示设备,音频数据通过 AO 端口输出到喇叭。同理,B 方传输自己的声音和图像给 A 机,这样就实现了双方的可视通信。

3 结束语

主要介绍了嵌入式多媒体通信终端的设计。该终端的实现可以应用于楼宇智能化远程抄表系统、智能家居系统、自动售票检票交通收费系统、网络控制显示屏系统、工业智能化从站系统、数码影像网络监视防盗系统、远程医疗诊断系统等等,具有很高的实用价值。

文中主要从技术的角度对商业智能的基本概念、技术架构和涉及到的技术进行了探讨,并将商业智能应用于科技服务决策支持系统中。通过商业智能系统的构建,科技信息供需双方和科技信息服务机构可以高效地利用系统中的科技信息,大大提高用户对信息查询分析和决策分析的能力。切实提高了技术成果转化双方的决策能力,提升科技服务机构的核心竞争力。商业智能引导科技服务体系的信息化建设进入了一个全新的境界。

参考文献:

- [1] 韩清池,陈世权. 商业智能及其应用的研究与发展[J]. 现代管理科学,2006(3):68-69.
- [2] 林 宇. 数据仓库原理与实践[M]. 北京:人民邮电出版社,2003.
- [3] Thomsen E. OLAP 解决方案:创建多维信息系统[M]. 第 2 版. 北京:电子工业出版社,2004.
- [4] 朱 明. 数据挖掘[M]. 合肥:中国科学技术大学出版社,2002.
- [5] Kimball R. 数据仓库工具箱:维度建模的完全指南[M]. 第 2 版. 北京:电子工业出版社,2003.
- [6] 袁 虹,何厚存. 联机分析及数据仓库的建模技术[J]. 计算机应用研究,1999,16(12):61-63.

- [3] Redhat, Ltd. Fedora Core Document[EB/OL]. 2007. <http://www.fedoraproject.org/>.

- [4] Bloom R B. Apache Server 2.0 技术参考大全[M]. 北京:清华大学出版社,2003.
- [5] Apache Group. Apache Web Server Document [EB/OL]. 2004. <http://httpd.apache.org/>.

参考文献:

- [1] Philips 公司. ilips Semiconductors, Inc. TM1300 Preliminary Data Book[M]. Netherlands: Philip 公司,1999.
- [2] Philips 公司. TM1300 Data Book[M]. Netherlands: Philip 公司,1999.
- [3] Philips 公司. Integrated Systems, Inc. pSOSystem system concepts[M]. Netherlands: Philip 公司,1997.
- [4] 曹羽承,莫德举. 基于 pSOS 的 TM1300 应用系统中的 BSP 研究[EB/OL]. 2004-10. 21IC 中国电子网. <http://embedded.21ic.com/>.
- [5] Philips 公司. Integrated Systems, Inc. pSOSystem Programmer's Reference[M]. Netherlands: Philip 公司,1997.
- [6] 赵 海. 嵌入式 - internet[M]. 北京:清华大学出版社,2001.