

基于 Analysis Services 的教学评估分析系统的设计

陈荣鑫¹, 陈维斌², 黄杰圣¹

(1. 集美大学 计算机工程学院, 福建 厦门 361021; 2. 华侨大学 计算机科学系, 福建 泉州 362011)

摘要:在本科教学水平评估的工作实践中, 如何对评估指标进行有效跟踪和自评, 为教学的科学决策提供依据, 同时避免传统事务统计型管理软件的固有缺陷, 增强对教学决策的支持, 是一个重要课题。提出构建基于 Analysis Services 的分析型应用系统, 系统的基础是数据仓库和 OLAP。在给出了主题分析与数据立方体的设计、ETL 设计、分析界面设计等技术方案的同时, 也阐述了系统实施的相关内容。

关键词:教学评估; 数据仓库; OLAP; OWC

中图分类号: TP311; G434

文献标识码: A

文章编号: 1673-629X(2007)08-0233-03

Design of Teaching Evaluation Analysis System Based on Analysis Services

CHEN Rong-xin¹, CHEN Wei-bin², HUANG Jie-sheng¹

(1. School of Computer Engineering, Jimei University, Xiamen 361021, China;

2. Computer Science Department, Huaqiao University, Quanzhou 362011, China)

Abstract: How to effectively track evaluation guideline, perform self-evaluation and support decision in the practice of UTLE, meanwhile avoid the defects of management software based on transaction and statistics is an very important problem to be solved. An application which is based on analysis services is presented. The infrastructure of the system is data warehouse and OLAP. Topics including subject analysis and data cube design, ETL design and interface design are presented, along with some content relative with the implementing of system.

Key words: teaching evaluation; data warehouse; OLAP; OWC

0 引言

本科教学水平评估是促进教育改革深化以及教育质量提高的重要途径^[1]。评估指标体系为各高校本科教育质量提供一个具备可比性的具体化目标。如何对评估指标进行有效跟踪和自评, 为教学的科学决策提供依据, 真正做到评建结合, 是一个重要课题。

目前出现了各种教学质量应用系统, 采用的评估指标大多是基于统计的显式指标, 这种系统基本属于事务型、统计型的系统。它有两个重要缺陷: 一是数据集成能力弱, 面对需要处理的庞杂的历史数据显得力不从心, 对于数据来源和格式的多样性, 这些系统往往需要人工进行导入和调整, 自动化程度不高, 效率低下, 而且容易出错; 二是分析功能有限, 指标需要预先

制订, 灵活性受限制, 很难形成高质量的分析报告。

基于数据仓库的联机分析处理(OLAP)方案可以成功解决以上问题。借助数据仓库的支持, OLAP 对大批量历史数据的卓越分析处理能力, 可以满足对各种教学活动产生的大量数据进行集成和实时分析处理的需要, 包括各类可视化报表的自动生成, 让处在不同层次、不同部门的教学管理者或领导能够从中发现教学活动中的主要问题, 以便及时改进, 进而辅助领导决策做好学校管理, 提高学校管理能力和水平。

本课题以集美大学评价办对院系级细化评估指标体系^[2]为蓝本, 旨在构建以教学评估数据仓库为基础的分析系统, 结合 OLAP 应用辅助教学决策。从开发、部署、运行维护成本等方面综合考虑, 选用 Analysis Services 作为系统设计和实施的软件平台。

1 Analysis Services 数据仓库与 OLAP

数据仓库提供了多维分析的物理基础。待分析的数据存在于关系数据库、多维数据库或客户文件中, 需

收稿日期: 2006-11-19

作者简介: 陈荣鑫(1975-), 男, 福建厦门人, 硕士, 研究方向为数据库技术及应用、数据仓库、数据挖掘; 陈维斌, 教授, 研究方向为数据库技术及应用、数据仓库、数据挖掘; 黄杰圣, 副教授, 研究方向为计算机应用技术。

要通过 ETL 工具把游离数据转换成数据仓库中的数据。OLAP 以多维分析为基础,可以根据需要通过下钻、上卷、旋转、切片、切块等基本操作,方便地对数据仓库中的数据进行多层次、多角度的分析处理,并以合适的方式输出分析结果。

Analysis Services 是优秀的分析服务器,它提供了构建企业级的数据仓库和在此基础上进行 OLAP 的软件环境^[3,4]。通过维度建模,可以把数据组织成星型或雪花型的多维数据集,由于该数据集包含了预计算的聚合数据,客户端可以从服务器中高效获取分析数据和结果。Analysis Services 提供良好的开发接口,比如 MDX 和 ADO MD 提供了有效操纵多维数据的强大途径,极大降低了分析系统的开发难度。

2 评估数据仓库的设计与实施

2.1 主题分析与数据立方体设计

本科教学水平评估的内容覆盖 7 个一级指标和 19 个二级指标^[2],在二级指标下又包含多个主要观测点。把体系分为教学条件、教学管理、教学效果等部分,每部分定义了多项分析主题,比如教学管理部分主要有学生学习质量和教师教学质量两大主题。各个主题还需要进行层次划分,每个主题基本映射到相应的评估指标。

以下是简化了的学生成绩的子主题建模实例:

立方体模型 MD:=(D,M);

维度 D={d_{time}, d_{course}, d_{department}};

时间维 d_{time}=({term, year, d_{time.all} }, term≤year≤d_{time.all}),

院系维 d_{department}=({student, major, department, d_{dept.all} }, student≤major≤department≤d_{dept.all}),

课程维 d_{course}=({course, type, exam_type, d_{course.all} }, course≤type≤exam_type≤d_{course.all}),

事实表 M={credit_hour, mark}, 包含学分和成绩,成绩是规整化的值。

这里设置了一个积点计算值 credit_mark=credit_hour×mark(见图 1)。

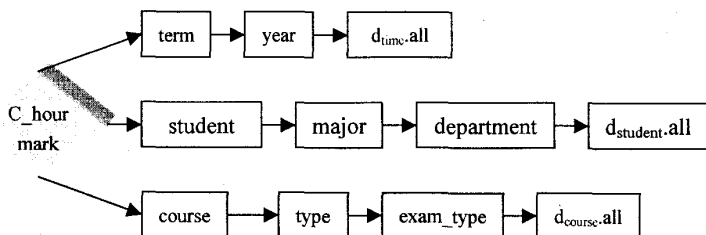


图 1 学生成绩立方体模型

由于评估维度的内容比较单一,交叉引用描述的

部分不多,这部分完全可以通过增加冗余字段来解决。采用星型数据仓库设计结构模式,比起雪花型模式有较高的效率查询响应速度,维度的更新和维护也比较容易。设计共享的维度,比如时间维度、院系单位维度等,提高维度复用性,保证了信息的一致性。

2.2 ETL 设计

庞杂的数据来源和不规范是数据仓库整合中的一大难题。数据仓库的数据来自不同院系单位,而且各种数据源的规格不同,有 Oracle, MS SQL Server, Access, Excel 数据表甚至平面文件。原始数据的格式和数据仓库的要求也有很多出入。为此,需要进行数据的析取、转换和合并等工作。

数据转换服务(DTS)通过提供一组工具,支持一个或多个数据源的导入、导出和转换。DTS 还提供包括记录软件包执行详细情况、控制事务和处理全局变量等服务,支持以编程方式创建和运行软件包,生成自定义任务和自定义转换^[5]。我们设计的 ETL 工具是基于 DTS 包的,可以将需要的数据从不同的数据源中抽取出来,根据数据仓库元数据要求进行必要的转换,加载数据到数据仓库中。下面是一个简单的 DTS 脚本实例:

```
Public Sub oCustomTask3.Trans.Sub1(ByVal oCustomTask3 As Object)
```

```
Dim oTransformation As DTS.Transformation2
```

```
Dim oTransProps As DTS.Properties
```

```
Dim oColumn As DTS.Column
```

```
Set oTransformation = oCustomTask3.Transformations.New("DTS.DataPumpTransformCopy") ' 定义转换规则,转换的基本操作为 copy
```

```
oTransformation.Name = "DTSTransformation-1"
```

```
oTransformation... ' 转换其它属性设置
```

```
Set oColumn = oTransformation.SourceColumns.New("StudentID", 1)
```

```
oColumn.Name = "StudentID"
```

```
oColumn... ' 源对象表中列的其它属性设置
```

```
oTransformation.SourceColumns.Add oColumn
```

```
Set oColumn = oTransformation.DestinationColumns.New("StudentID", 1)
```

```
oColumn.Name = "StudentID"
```

```
oColumn... ' 目标对象表中列的其它属性设置
```

```
oTransformation.DestinationColumns.Add oColumn
```

```
Set oTransProps = oTransformation.TransformServerProperties
```

```
oCustomTask3.Transformations.Add o-
```

```
Transformation ' 施加转换规则
```

```
End Sub
```

从易用性角度考虑,系统支持图形化的 ETL 流程设计,避免手工编制脚本。由程序界面进行脚本管理,用户可以定制并保存方案,提高复用性。设置排程(Scheduling)后,可以让系统定期执行脚本,进行数据的自动抽取。

2.3 分析界面设计

数据仓库中数据分析的基本操作包括钻取、切片、旋转等,这些操作有效地支持交互和可视化实现关系到软件的可用性。Analysis Services 自带了分析界面,属于胖客户端的 C/S 模式,客户端的维护成本较高。Office Web 组件(OWC)的数据透视表组件 PivotTable 用于在 Web 浏览器和传统的开发环境中建立数据分析和数据报表,支持多维数据的可视化^[6]。考虑到分析使用者分布在各个地方,采用更为流行的 B/S 模式设计分析界面,把 OWC 组件嵌入到 Web 页面中,供 Web 瘦客户端访问,这样可以避免烦琐的胖客户端配置,提高易用性和可维护性。

客户端执行的页面脚本是通过 HTTP 与 Analysis Services 下的多维数据集实现连接的。需要在 Microsoft IIS 服务器配置 Analysis Services 的请求代理。可视化分析的配置与使用步骤包括:

(1)通过以下字串,建立与数据源的连接;

```
strConnect = "Data Source = HTTP:// & myDataSource &  
";Provider = MSOLAP.2;Initial Catalog = " & myCatalog
```

(2)连接数据源,指定数据库和维表;

(3)通过拖动维度项目信息至数据透视表,展开、合并、对换维度信息等操作,进行诸如钻取、切片、旋转等分析,自动显示分析结果。

图 2 为所设计的学生成绩分析界面。

2.4 实施概要

采用 MS SQL Server2000 作为后台数据库,Analysis Services 作为分析服务器。采用基于 OWC 可视化引擎进行可视化分析界面设计。

目前在系统中定义了包括数据仓库管理员和普通分析人员这两大类用户角色。在此基础上,根据不同分析应用模块进行权限的层次细化。数据仓库管理员是专业人员,负责数据仓库的规划,数据立方体的设计以及数据的 ETL 操作等工作;普通分析人员是赋予相应层次权限的各级决策人员和部门负责人,主要利用

分析结果进行决策。

后台数据库、数据仓库和 IIS OWC 引擎配置在一个服务器上,采用紧耦合的实施方案,避免通过网络传输分析数据造成效率低下的问题。数据的 ETL 操作以及数据立方体的建立需要在该服务器上完成。普通分析人员在微机上用 Web 客户端进行访问。

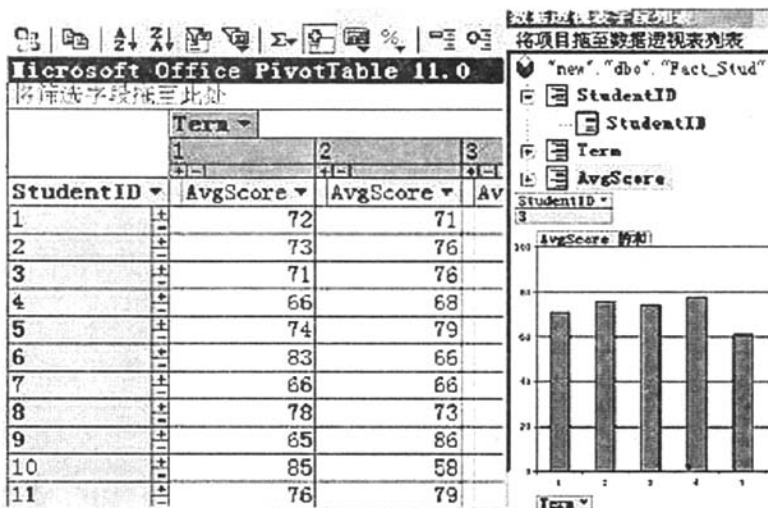


图 2 学生成绩分析界面截图

3 小 结

数据仓库实施有多种方案可选,最优的应该从实际出发,本着在满足分析需要的基础上尽量降低各种成本这个基本思路,设计了教学评估分析系统。

未来除了进一步规划面向院系以及特定部门的数据集市,考虑加入基于数据仓库的数据挖掘功能模块,进一步从教学活动数据中发现隐藏其中的模式规律。

参考文献:

- [1] 刘凤泰. 在 2005 年普通高校本科教学工作水平评估研讨班上的讲话 [EB/OL]. 2005. <http://pjb.jmu.edu.cn/2005>.
- [2] 集美大学评价办. 集美大学(系)级本科教学工作水平评估自评结果及依据 [EB/OL]. 2005. <http://pjb.jmu.edu.cn/2005>.
- [3] 罗运模. SQL Server 2000 数据仓库应用与开发 [M]. 北京:人民邮电出版社,2001.
- [4] Iseninger D. SQL Server 2000 分析服务 [M]. 北京:北京大学出版社,2001.
- [5] Thomsen E, Spofford G, Chase D. Microsoft OLAP 解决方案 [M]. 北京:人民邮电出版社,2000.
- [6] Hasan J, Tu K. OLAP: Build an OLAP Reporting App in ASP.NET Using SQL Server 2000 Analysis Services and Office XP [EB/OL]. 2003. <http://msdn.microsoft.com/2003>.