

DOI: 10.3969/j.issn.1672-8874.2009.02.010

多维分析在军校教学管理中的应用研究^{*}

刘青宝, 吴俊, 陆昌辉

(国防科学技术大学 信息系统与管理学院, 湖南 长沙 410073)

[摘要] 针对当前军校教学管理过程中长期积累的大量业务数据得不到充分利用的现状, 本文采用数据仓库和联机分析及数据挖掘技术, 建立了军校教学管理多维数据模型, 并进行了多维分析系统的设计与实现, 从而支持教学管理者对大量繁杂的业务数据进行多角度的综合分析, 为军校教学资源的优化配置和教学训练方案的科学制定提供决策支持。

[关键词] 教学管理; 多维数据分析; 数据仓库; 数据挖掘技术

[中图分类号] E251.3 [文献标识码] A [文章编号] 1672-8874 (2009) 02-0034-03

Research on the Application of the Multi-dimensional Data Analysis in the Teaching Management of the Military Academies

LIU Qing-bao, WU Jun, LU Chang-hui

(College of Information System and Management, National University of Defense Technology, Changsha 410073, China)

Abstract: To make full use of the massive teaching management data accumulated in teaching, this paper focuses on the research of the application of multi-dimensional data analysis in teaching management. Using the technique of data warehouse and on-line analysis processing, this paper proposes a multi-dimensional data model and designs the multi-dimensional data analysis system of the teaching management in the military academy to support the optimized allocation of the teaching resources and the scientific making of teaching plans.

Key words: teaching management; multi-dimensional data analysis; data warehouse; data mining

随着各类军校教学业务系统的逐年完善, 每年都积累了大量的学习训练情况、教员队伍建设、教学保障、教员学员奖惩情况等各种教学业务数据。如何充分利用这些数据, 从中分析并挖掘出有用信息或规律, 为军校教学管理进行科学决策提供全面的信息服务, 已成为管理决策者迫切关心的问题。

传统面向决策分析型的教学管理系统大多是基于指标统计报表的教学评估与分析系统。这些传统分析系统面临着一系列的固有缺陷, 如数据堆砌问题(表格堆砌了大量数据, 到底这些数据代表了什么信息、什么趋势? 高层的决策者对报表数据越来越难以理解)、难以交互式分析(定制好的指标统计报表过于死板, 难以动态交互分析)、难以追溯历史, 形成数据孤岛(教学业务系统虽大量增多, 但数据存在于不同教学管理部门; 旧的数据往往被业务系统备份出去, 导致宏观分析、长期历史分析难度大)。数据仓库是一种新型的信息管理技术, 自上世纪 90 年代以来, 数据仓库技术的研究和应用得到了广泛关注。基于数据仓库的多维分析和数据挖掘技术为克服上述问题提供了有力的技术支持。

一、教学管理多维数据模型设计

当前军校教育管理的数据规模空前增长, 而且数据门类繁多。为了对这些长期积累起来的大量繁杂数据进行科学分析, 必须采用一种新的模型来实现数据的组织。一般决策所需的数据或指标总是与一些维度(每一维度代表对数据的一个特定的观察视角)和级别(如部门级别、学历等级)的统计和计算有关。以多维数据模型为核心的多维数据分析是数据辅助决策的主要内容, 数据仓库的多维特征满足了数据分析与决策支持对数据的要求, 克服了业务型数据库的数据组织性差、利用率低的缺点。

教学管理多维数据模型以描述数据的多维特征为目标, 形成一个模拟决策者实际观察和分析数据的多维逻辑视图。在多维数据模型中, 数据是按主题进行组织的, 这些数据可以分为两部分: 一部分是决策者所要分析的对象, 通常称为事实, 它是一些度量信息; 另一部分是决策者进行分析时考察的角度(或视角), 通常称为维, 它包含的是关于度量的描述性信息。

* [收稿日期] 2009-04-12

[基金项目] 国防科学技术大学“十一五”重点课程《决策支持系统》建设项目。

[作者简介] 刘青宝(1967-), 男, 江西永新人, 国防科学技术大学信息系统与管理学院副教授, 博士。

(一) 多维分析主题的确立

主题是一个抽象的概念,是指用户使用数据仓库进行决策分析时所关心的重点方面,是在较高层次上将教学业务系统中的数据进行综合、归类,以支持对教学管理问题的全方位分析。

根据对原有分散的各教学业务管理系统的分析,结合军校教学管理工作的现实需求,可得到如下教学管理主题:学习训练情况、教学保障情况、师资建设情况、教学质量情况、安全管理情况等多个主题。以学习训练主题为例,它综合了来自学员学籍管理系统、选课系统、成绩管理系统、军事体能训练管理系统、实践活动以及兴趣科研小组管理系统等反映学员学习训练情况的所有信息。

(二) 多维分析的维度设计

首先建立以下传统的维度:时间维度、单位维度、学员维度、课目维度、教员维度和教学设施设备维等。并根据军校教学管理的特点,时间维分为学期、学年、阶段(含军训、基础合训、专业培训等阶段)、整个学制五个层次;单位维分为学校、学院、系、学员队、区队等多个层次。由于学员是教学管理的核心,为了满足多视角分析的需要,对学员维建立多个维层次结构:学历层次结构、学科专业层次结构、培养目标分类结构(技术类、合训类、指挥类)。

在维度划分过程中,还存在一些数据维度不清晰,无现存的维度层次结构可以参考的情况。如对学员的综合评价分类,不好划分其维层次结构。对于这类数据我们采用非传统的方法,引入层次聚类算法。通过层次聚类把无层次的数据聚合成几个层次,建立数据挖掘维层,来弥补传统维度层次划分的不足,使维度层次划分更加科学、合理和规范。为了使军校教学管理与实施单位更有针对性地制定教学策略,采用聚类分析算法,对学员进行非传统的动态区分。由于聚类算法的冗长,这里不作介绍,具体的计算过程可以通过统计软件SPSS来实现。以某军校教学管理系统中学员的考评成绩数据为例,进行以下数据预处理:以单个学员为数据对象,所有课目为属性,对课目成绩进行布尔转换:优良以上为1,否则为0。对转换后的成绩数据采用布尔属性聚类算法。

(三) 多维分析模型的建立

根据以上所确立的主题和分析维度,建立教学管理的多维数据模型。学员学习训练情况多维数据模型如图1所示。

由于篇幅有限,其它模型不作一一描述。

二、教学管理数据仓库的构建

建立教学管理数据仓库是一个复杂的任务,因为涉及这样一个系统如何存储和管理它的数据;如何构造一个提取程序,将数据从各部门的教学业务数据库转换到面向分析的数据仓库;如何构造一个数据仓库刷新程序,合理地保持数据仓库中的数据相对当前性和一致性。

(一) 数据视图的设计

在教学管理数据仓库系统建设中,首先要建立如下三类视图:业务数据源视图、数据仓库视图、多维数据查询视图。

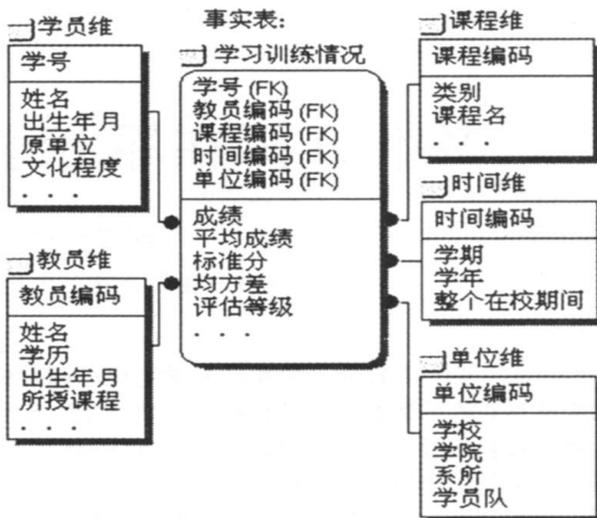


图1 学习训练情况多维分析模型

教学业务数据源视图描述各业务数据库中为决策分析所能提供的各种原始教学数据及其业务含义。这些视图信息以不同的详细程度和精度建档,为仓库数据的抽取、预处理和集成装载提供指导。如学员学习训练情况分析主题的业务数据源包括学生学籍库、学习成绩库、军事体能训练库、学生奖惩信息库、实践活动数据库等。在这些数据源中,有学生档案表、选课表、成绩表以及军体训练记录表、实践活动记录表等,其中大部分数据项是全校规范的和一致的,如学号和课程号等,但也有不少非规范、非结构化的数据项,如一些非考试科目的考核记录,缩写很不规范;对实践活动的记录甚至就是一段视频。数据仓库视图包括分析主题的事实表和维表,它们提供存放在数据仓库内部的信息。多维查询视图是从最终用户的角度透视数据仓库中的数据。

(二) 数据预处理

军校教学管理的历史数据时间跨度比较长,数据内容也因院校整合、专业调整等因素的影响发生很大改变。要实现数据的抽取、转换、清洗、载入等数据转移过程是一项繁杂而艰巨的工作。本文主要利用了DTS的强大数据转换功能来实现从Visual FoxPro、Excel、Access、SQL Server、Oracle数据源导入数据到SQL Server2005中。针对某些应用需求,还编写了特定的程序来实现有关的数据转换工作。例如,根据学习训练成绩分析的要求建立了标准分这一新的度量值,以满足学员的学习训练成绩与前面学年学习训练成绩进行比较分析的需要。如图2所示为学习训练中数据转移过程的DTS包。

三、军校教学管理多维分析系统的设计与实现

本系统在完成决策分析的同时也兼顾了军校教学管理综合信息的常规查询。主要功能包括:抽取业务数据,并对数据进行清洗转换等预处理;集成并装载细节数据至数据仓库基本库;对细节数据进行初步分析,聚集、综合形成各种粒度的决策数据;以数据可视化技术把数据仓库中的数据按照用户分析的维度进行多维交互式展现,支持对多维数据的切

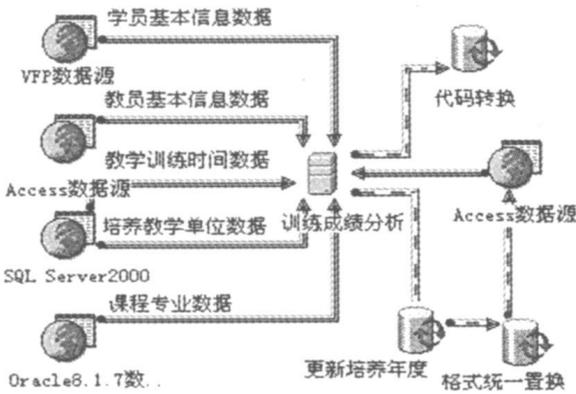


图2 学习训练成绩情况数据的 DTS 包

片、钻取等联机分析操作。系统功能结构如图3所示。

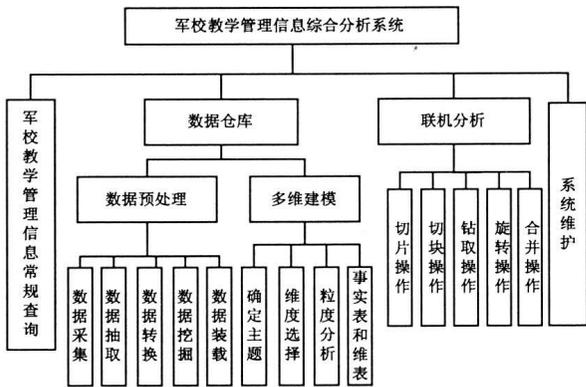


图3 系统功能结构图

(一) 系统体系结构设计

数据仓库系统的体系结构决定了数据加载、访问和传递的方式。军校教学管理多维分析系统的总体结构按三层设计：数据抽取转换层、仓库数据管理层、多维分析与展现层。系统总体结构如图4所示。

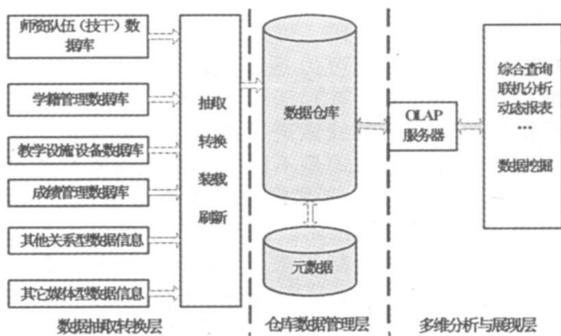


图4 系统总体结构图

(二) 多维分析与展现层的实现

多维数据数据分析与展现层是直接面向用户的，是支持教学管理决策的核心部分。用户通过该层界面对数据仓库中的教学管理数据进行多维分析、汇总，形成图表或报表的形式，使决策者可以清晰、直观地看到教学管理的决策分析结果。

多维数据数据分析与展现层主要由三个模块构成：多维数据集解析模块、图表展现模块以及数据导出模块，如图5所示。

多维数据集解析模块负责对连接的多维数据集进行解

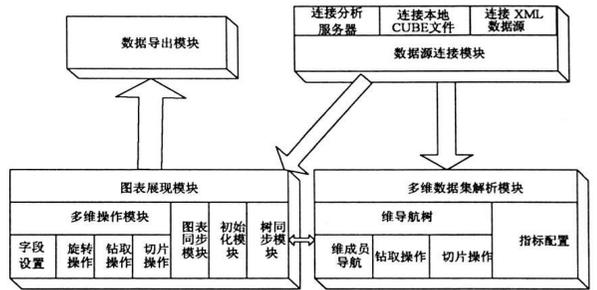


图5 多维分析与展现层结构图

释，分为维导航树和指标配置两部分。通过它不仅可以获得多维数据结构（其中包括维结构信息和立方体结构信息），还可以根据需要对其进行编辑。图表展现模块主要负责处理图表初始化设置、交互式多维操作、基于导航树的图表数据同步等。数据导出模块主要是将联机分析的结果导出成 Excel 文件以便用户进行打印存档，同时，此模块还提供视图保存的功能。用户可以将当前的分析状态以视图的形式保存在一个 XML 文档中，下次分析时只要导入该 XML 文档就可以从保存的状态继续进行分析。

本系统所使用的 Microsoft SQL Server 2005 Analysis Services 平台，支持应用程序

通过 OLE DB 和 ADO 使用 MDX 语句来建立本地化的立方体和查询立方体数据。当用户需要执行一多维数据分析操作时，首先应通过用户界面来得到用户的具体查询请求，并将此请求转化为适合于 OLAP 查询的 MDX 语句。然后使用 OLE DB 或 ADO 将此 MDX 语句送交给 OLAP 服务器执行并得到查询结果。最后将此结果以直观易懂的形式显示给查询者。

对于多维数据库的显示，可以有很多种实现方式，本系统采用了在 NET 中调用 Dundas Olap services 控件来完成对多维数据集的显示。

[参考文献]

- [1] 邓苏等. 决策支持系统[M]. 北京: 电子工业出版社, 2009.
- [2] 孟跃红, 黄智. 数据仓库技术在高校教育管理中的应用[J]. 徐州师范大学学报(自然科学版), 2003, (2).
- [3] 杨晓文. 学生成绩分析数据仓库的模型设计及实现[J]. 山西电子技术, 2005, (1).
- [4] 李继云等. 高校老师教学质量评价方法的研究[J], 统计与咨询, 2005, (3).
- [5] 安淑芝等. 数据仓库与数据挖掘技术[M]. 北京: 清华大学出版社, 2005.
- [6] 董一鸿. 大型数据库中基于领域连接的层次聚类算法[J]. 计算机工程与应用, 2003, (32).

(责任编辑: 阳仁宇)