

基于OLAP技术实现专利分析系统 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/141/2021_2022__E5_9F_BA_E4_BA_8EOLAP_c29_141404.htm

摘要：为适应全球化竞争形势的需要，本文为企业提供强大的专利分析系统，设计了一个OLAP和OLTP综合使用的方案。数据仓库采用数据集市形式。OLAP技术的使用，为企业进行技术创新和技术选择提供决策依据。引言 专利分析不但对发明人、中小企业、实验室、大学等知识创造者有极大的价值，而且对大企业、政府机构也有极大的价值，而且还可以指导企业通过自主研发部署专利，在竞争对手的核心技术网络中秘密部署杀手锏，使企业在相关领域的竞争中获得有利的位置。因此企业需要分析处理大量的专利信息(中国,美国,欧洲,日本等国家专利局的专利约1000万)。专利信息量的迅速增长以及查询的复杂化，使传统的OLTP（联机事务处理）系统不能满足对数据进行深层次多维分析的要求，于是人们提出了数据仓库和OLAP(联机分析处理)技术。数据仓库是支持管理决策过程的，面向主题的、集成的、随时间而变的、持久的数据集合，对分布在企业内部各处的OLTP数据进行抽取、净化，为企业决策分析提供所需的基础数据 [1]。OLAP则利用存储在数据仓库中的数据完成快速、一致、交互的分析，然后以直观的形式将分析结果返回给决策者。在一个针对专利进行决策分析的系统中,需要对专利信息进行多维分析处理,并将分析结果以多维视图的方式展现给决策者,企业决策者从而做出相应的决策.因要求该系统具有对大量多维专利信息进行快速分析的特点,所以在此采用了数据仓库和OLAP技术作为专利信息处理系统的核心。 1

系统设计 开发基于数据仓库的OLAP系统，一般有两种方案：一是利用著名数据库厂商现有数据仓库和OLAP产品(如:Oracle Express,DB2 OLAP等)来创建；二是利用开发工具将专利分析需求进行多维化处理。第一种方案因只是在现有成熟产品上作二次开发，所以能够创建较完善的基于数据仓库的OLAP系统，但不灵活，且费用极其昂贵，所以在此，我们采用第二种方案，采用能处理大规模数据的ROLAP技术,手动构建多维立方体,并构造自己的OLAP多维引擎，以设计适合专利分析OLAP系统。整个OLAP系统,大致分为四层(如图1所示)：第一层是数据源，存储了所有相关领域的专利信息,包括内部和外部信息. 第二层是OLAP数据仓库层。源数据库中的结构未经过系统化，其组织形式是面向具体应用的。数据抽取程序将源数据按主题进行归纳整理，存入OLAP数据库中，提供适合OLAP分析的详细、集成、准确的客户基础数据。第三层是OLAP服务器层，保存了分析所需的客户聚集数据和相关的元数据，代理用户的分析请求，获取分析数据并返回给用户。第四层是应用层，这层让用户根据模型信息，提交分析请求，然后将获得的数据按用户所需要的方式展现。图1 OLAP服务器层结构 OLAP服务器层是本系统的核心,其详细结构如图1所示,主要由OLAP立方体计算和OLAP多维引擎两大模块组成.我们将详细探讨这两方面的技术实现,并给出优化策略。 100Test 下载频道开通，各类考试题目直接下载。详细请访问 www.100test.com