

基于商业智能的银行分析型 CRM 数据仓库的设计与实现

Bank-based business intelligence analytical CRM data warehouse design and implementation

郑华

ZHENG Hua

(广西工商职业技术学院, 南宁 530003)

摘要: 根据银行企业分析型 CRM 数据仓库的应用需求, 详细论述了 CRM 数据仓库的设计思路和设计过程, 深入研究了数据仓库的组织与结构设计, 并结合企业的实际情况, 对多维数据模型进行了设计。为下一步 CRM 系统的实施奠定了坚实的基础。

关键词: CRM; 商业智能; 设计

中图分类号: TP242

文献标识码: B

文章编号: 1009-0134(2010)01-0136-04

0 引言

中国的银行业在发展的过程中, 已逐步积累了大量的客户数据和经营数据, 利用这些数据, 发掘有价值的信息, 已经成为银行业普遍关心的问题, 而解决问题的关键就是建立银行企业级的数据仓库, 在数据仓库的基础上, 施行客户关系管理, 以满足银行客户分析需要和管理决策需要, 原有的客户关系系统作为一种管理系统难以做到这一点, 而商业智能系统则可以很好的解决这个问题, 它使我们能够将数据转变为有效的信息和知识。

1 商业智能技术

1.1 商业智能产生的背景

如今, 在金融行业大变革之后, 一系列的新问题让我们不得不关注。例如, 如何才能分析上市公司的财务情况并规避财务风险? 如何才能让银行、信托机构更好地分析、评价客户的信用程度和还款能力? 如何大规模地分析用户的消费行为? 这时商业智能应运而生。商业智能可以为企业或机构提供大规模数据联机处理、数据挖掘、数据分析, 以及报表展现等服务, 通过对大量的特定数据进行分析、整理, 从而形成量化的决策信息, 以供企业决策者采用。

1.2 商业智能的定义

商业智能这一术语 1989 年由 Gartner Group 的

Howard Dresner 首次提出, 它描述了一系列的概念和方法, 通过应用基于事实的支持系统来辅助商业决策的制定。商业智能系统从企业运作的日常数据中开发出结论性的、基于事实的和具有可实施性的信息, 使企业能够更快更容易的做出更好的商业决策, 使企业管理者和决策者以一种更清晰的角度看待业务数据, 提高企业运转效率、增加利润并建立良好的客户关系, 以最短的时间发现商业机会捕捉商业机遇。

1.3 商业智能的核心技术

商业智能的技术体系主要有数据仓库 (DW)、联机分析处理 (OLAP) 以及数据挖掘 (DM) 三部分组成。

数据仓库 (Data Warehouse) 是支持管理决策过程的、面向主题的、集成的、动态的、连续的数据集合。可以从容量庞大的业务处理型数据库中抽取数据, 清理、转换为新的存储格式。

联机分析处理 (OLAP) 技术则帮助分析人员、管理人员从多种角度把从原始数据中转化出来、能够真正为用户所理解的、并真实反映数据维特性的信息, 进行快速、一致、交互地访问, 从而获得对数据的更深入了解的一类软件技术。

数据挖掘 (Data Mining) 技术以数据仓库和联机分析处理 (OLAP) 为平台, 借助企业拥有的大量数据, 通过清洗、转换、装载等数据处理方法, 发

收稿日期: 2009-08-21

作者简介: 郑华 (1977 -), 女, 讲师, 工程硕士, 研究方向为数据挖掘和平面设计。

现大量资料间的关联与趋势，探寻一种独特的、通过其他方法发现不了的业务规律和模式。

2 银行分析型CRM数据仓库体系结构

数据仓库的建设，是以现有企业业务系统和大量业务数据的积累为基础。数据仓库不是静态的概念，只有把信息及时交给需要这些信息的使用者，供它们做出改善其业务经营的决策，信息才能发挥作用，信息才有意义。而把信息加以整理归纳和重组，并及时提供给相应的管理决策人员，是数据仓库的根本任务。因此，从产业界的角度看，数据仓库建设是一个工程，是一个过程。

整个数据仓库系统是一个包含四个层次的体系结构，具体如图1所示。

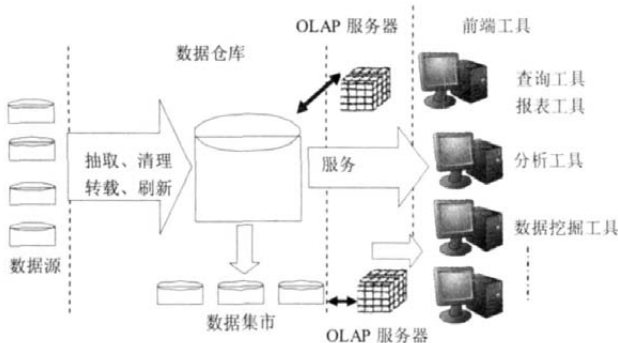


图1 体系结构

数据源：是数据仓库系统的基础，通常包括企业内部信息和外部信息。内部信息包括存放于关系型数据库(RDBMS)中的各种业务处理数据和各类文档数据，外部信息包括各类市场信息、竞争对手信息和各种手工收集的信息等等。

数据的存储与管理：是整个数据仓库系统的核心。数据仓库的真正关键是数据的存储和管理。针对现有各业务系统的数据，进行抽取、清理，并有效集成，按照主题进行组织。

OLAP (Online Analysis Process)服务器：对分析需要的数据进行有效集成，按多维模型予以组织，以便进行多角度、多层次的分析，并发现趋势。

前端工具：主要包括各种报表工具、查询工具、数据分析工具、数据挖掘工具以及各种基于数据仓库或数据集市的应用开发工具。其中数据分析工具主要针对OLAP服务器，报表工具、数据挖掘工具主要针对数据仓库。

3 系统分析

在银行各业务系统的基础上建立两层数据仓库

(DB-DW)体系结构，利用数据挖掘等商业智能技术可以实现客户细分、客户保持、客户最大价值等分析。这样，银行的管理者就可以对客户的动向进行预测，从而调整产品策略，制定正确的决策。但按照发展的需求，DB-DW 两层结构不能涵盖银行所有的数据处理要求，使用效果不能令人满意。

没有支持战术决策的数据结构基础，难以实现实时分析和挖掘应用。

企业数据集成接口复杂，难以满足企业 OLTP 系统的需要。

难以承受大量的分析查询。

在银行网络化、智能化应用领域，最重要的就是系统功能的扩展，随着生产方向的调整，系统也应该随之进行相应的升级扩展。因此，两层体系结构的应用不适合应用于网络智能化银行各业务的实际生产过程。

针对传统三层B/S结构和两层C/S的不足，本系统提出一种采用Windows环境下，可以满足网络化集成的系统开发、系统维护和跨平台等方面特殊要求的五层B/S模式开发。如图2所示：

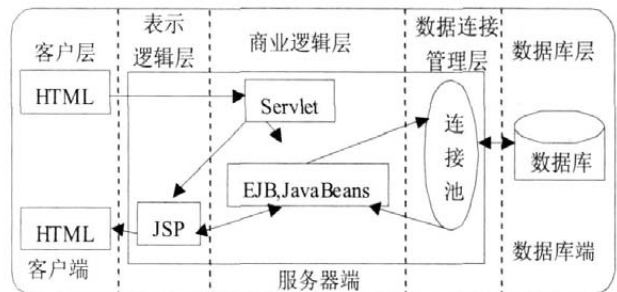


图2 五层B/S模式

该模式采用JBuilder9.0和数据连接池技术，将传统三层B/S结构的第二层（服务器端）划分为表示逻辑层、商业逻辑层和数据连接管理层。这样，CRM系统的服务请求及响应实现过程为：客户端的浏览器通过超文本链接标记语言(Hyper Text Markup Language, HTML)向应用服务器发出请求；服务器端商业逻辑层中的Servlet对请求进行分派，调用EJB (enterprise Java beans), JavaBeans 组件进行商业逻辑处理，涉及到数据库操作时，从数据连接管理层中的连接池中取出一个数据连接，使用Java数据库连接(Java DataBase Comectility, JDBC)技术访问数据库取得所需数据后，将数据连接放回连接池，以释放所占用的资源。最后，Servlet调用表示逻辑层的结果处理JSP页面对获得的数据进行格式化，形成

最终HTML页面并发往客户端,以完成对该次服务请求的响应。服务器端运用SQL Server 2000数据库,完成相关业务的后台处理。

4 数据仓库的组织与结构设计

4.1 选取业务处理

设计的第一步是通过对业务需求的理解与对可用数据的理解组合起来而确定建模的业务处理内容,如图3所示。业务处理过程并不是指业务部门或者职能,如果建立的维度模型是同部门捆绑在一起的,就无法避免出现具有不同标记与术语的数据拷贝的可能性。多重数据流向单独的数据模型,会使用户在应付不一致性的问题方面显得很脆弱。

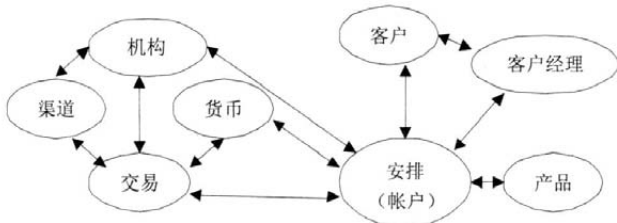


图3 银行业务关系

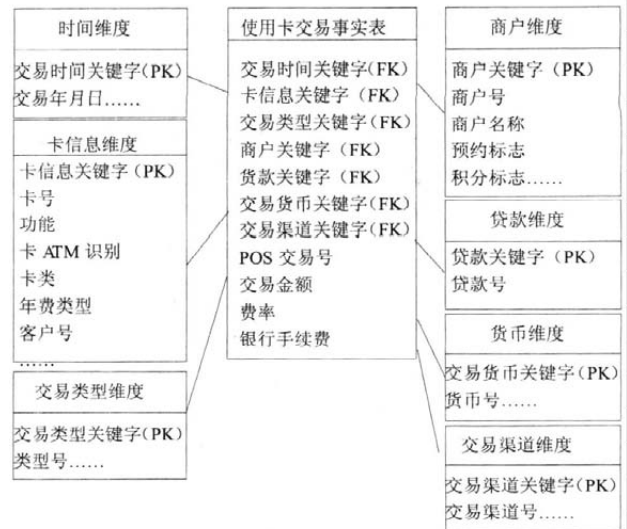
4.2 维表的设计

目前银行账务管理都采用以账户为核算的单位,同一客户在不同的网点拥有账户;同一客户在同一网点拥有不同产品的账户;甚或同一客户在同一网点、相同的产品目录中拥有不同的账户。客户同银行打交道会选择不同的账户,因此在分析系统中账户信息应当设置成为一个维表。同一客户拥有多个账户,并且客户的大部分信息在一定的时间段中保持不便,设计客户信息维表来统帅客户的账户信息,通过客户号可以归约得出客户在该分行的存款、贷款等信息。

客户交易可以选择卡、存折等不同的载体;可以选择通过网点柜台、ATM、POS、网上银行、电话银行等不同的渠道;目前银行所提供的全国联网实时交易系统,客户可以在全国不同的地区进行存取款、消费等交易。这些不同显示出客户不同的交易行为、习惯,也反映出客户对金融服务的不同需求。设计维表时,交易载体、交易渠道、交易地点应当首要考虑。

将客户信息作为一个维表设计,主要依据客户的信息在某一个时期内保持不变,而将交易信息作为事实表。客户信息维表包含客户的所有属性,如

出生年月、性别、学历等等。如图4所示。



```
public class DBConn
{
    public static synchronized Connection
    getConnection() throws Exception{
        try{
            Context initCtx=new javax.naming.InitialContext
            ();
            Context envCtx=(Context)initCtx.lookup("java:
            comp/env");
            DataSource ds=(DataSource)envCtx.lookup
            ("jdbc/mingri");
            return ds.getConnection();
        }
        catch(SQLException e){
            throw e;
        }
        catch(NamingException e){
            throw e;
        }
    }
}

public ResultSet getResult(String sql){
    try{
        Statement stmt=con.createStatement();
        ResultSet rs=stmt.executeQuery(sql);
        return rs;
    }
}
```

```
catch(Exception e){}
return null;
}
```

4.3 数据装载

将多种银行业务系统数据操作型环境数据移入数据库必须作出许多决策,其中首要的一个关键技术问题是决定从操作型环境中选取哪些数据移入 Oracle 数据库。我们须全面考虑具体需求,对一组操作型系统进行分析并决定哪些拥有对决策支持过程有用的数据。一旦选定了源数据,我们就要着手考虑如何将它们移入 Oracle 数据库中。

利用 Oracle 企业管理器(OEM)部件中的 SQL*Loader(SQL*装载器)可以将平面文件中的数据移入 Oracle 数据库中。SQL*Loader 的输入通常是一个在 Oracle 或非 Oracle 的操作型环境下产生的 ASCII 文本平面文件。当向 Oracle 转载数据时,由一个控制文件控制 SQL*Loader 如何将输入文本映射至 Oracle 数据库表的一列或多列中。

激活 SQL*Loader 有两种方法,一是用一系列命令行参数控制会话行为:

```
Sqldr80 userid=/control=regional bad=regbad.dat
discard=regdsc.dat
```

第二种是用一个命令行参数的最小集,并同时在控制文件中说明会话的参数,两种方法实质上是一致的:

```
Sqldr80 userid=/control=mgional
infile 'regional.dat' badfile regbad.dat discardfile
regdsc.dat
```

使用以上操作方式有两个关键性要点:

1) 源数据库的表的属主不必与目标数据库的表的属主相同,数据可从一个用户的表转移到另一个用户的表中。

2) 当导入表时,如果这些表事先不存在,可在它们的数据被导入之前利用在导出文件中的代码产生。

3) 导出/导入有三种操作方法:

(1) 交互式对话:使导出/导入进入一种对话状态,操作者可询问一些问题并接受回答。

(2) 命令行一参数:激活在命令行中带有参数及其值的导出得入。

(3) 参数文件:激活导出得入并为关键字 parfile 提供一个值,该值是一个文件名,在这个指定的文

件中存放着导出/导入操作需读取的关键字和它们的值。

导出命令为:

```
exp80{parameter1=value1, parameter2=value2,
..... parameters=values}
```

导入命令为:

```
imp80{parameter1=value1, parameter2=value2,
..... parametern=values}
```

5 结束语

根据银行企业分析型 CRM 数据库的应用需求,深入研究了数据库的组织与结构设计,并结合企业的实际情况,对数据模型进行了设计。为下一步 CRM 系统的实施奠定了坚实的基础。

参考文献:

- [1] 李启炎.企业商业智能教程[M].上海:同济大学出版社,2007,10-23.
- [2] 朱德利.SQL Server 2005 数据挖掘与商业智能完全解决方案[M].北京:电子工业出版社,2007,46-56
- [3] 郭宁,郑小玲.管理信息系统[M].北京:人民邮电出版社,2006,100-110.
- [4] 吴志军.客户关系管理中的数据挖掘[J].商场现代化,2005,(20):15-30.
- [5] 曼蒂,桑斯维特,等.数据库工具箱:面向 SQL Server 2005 和 Microsofr 商业智能工具集[M].北京:清华大学出版社,2007,57-59.
- [6] 杨路明,等.客户关系管理理论与实务[M].北京:电子工业出版社,2004,89-95.
- [7] 张云涛,龚玲.解析 BIRT-Eclipse 商业智能和报表工具[M].北京:电子工业出版社,2007,215-220.
- [8] 朱顺泉.数据挖掘技术在商业领域中的应用[J].中国管理信息化,2006,(1):8-9.
- [9] 李东.决策支持系统与知识管理系统[M].北京:中国人民大学出版社,2005,120-130.
- [10] 陈文伟.数据库与数据挖掘技术[M].北京:清华大学出版社,2006,69-78.
- [11] Tang,z.h.,MaccLennan,等.数据挖掘原理与应用:SQL Server 2005 数据库[M].北京:清华大学出版社,2007,215-2.