

三峡工程泥沙信息系统文档数据库的设计与实现

王 东, 刘兴年, 曹叔尤

(四川大学高速水力学国家重点实验室, 四川 成都 610065)

摘要:针对三峡工程泥沙科学和技术的大量资料和文献非结构化特征,设计文档数据库来维护和查询这些数据.首先根据需求分析建立了实体-关系模型,再将其转化成在 Oracle 环境下 DBMS 支持的关系模型.在文档数据库前端应用程序开发过程中,运用了 VB 数据库最核心的对象群 ADO2.6 与 Oracle8i 对接,实现了用户通过友好的应用程序界面对 Oracle 数据库中各类文档资料的维护和查询.在多用户 C/S 系统中测试表明,数据的完整性、安全性、并发性和整个系统的稳定性都能够较好地满足用户提出的要求.

关键词:三峡水利枢纽;泥沙信息系统;文档数据库;关系数据模型;VB;Oracle

中图分类号:TV152;TP311.135.1 **文献标识码:**B

文章编号:1006-7647(2003)01-0022-03

1 工程背景

众所周知,三峡工程泥沙问题复杂,影响重大,涉及因素多.在三峡工程施工后期和蓄水初期,为尽可能发挥三峡工程的综合效益,同时也为了系统全面地总结三峡工程泥沙现有大量研究成果,四川大学高速水力学国家重点实验室受中国长江三峡工程开发总公司的委托,开发了供业主、决策者和管理人员使用的决策支持系统软件,使运行管理现代化、科学化.

系统总体逻辑结构如图 1 所示,其中文档子系统内容包括泥沙决策过程中水库泥沙设计规范、泥沙调度规则、历史资料、政策法规、泥沙专家对三峡泥沙关键问题的论述、各高校与科研院所针对三峡工程所作的数学模型、物理模型和断面资料等信息,并且这些信息主要以 Word 文件的形式存放.文档子系统主要处理非结构化的数据,故采用文档数据库

而非纯粹的高度结构化的关系数据库,其功能为对各类文档资料的维护和查询,整个系统的数据都存储于文档数据库中.本文仅针对整个系统中文档数据库的设计和实现作一点基本探讨.

2 文档数据库的设计和实现

2.1 需求分析及子系统目标

a. 能够对政策法规、历史资料、数学模型资料、物理模型资料、专家意见以及三峡工程泥沙会议纪要 6 大类的文档进行管理,包括对每一大类文档中的子类型文档中的记录添加、删除、更新操作.

b. 在 C/S 体系结构中,能够对 6 大类文档的内容按用户的输入进行查询,并要求一定的响应时间.这里假定用户并无数据库知识.

c. 有较好的用户错误捕获功能,保证数据的完整性.由于是多用户系统,还要保证数据的并发和一致,防止添加、删除、更新异常.

d. 对系统的不同用户设置不同的访问和处理权限,防止非法用户进入网络访问和破坏数据,保证数据的安全性^[1].

2.2 关系数据模型的建立

文档数据库不同于关系数据库的方面在于允许创建许多不同类型的非结构化的或任意格式的字段,但其也属于关系数据库的一种,在概念模型设计中,两者没有任何区别.根据需求分析得到的目标,在概念模型^[2,3]中提出了 8 个实体(Entity),分别是

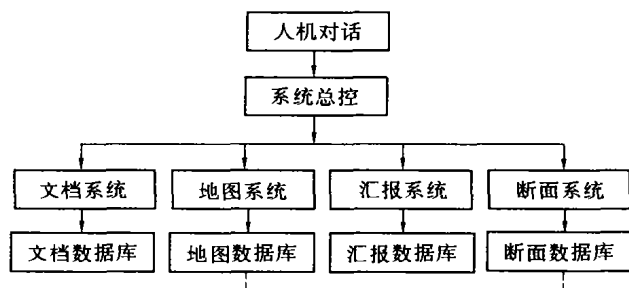


图 1 三峡工程泥沙信息系统总体逻辑结构

文档大类 (DocType)、文档子类 (DocCatalog)、政策规范 (TabRule)、历史事件 (TabHstDC)、专家意见 (TabEptViews)、会议纪要 (TabMtKeys)、物理模型 (TabPhysics) 和数学模型 (TabMath). 形成的关系是文档大类型、文档小类型对其他 6 个实体进行管理, 而其他 6 个实体之间则互相独立. 将实体-关系 (E-R) 模型转化成在 Oracle 环境下 DBMS 支持的关系模型, 其关系模型结构如下:

a. 独立对象 (即 6 个基本表):

基本表 1: 政策规范 (编号 (RuleID), 子类名称 (CatalogName), 规范名称, 发布单位, 文件内容……)

基本表 6: 专家意见 (编号 (EptViewsID), 子类名称 (CatalogName), 观点名称, 专家单位, 文件内容……)

b. 管理关系:

文档大类 (ID, TypeName, Desp, TableName)

文档子类 (ID, TypeName, 子类名称 (CatalogName))

以上关系模型中加有下划线的 ID 字段是表的主键, 保证了记录的唯一性. 该关系模型首先通过关键字段 TypeName 实现在文档大类中操作管理文档子类的功能, 而通过文档子类中的 CatalogName 字段 (划有波浪线), 可以确定需要操作的记录是属于 6 个基本表的哪一个.

整个系统采用的是 C/S 体系结构, 客户机上是前端应用程序, 在文档子系统模块中, 程序通过友好的界面, 一方面接收用户的输入, 另一方面将这种输入转换成 SQL 语句, 提交给在服务端的 Oracle 数据库管理系统. 服务端在完成对数据库的各种操作后 (加入、删除、更新、查询等), 将结果返回给数据库应用程序.

2.3 网络设计与系统硬软件

a. 网络接口. 整个系统在 VB6.0 环境下开发, 主要技术是: ①在各工作站上用 Oracle 客户端程序 Net Configuration Assistant 配置监听器和指向服务端数据库的网络服务名; ②用 VB 数据库最核心的对象群 (ActiveX Data Objects) ADO2.6 完成与 Oracle8i 对接^[4], 如图 2 所示.

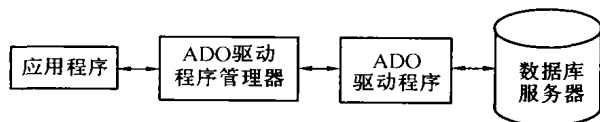


图 2 数据库接口示意图

b. 系统开发与运行的软硬件环境. 系统开发与运行硬件环境采用 Intel P4 系列微机. 开发与运行平

台: Win Server2000. 系统开发与运行软件环境为 VB6.0, 服务器安装 Oracle8i 服务端软件和 Microsoft Office 办公套件, 工作站安装 Oracle8i 客户端软件.

2.4 文档数据库的实现技术

文档子系统开发完成后的功能为: ①实现文档子类和各类文档记录的维护, 即添加、删除、更新、保存; ②实现文档的浏览、查询、显示.

文档数据库的主要实现技术:

a. 由于基本表操作的每一条记录都属于相应的文档大类和文档子类, 故在添加、删除、更新、保存、查询记录时存在一个系统对当前记录的识别与分类问题. 解决办法为: 在各个基本表的表结构设计中增加了 CatalogName 字段以识别文档子类, 文档大类由 VB 的控件 TabStrip 识别 (见图 3); 再利用 Adodc 控件的 Rcordset 对象, 在应用程序执行过程中, 根据用户在客户端的输入和操作动态绑定各个基本表或 SQL 查询语句, 从而实现对记录的操作.

在服务端数据库中, 由于每一条记录中都可能包含 Word 文档, 而一般的 Oracle 字段类型的容量不足以存放这些内容, 故在各个基本表中设置了一个 Blob 字段, 用于专门存放 Word 文档内容; 在客户端应用程序中, 用到了一项数据压缩、解压缩技术, 即操作 Word 文档数据的读出和写入 Blob 字段时, 都以二进制的方式来处理. 譬如在浏览记录的过程中, 当用户按下显示按钮时, 应用程序调用解压模块, 将当前记录 Blob 字段中 Word 文档从服务端数据库解压到客户端的 Temp 文档里, 再用 API 函数 ShellExecute 打开, 实现对文档的浏览. 当当前记录移动后, 新的 Temp 的文档将覆盖原有的内容.

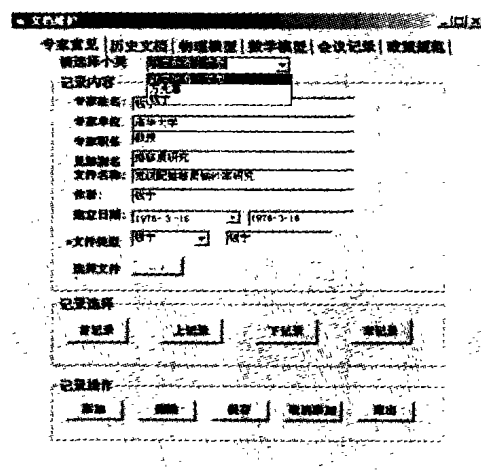


图 3 记录维护界面

b. 在数据库查询模块中, 因为各个基本表中的字段类型各异 (一般有 Varchar2, Number, Date, Blob 等), 所以最后的 SQL 查询语句中的的关系运算符也必然不一样. 为了确保用户在选定需要查询的字段