

# 基于多维数据建模的一体化平台

陆昌辉, 刘青宝, 邓 苏, 张维明

(国防科学技术大学信息系统与管理学院, 长沙 410073)

**摘要:** 基于快速原型开发思想, 设计并实现了一个一体化的多维数据建模平台, 该平台为 OLAP 应用系统设计人员提供了一个快速构建原型系统的支撑环境, 采用图示化的方法描述了用户需求, 经过简单的操作, 可以演示对应的分析结果, 便于用户修正与确认, 并为系统设计人员提供了一个导向, 加快了系统的开发速度。

**关键词:** 数据仓库; 多维数据建模; 一体化平台

## Unified Platform Based on Multi-dimensional Data Modeling

LU Chang-hui, LIU Qing-bao, DENG Su, ZHANG Wei-ming

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

**【Abstract】** Based on the idea of rapid prototyping, this paper introduces a unified platform for multidimensional data modeling. This platform provides a design environment of prototype for the OLAP applications designers, describes the users' requirements with diagram, and shows the corresponding analysis results to the users immediately after a series of simple operations. Through this way, users can adjust and affirm their analysis requirements easily. And it provides the application designers with a guide, expedites the process of applications development.

**【Key words】** data warehouse; multi-dimensional data modeling; unified platform

近几年来, 数据仓库技术被应用到决策支持系统中, 许多机构建立了相对完善的 OLTP 系统。随着时间的推移, 这些系统积累了大量的历史数据, 蕴含了许多重要的信息。

多维数据建模技术能够帮助人们在大量的数据中, 挖掘出有用的知识; 在竞争日益激烈的市场中, 迅速地作出正确的决策。借助该技术, 各种 OLAP 工具能够方便地对数据进行复杂分析和可视化处理。多维数据建模技术借鉴了分析师考虑问题的方式, 提高了系统的查询性能<sup>[1]</sup>。

对于任何产品来说, 只有满足用户的需求才能体现其使用价值。在 OLAP 系统中, 用户需求显得尤为重要。然而, 用户需求往往是难以把握的。在传统工业中, 为了准确地描述用户需求, 往往会提供一个样品来与之交互, 并取得确认。在开发 OLAP 系统时, 需要一个平台, 它可以让系统设计人员便捷地描述用户需求, 经过简单的处理, 就可以向用户演示相应的分析结果。该平台的优点为:

(1) 利于用户更加准确地描述分析需求。借助于该平台, 用户可以根据其所见到的分析结果, 完善分析需求, 促进用户与系统设计人员间的交流, 从而减少了理解上的误差。

(2) 该平台有助于获取用户的需求。由于 OLAP 系统面向的用户一般是政府部门或者企事业单位的高层管理人员, 因此他们很难专门抽出一段很长的时间来与系统设计人员进行沟通。在军队的高层管理者中, 这种状况尤为明显。因此, 一体化多维数据建模平台可以在相当短时间内非常便捷、准确地获得用户需求。

### 1 总体思路

总体流程从上到下分为 3 层(图 1), 即:

(1) 用户需求描述层。用户通过图示化标记, 快速地描述分析需求, 并根据设定的建模规则进行语法检查。语法检查通过后, 将建立的图示化模型转换成对应的 XML 文档, 为后面要进行的模型匹配与模型映射操作提供接口; 若存在语法错误, 则给出相应提示, 方便用户进行调整。

(2) 模型匹配与映射层。以描述整个数据仓库存储内容的全局元

数据为依据, 对转换后的模型进行匹配评估, 若匹配不成功, 则给出相应的提示信息, 告知用户其分析需求得不到底层数据的有力支持, 提示他进行相应的调整, 同时也对数据仓库提出新的数据需求; 若模型匹配成功, 则将该概念模型映射为对应的多维数据集。在综合指标的指引下, 从数据仓库中抽取数据进行综合, 并装载到该多维数据集, 为后面的分析结果展现提供了数据准备。

(3) 分析结果展现层。通过动态的方式(OLAP 分析)与静态的方式(报表展现)对用户描述的分析需求进行快速展现, 用户根据所见到的效果对需求进行完善与确认, 为系统分析师在进行系统设计时提供了一个导向。

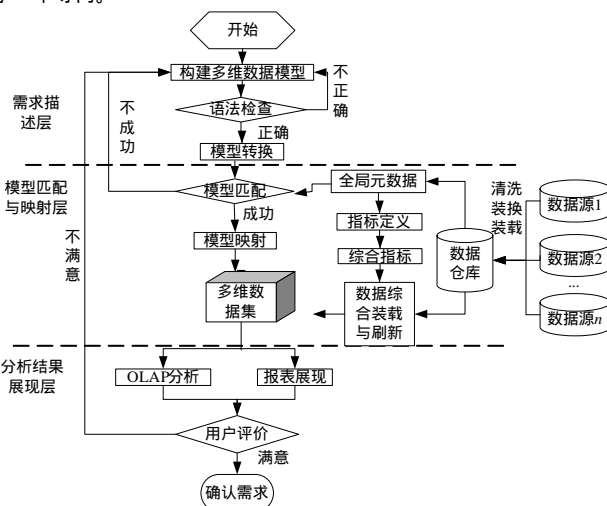


图 1 总体流程

**基金项目:** 国家自然科学基金资助项目(60172012)

**作者简介:** 陆昌辉(1976 - ), 男, 博士、讲师, 研究方向: 数据仓库, 决策支持技术; 刘青宝, 副教授; 邓 苏, 教授; 张维明, 教授、博士生导师

**收稿日期:** 2006-12-03

**E-mail:** lch829@vip.sina.com

## 2 支撑平台

根据图 1 描述的总体流程,设计并实现了该建模技术的支撑平台——MDMP,它由模型构建子系统、元数据管理系统、综合指标管理系统、模型匹配与映射子系统、数据综合与装载子系统、OLAP 分析子系统、报表展现子系统等组成,下面分别对其进行介绍。

### (1)模型构建子系统

模型构建子系统为用户提供了一个图示化的描述环境,它包含模型构建、语法检查、模型转换、逆向工程等功能模块。其中,模型构建模块就是利用一套图示化标记根据用户需求来建立多维数据模型,在具体的构建过程中,可以从顶层视图、中层视图、底层视图 3 个不同层次来对需求进行描述。为了与后面的模型匹配与映射子系统进行接口,通过模型转换模块可以将图示化的形式转换为对应的 XML 描述形式,该 XML 文档的定义格式与转换算法可以参见文献[2]。逆向工程就是将用 XML 文档描述的多维数据模型转换为对应的图示化形式。关于模型的构建方法以及该子系统的描述可以参考文献[3]。

### (2)元数据管理系统

元数据管理系统对下提供对各数据源(该数据源可以是存放数据的集中式数据仓库,也可以是提供基础业务数据的各数据源)的元数据抽取功能,对上为多维数据模型匹配与映射,以及数据综合、装载与刷新等提供元数据访问接口,对内提供元数据管理、交换和导航功能,系统结构见图 2。元数据访问接口有 2 种方式:XML 文档访问和 SQL 访问接口,其中,XML 文档的 EBNF 编码语法可以参考文献[4]。

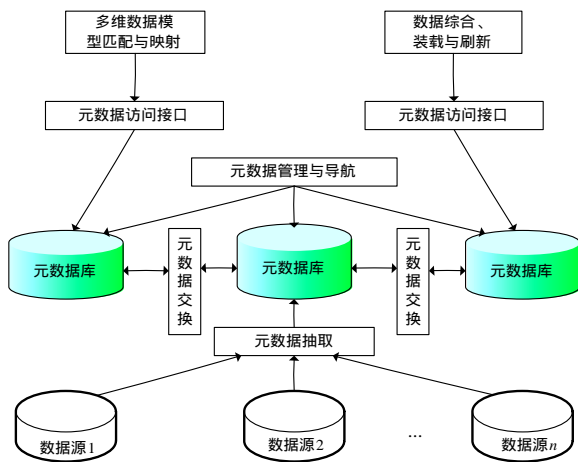


图 2 元数据管理系统功能结构

### (3)综合指标管理系统

指标是对分析内容的一种汇总和计算,指标体系将相关的指标按照一定的分类原则进行组织划分,在结构上采用线分类法<sup>[5]</sup>(层级分类法、指标树)进行划分。

综合指标管理包含指标分类的管理和指标的管理 2 个部分,它具有两重作用:

- 1)对数据仓库中所支持的分析焦点进行统一分类管理;
- 2)在对经过映射生成的多维数据集进行数据综合与装载时提供指导。

### (4)模型匹配与映射子系统

用户在构建多维数据模型时,不清楚数据仓库中存储的数据能够提供的分析支持,必须根据底层的数据支持情况(在这里体现为通过元数据管理系统抽取的全局元数据)对其

构建的多维数据模型进行匹配分析,匹配成功后,则转换为对应的多维数据集,为后面的 OLAP 分析与报表展现提供访问接口;不成功则给出相应提示信息,便于用户进行修改。这就是模型匹配与映射子系统的主要工作目的。

### (5)数据综合与装载子系统

数据综合与装载子系统在综合指标的指引下,从数据仓库中抽取相关数据进行汇总和计算,并按一定的装载策略导入到多维数据集中,为后面分析结果的展现提供数据基础。

### (6)OLAP 分析子系统

OLAP 分析子系统主要分为 4 个部分:数据源连接模块,多维数据集解析模块,图表展现模块,数据导出模块。具体见图 3。关于该子系统的设计与实现方法可以参考文献[6]。

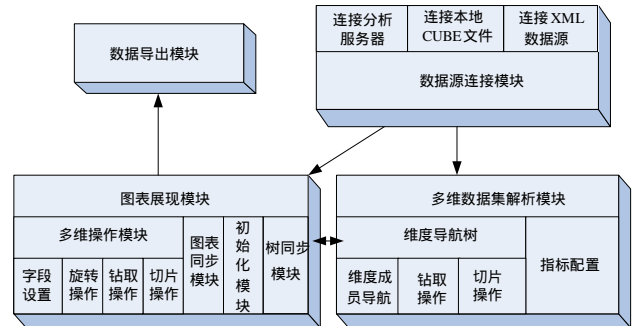


图 3 OLAP 子系统功能结构

### (7)报表展现子系统

该子系统采用汇总报表的形式对用户分析的内容进行展现,便于用户将得到的汇总结果进行打印存档。在该子系统中,用户可以灵活定制报表的样式,并用配置文件的形式对这些模版信息进行保存,免得重复定制。

## 3 应用范例

图 4 为所得模型的中层视图。

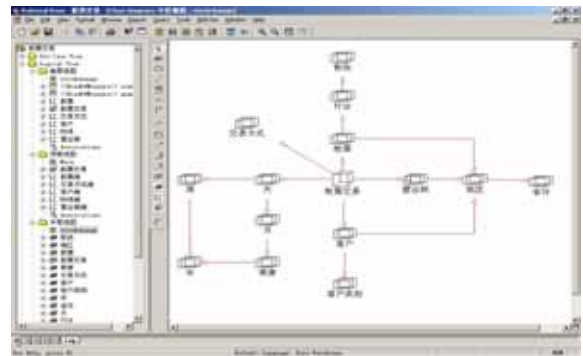


图 4 股票交易模型的中层视图

在股票交易 CRM 系统的建设过程中,可以发现证券公司的领导经常关心某一客户在什么时候、什么地方、对什么样的股票进行了交易的问题。为了迅速而准确地把握分析需求,可以运用该 MDMP 平台,用图示化的方法对其需求进行描述,具体描述方法可以参考文献[3]。

该模型经过语法检查,转换为对应的 XML 文档后,根据全局元数据的描述,进行匹配与映射,在综合指标的指引下,通过数据综合与装载子系统对映射得到的多维数据集进行数据填充,图 5 就是在 OLAP 分析子系统中展现的结果,用户根据其所见到的分析结果,进一步完善分析需求。该技术在实际工程中得到了很好的应用,明显地缩短了系统的开发周期。

(下转第 41 页)