

文章编号: 0253-2697(2003)04-0077-04

基于数据仓库的钻井工程智能决策支持系统研究

李琪¹ 徐英卓²

(1. 西安石油大学 陕西西安 710065; 2. 西北工业大学 陕西西安 710072)

摘要: 以实现钻井工程网络化、信息化以及管理科学化为目标,提出了利用数据仓库技术建立计算机支持的钻井工程决策智能系统,阐述了系统的结构设计、数据仓库系统的体系结构及实现方法。运用数据仓库技术和网络通信技术,针对整个钻井工程的各个环节,研究和开发了以强大的钻井工程数据仓库为支撑,可满足不同层次和部门的工程管理及技术人员决策需求的智能支持系统,以实现科学化决策。

关键词: 数据仓库;钻井工程;智能决策;支持系统;网络;元数据

中图分类号: TE249 **文献标识码:** A

Intelligent decision support system based on data warehouse for drilling engineering

LI Qi¹ XU Ying-zhuo²

(1. Xi'an Petroleum Institute, Xi'an 710065, China; 2. Northwest Industry University, Xi'an 710072, China)

Abstract: An intelligent system for drilling engineering decision was established using the data warehouse technology, which aims to implement the idea of information, automation and scientific management in drilling engineering. The structure design of the system, architecture and realization method of the data warehouse system were discussed. An intelligent support system for each phase of drilling engineering was developed by the use of data warehouse and communication technology. This system supported by powerful drilling engineering database can satisfy the requirement of decision-makers in different levels and departments, which makes it possible to implement science decision.

Key words: data warehouse; drilling engineering; intelligent decision; support system; network communication technology; meta data

现代钻井工程由钻井设计、钻前工程、钻井作业、固井作业和完井作业等多个环节组成,涉及分布在不同地域的众多职能部门。信息种类繁多、数据量大,综合利用这些分散的异构环境的数据源,及时得到准确、科学而有效的决策依据对决策人员至关重要^[1]。在传统的钻井工程管理中,信息的组织与处理缺乏集中存储和管理机制,因而缺少综合性、分析性数据,不能满足中上层管理决策人员对信息的查询、分析。为此,笔者提出建立基于数据仓库的钻井工程智能决策支持系统,简称钻井工程决策系统 DEDES。它运用数据仓库技术将分布在不同地点的数据转换成集中、统一、随时可用的信息,能支持各种类型的数据查询、分析和决策,以实现钻井工程网络化、信息化和管理科学化,从而全面提高钻井技术和综合经济效益。

1 数据仓库技术

数据仓库^[2]DW(Data Warehouse)是面向主题的、

集成的、稳定的、不同时间的数据集合,用于支持生产管理中的决策制定过程,其信息来自不同地点的数据库或其他信息源,且信息具有分布和异构的特点。数据仓库的主要功能:①从各信息源提取所需的数据,将其加工处理后存储到数据仓库中;②直接在数据仓库上处理用户的查询和决策分析请求。

1.1 数据仓库结构

数据仓库是存储数据的一种形式,它从传统的业务数据库中获得基本数据。先按决策的主题要求形成当前基本数据,再按综合决策的要求形成轻度综合数据和高度综合数据。当前基本数据随着时间的推移,由数据仓库的时间控制机制转为历史基本数据(如图1所示)。整个数据仓库由元数据来组织,元数据不包括任何业务数据中的实际数据信息。它用于定义数据结构、存储数据模型、综合算法以及从作业环境到数据仓库的规划等。

1.2 数据仓库系统

数据仓库系统是由数据仓库、数据仓库管理系统

基金项目:国家自然科学基金资助项目(No. 59974021, No. 59774006)。

作者简介:李琪,男,1963年12月生,1986年毕业于西南石油学院获硕士学位,2000年获博士学位,现为西安石油大学石油工程学院教授。

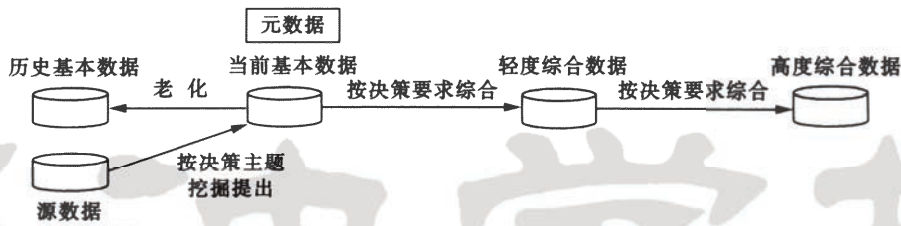


图1 数据仓库组织结构

Fig.1 The structure of data warehouse

和数据仓库工具集组成。其中数据仓库是核心；数据仓库管理系统负责整个系统的运转；数据仓库工具集是数据仓库发挥作用的关键，只有通过有效的工具数据仓库才能提供支持决策的信息。工具集分为两类^[3]：一类是查询检索工具；另一类是从大量数据中挖掘具有规律性知识的数据挖掘工具。

2 DEDS 系统结构及工作流程

DEDS是一套完整的信息管理技术体系，分为基础系统网络、数据采集、数据仓库系统和数据呈现4个部分。它们之间相互联系，共同构成层次分明的钻井工程信息分析环境(图2)。

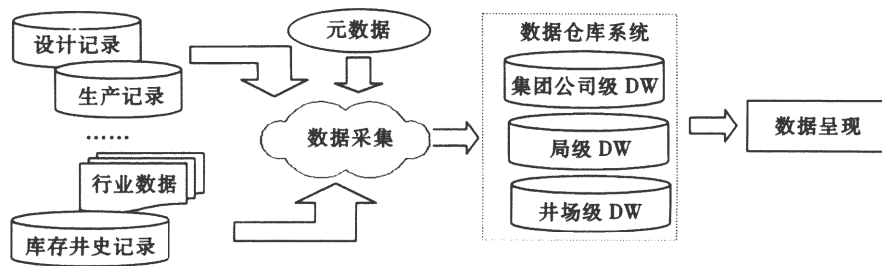


图2 DEDS系统结构

Fig.2 The structure of drilling engineering decision support system

DEDS的工作流程为：①数据采集系统在元数据的控制下采集各类工程数据，进行重整归类后存放到井场级数据仓库中，②根据汇总级别不同，在井场级数据仓库的基础上，面向不同层次的管理人员分别形成局级和集团公司级数据仓库；③通过数据呈现工具以多维视图、报表或图表形式提供给用户。

田基地的现场，各油田又受集团公司统一管理。因此，从系统的功能、联动方式以及用户级别和用户所处的工作环境考虑，将DEDS分为井场级、油田级和集团公司级的三级结构。整个网络共分集团公司级、油田(局)级和井场(前指)级3个层次(图3)。从井队现场

3 DEDS 系统设计

3.1 基础系统网络

基础系统网络是DEDS得以实现的物质技术保障，它把分布在不同地点、不同功能的作业系统连接到一个统一的物理网络中。并根据信息处理的特点，为数据仓库的建立提供完整的技术支持手段，包括网络连接、数据库及其互连、数据访问技术和客户机标准等，使得在整个集团公司范围内建立统一的Client/Server结构，从而实现跨系统的数据采集、重整和存储数据。

考虑到钻井工程的特殊性和复杂性，通常钻井工程设计是在油田基地完成。而钻井作业分布在远离油

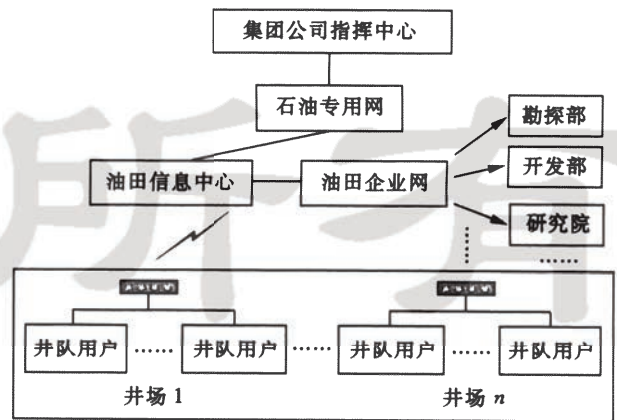


图3 基础网络结构

Fig.3 The architecture of the basic network

采集的数据通过远程传输汇总到油田信息中心，利用

油田企业网呈现给油田各职能部门,同时信息中心将传来的信息进行综合处理后再通过石油专用网上报集团公司。整个网络由油田企业网和石油专用网组成,油田企业网是以 IBM Netfinity 5500 服务器为中心的星型结构。网络操作系统采用 Windows NT,负责整个系统网络的安全运行;网络协议为 TCP/IP,不仅支持局域网内部的通信,还支持远程用户的数据库访问、文件传输和电子邮件等。采用 SQL Server7.0 作为数据仓库的数据库平台。

3.2 数据采集

DEDS中数据来源于各作业系统的数据库,这些源数据和目标数据仓库的数据因类型、含义的不一致而存在着差别,因此在进入数据仓库之前须进行转换。数据采集是为数据转换而设立的中间处理过程,具体包括数据的抽取、集成和加载。它将源数据库按数据仓库的设计要求通过综合、重构、筛选及转换等数据变换操作,生成纯净、统一的能支持决策分析的分析型数据后再存到数据仓库。

由于源数据库中常有数据的插入、修改等操作而引发数据的改变,在数据采集过程中,为了能把变化的数据及时反映到数据仓库,以保证其准确性,必须解决数据变化的捕获问题。为此,在源数据库和数据仓库之间设立一个中间存储空间,采用 SQL Server 的触发器机制^[4]。触发器把变化的数据及时发送到中间存储空间进行转换和集成,再利用 SQL Server 的复制功能更新数据仓库。

3.3 DEDS 中的数据仓库系统

3.3.1 系统结构

根据钻井工程行业的组织结构、作业内容和管理决策人员分析问题的特点,DEDS 中数据仓库系统是一个对行业数据进行重新集成、适应多维、快速查询、进行联机分析处理(OLAP)操作以及决策支持的数据采集、处理、管理和展现的体系结构,分为基础数据层、井场级数据仓库层、局级数据仓库层和集团公司级数据仓库层,其结构如图 4 所示。

基础数据层主要是各井队的现场数据,是以上各

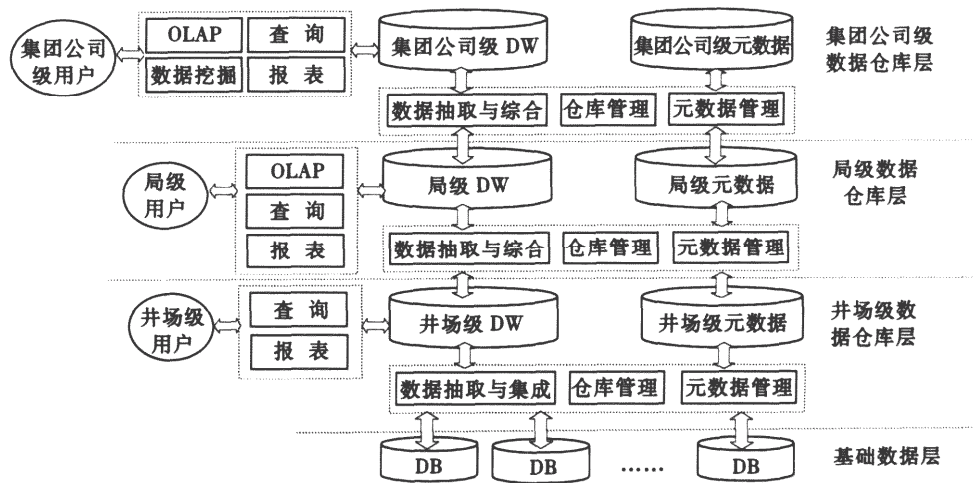


图 4 DEDS 的数据仓库系统结构

Fig.4 The architecture of the data warehouse of DEDS system

层的数据来源。井场级数据仓库层的数据直接来自各井队现场信息数据库,主要是大量的细节数据,用于支持详细数据的查询、报表和施工决策。在整个体系结构中它居核心地位,负责基础数据层的数据集成处理,以保证数据的完整、统一,确保数据仓库的质量。

局级数据仓库层是对井场级数据仓库进行轻度汇总,抽取适于局级指挥部门的数据(如:重点探井施工情况),面向油田级管理和设计人员。

集团公司级数据仓库层面向集团公司级高层管理人员,是在局级数据仓库的基础上抽取适于高层决策

分析用的数据并进行高度汇总后放入集团公司级数据仓库,同时生成相应的元数据。

每 1 层(除基础数据层外)中,均包括由下层数据到上层数据仓库的数据抽取与综合的功能模块、数据仓库和元数据、数据仓库和元数据的管理模块以及供用户从数据仓库中获取信息的工具(报表、优化查询、OLAP 和数据挖掘等)。

各级数据仓库工具是整个系统发挥作用的关键。其中联机分析处理操作工具^[5]用于提供面向分析的多维数据模型,并使用多维分析方法从多个角度、多个

侧面对多维数据进行分析、比较,使用户以更接近自然的方式分析数据。数据挖掘工具用于从数据仓库中提取意料之外或未知的、能反映大量数据内在特征和规律的有用信息与知识,例如,通过从井史资料中抽取有关钻时记录及岩石数据,可挖掘出地层压力的变化规律,从而为决策人员提供智能决策支持。

3.3.2 体系结构

数据仓库是面向主题来组织的,主题是在较高层次上将数据归类的标准,通常反映管理决策者最关心的问题。本系统从油田的实际需求出发,围绕主要的作业功能,设计了井眼稳定性、轨迹控制、地层伤害、事

故、复杂情况预测、钻头、钻井速度、钻井时效、钻井质量、消耗材料、成本等共 18 个主题。对每个主题,其数据处理都是多层次的(从井队、井场、局级到集团公司),同时也是多维的,即从不同观察角度进行的。如“钻井速度”主题,在钻井速度影响因素分析中,常按地层、岩性、钻头、钻井参数、钻井方式、水力参数、钻时七维进行数据的处理。在钻井工程作业中,不同层次、不同部门的管理人员对信息的需求均不相同,因此,将 DEDES 中的数据仓库设计成多级多维的体系结构,其总体框架是二维矩阵结构。以“钻井时效”主题为例(如图 5)。该体系结构允许用户从不同的角度、不同层

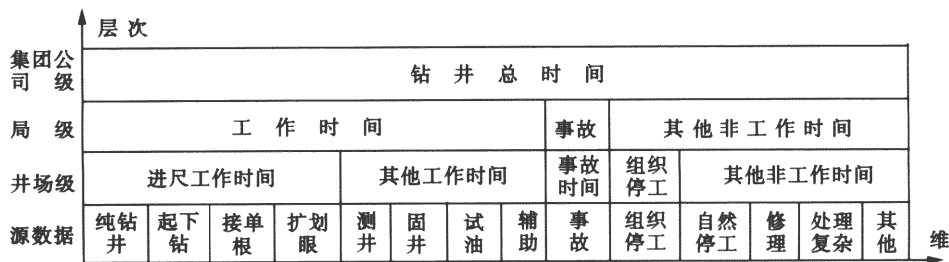


图 5 钻井时效主题数据仓库的多级多维体系结构

Fig. 5 The multi-level multi-dimensional architecture of the topic data warehouse about drilling benefit

次来考察工程数据,以便进行数据查询和分析。

3.3.3 元数据

DEDES 中的元数据是对整个数据仓库环境的描述,分为数据字典和数据处理规则,数据库设计和数据采集系统的开发均根据元数据并在其控制下进行。数据字典包括数据类型定义、表结构定义、表说明和主题的维结构定义。数据处理规则定义了数据仓库中有关数据处理的规则,包括由各井队的现场数据到数据仓库的数据映像和数据更新规则。

3.4 数据呈现

DEDES 中构造数据仓库的目的是为了满足钻井工程行业中各种用户对决策支持的需求,他们在进行不同决策和分析时都须从数据仓库中查询或检索数据。这些查询不仅指对记录级数据的查询,也包括对分析结果(发展趋势或模式总结)的查询。数据呈现是利用 OLAP、优化查询、统计分析报表以及数据挖掘等工具按照不同层次的用户查询和分析需求来处理与显示数据,它包括创建数据的多维视图供用户进行多维分析、生成各种统计报表、生成各种图形和图像等。

4 结 论

数据仓库是一种面向应用的数据管理技术,它提供了集成化的数据管理功能。数据仓库中集成的不同层次的汇总数据可供各级决策人员进行综合性的数据分析,特别是战略分析。基于数据仓库的钻井工程智能决策支持系统能有效地为石油行业各级工程管理与技术人员提供决策支持,从而为实现钻井的网络化、信息化提供途径。

参 考 文 献

- [1] 陈文伟. 智能决策支持系统[M]. 北京:电子工业出版社, 1998:150-163.
- [2] 王珊. 数据仓库与联机分析处理[M]. 北京:电子工业出版社, 1998:120-144.
- [3] 刘东波. 数据仓库技术的现状与未来[J]. 微型机与应用, 2000, 19(7):42-44.
- [4] 袁鹏飞. SQL Server 数据库应用开发技术[M]. 北京:人民邮电出版社, 1998:56-64.
- [5] 姚卿达, 张凤荔, 周明天, 等. 数据仓库和数据采掘应用研究[J]. 计算机科学, 1998, 23(6):57-59.

(收稿日期 2002-02-26 改回日期 2002-09-09 编辑 杨 苗)