

基于数据仓库技术的客户信用管理系统的设计与实现

Customer Credit Administration System Design and Realization Based on the Data Warehouse Technology

霍 红 (天津工业大学 天津 300160)

摘要:企业中的数据存在于各个异构的传统数据库系统中,数据仓库是企业整合数据的一种有效途径。数据仓库使得企业的信息变得易于获取,并且可为企业决策提供可靠依据。本文概述了基于数据仓库技术的客户信用管理系统的设计思想、主要功能、实现平台及其主要技术实现方法和手段。

关键词:分析模型 数据仓库 数据挖掘

1 引言

数据仓库(Data Warehouse DW)能够把日常业务处理中所搜集到的各种信息进行集成,从事物的历史 and 发展的角度来组织和存储数据,供用户进行数据分析,并支持管理决策。客户信用管理系统的解决方案的基础就是基于客户为中心的数据仓库中的详细数据,借助集中管理详细的客户信息,来分析各种数据之间的错综复杂的关系,并通过数据挖掘技术透过无序的、表层的客户信息挖掘内在知识和规律,对未发生行为做出结果预测,为企业的综合经营决策提供依据,从而降低企业信用风险。

2 系统的设计与实现

2.1 客户信用数据仓库的建立

数据仓库是面向主题的、集成的、不断更新的、随时间不断变化的数据集合。并能对海量复杂、分散的客户信息进行抽取、清洁和转换,建立一个整合的、标准化、结构化的数据模型,形成全面、一致和面向决策的数据。该系统有信用申请处理、客户档案管理、客户信用评级、客户信用额度评审、帐龄分析、信用决策六大模块,每个模块处理并存储着既独立又有关联的管理数据。运用数据仓库的理论和方法,对这些管理数据进行综合,提取并存储。根据数据仓库体系结构,实现客户信用数据仓库的建立有以下几个步骤。

(1) 数据抽取

数据仓库需要通过抽取过程将数据从联机事物处

理系统、外部数据源、联机的数据存储介质中导入数据仓库。技术上主要涉及互联、复制、增量、转换、调度和监控等。数据抽取可以定时进行,但多个抽取操作执行的时间,相互顺序对数据仓库中信息的有效性十分重要。**SM** 数据抽取,即从相关数据库历史数据源抽取数据,定期复制数据存入数据仓库。开发数据仓库的数据抽取应用程序。

(2) 建立数据仓库

建立四大主题数据库,包括:客户信息主题数据库(包括客户经营管理信息、财务信息、宏观经济信息等主题数据库)。客户分类主题数据库(包括客户信用评级、信用额度等信息,是企业赊账管理的基础)。客户帐龄主题数据库(采集客户应收账款清单,计算信用风险指数)。销售主题数据库(客户交易的历史数据,如订单、合同中的数据处理过程)。

(3) 建立数据集市

从最终的用户出发,形成高度专门化的面向部门和最终用户的小型数据仓库。采用元数据方法,定义和存储目录和数据抽取规则,数据仍存储在数据仓库中。用户通过访问数据集市存储的数据目录及数据抽取规则,达到访问数据的目的。

(4) 建立用户访问界面

进行数据集市的 Web 页面设计。编制通过 Web 页面动态连接各主题数据库或数据集市小型数据仓库的应用程序。

2.2 系统功能设计

(1) 信用申请处理子系统。主要是流水式地处理客户的信用请求,内容包括打印标准的客户信用申请表、审查客户信用登记表、评价客户信用申请、记录审批结果、回复客户、记录客户争议、向信用优良客户推销等。此外,子系统可根据客户的种类和不同的级别,制作不同的客户信用申请表,便于客户信息的收集和客户的分类管理,并且在未来可能发生的委托追帐时,作为重要证据使用。信用申请处理系统的建立,目的主要有两个,一是使授信工作系统化;二是及时回复客户的书面请求,对外塑造企业的管理水平高的形象。

(2) 客户档案管理子系统。客户档案是客户信用信息管理系统的重要部分。包括客户档案版本的兼容性、客户信息欠缺提示、对非量化的指南、对于数据动态化的提示、客户企业的行业和规模的国际分类、相关变量加权指南、对于流行报告常见编码系统的识别、根据即期销售合同额度对客户分类、客户财务数据分析、客户财务状况趋势分析、行业平均值方差分析等。此外,子系统还包括功能强大的数据检索和数据维护功能,可根据档案编号、行业分类(SIC)、客户分类、客户名称、客户所在地理位置、客户规模、信用评级、建档日期等进行检索。在此系统中可提供数据存储、信用评级体系、合乎国际惯例的检索等通用程序工具箱。由于此子系统面对的是决策支持层,所以提供简单的预测曲线和客户资信打分方式显示。

(3) 客户信用评级子系统。对客户企业进行资信评级是赊账管理的一个重要环节,是客户授信工作的基础。此子系统要求把企业归入一个行业,并依系统的提示输入行业代码。然后,系统自动采集客户数据仓库中的财务及交易的历史数据,并直接计算出该客户的信用风险指数,给出企业的信用等级。

(4) 客户信用额度评审子系统。信用额度是一种事先的预防控制措施,系统根据客户的信息数据计算出其相应的信用额度,当客户申请在信用额度之内,管理人员即可作出决策;当客户申请高于信用额度时,企业要对该客户进行进一步的分析,然后作出决策。本系统包括两部分内容,整体信用额度和具体客户的信用额度。

(5) 帐龄分析子系统。帐龄分析结果是企业信用部门应收账款和安排催帐工作的依据。可输出以下报

表:客户欠款清单、大额应收账款清单、有争议项目清单、客户付款记录、预警期客户名单、逾期应收账款清单、帐龄表。用户可在任何一个子系统中调用其中的表格。

(6) 信用决策模块。在该子系统中,通过调用以上各个子系统的数据,实现信用统计。包括催帐名单、信用额度利用率名单、客户财务比率分析、客户付款行为分析、应收账款质量分析等模块。各模块功能如下:

催帐名单:系统综合分析信用额度和帐龄分析等因素,列出催帐顺序名单,按照最应追讨的顺序排序。

信用额度利用率名单:对客户利用其信用额度的情况进行分析,按照客户利用率高低进行排序,根据顺序考虑改变对一个客户的授信额度。

客户财务比率分析:根据客户信息数据,计算客户财务比率,用作授信的参考依据。

客户付款行为分析:根据客户的付款记录,对客户进行分类,并对客户未来付款作出预测。

应收账款质量分析:在一个客户名下,分析该客户对应的应收账款质量。

2.3 系统的实现

系统以 Microsoft 公司的 SQL Server 2000 作为开发环境。

SQL 关系数据库——提供数据仓库的创建和维护;

数据传输服务 DTS——提供数据输入/输出功能,在传输过程中完成数据验证、清洗和转换操作;

Analysis Services——支持在线分析处理 OLAP、数据挖掘;

Print Table Services——提供客户端 OLAP 访问功能,开发人员可用编程语言开发前端数据展示程序;

MMC(Microsoft Management Console)——提供日程安排、存储管理、性能鉴别、报警、通知等服务;

扩充服务 MDC——支持对多维数据对象的查询和操作。

3 主要的技术手段的实现

3.1 客户信用额度分析模型

客户信用额度的计算是客户信用额度评审子系统的关键。系统根据客户的信息数据计算出其相应的信用额度,对客户的信用等级做出准确评价。

信用额度的计算方法:利用营运资产分析模型对不同风险的客户确定其信用额度。

第一步计算营运资产规模:营运资产=(营运资本+净资产)/2 其中

营运资本=流动资产-流动负债

第二步衡量营运资产质量:流动比率=流动资产/流动负债 ①式

速度比率=(流动资产-存货)/流动负债 ②式

短期债务净资产比率=流动负债/净资产 ③式

债务净资产比率=负债总额/净资产 ④式

第三步计算评估值:①式+②式-③式-④式

第四步计算信用额度:信用额度=营运资产×经验型百分比率

经验型百分比率根据评估值给出。

赊销额度的确定:将营运资产分析模型与特征分析模型结合起来确定赊销额度。

3.2 客户信用等级评价分析模型

对客户进行企业资信评级是赊账管理的一个重要环节,是客户授信工作的基础。根据客户的财务及交易的历史数据,计算出该客户的信用风险指数,并给出企业的信用等级。信用等级按三等九级制来划分,即AAA级、AA级、A级、BBB级、BB级、B级、CCC级、CC级、C级。信用等级评价模型采用模糊综合评判法。

(1) 选取因素集。从大量的历史数据观测样本中挖掘出若干评价因子特征,构成因素集,即 $U = \{U_1, U_2, \dots, U_n\}$ 。 U_1 为流动比例, U_2 为速动比率, ..., U_n 为赊购付款履约情况。

(2) 确定备择的评语集, $V = \{U_1, U_2, \dots, U_n\}$, 按照国际惯例,常用 AAA, AA, A, BBB, BB, B, CCC, CC, C 三等九个级别来划分企业的信用等级,故取 $V = \{AAA, AA, A, BBB, BB, B, CCC, CC, C\}$ 。

(3) 确定评判矩阵 R 。寻找因素集中各元素对备择集中各元素的隶属关系,建立隶属函数。单个因素构成一个模糊评判向量: $R = (r_1, r_2, \dots, r_n)$, 所有单个因素的模糊评判向量构成了因素模糊评判矩阵: $R = (r_{ij})_{i=1,2,\dots,m; j=1,2,\dots,n}$ 。其中 r_{ij} 表示对第 i 个因素作出第 j 级评语的隶属度。

(4) 确定各因素的加权系数。加权系数 $A = (a_1,$

$a_2, \dots, a_n)$ 。

设 a_i 表示第 i 位专家确定的第 i 个因素的权重,则因素 U_i 的权重系数为:

$$b_i = \sum_{j=1}^n a_{ij} / n \quad \text{归一化后为: } a_i = b_i / \sum_{i=1}^n b_i$$

(5) 计算综合评价向量及综合评价值

$$P = AR = (p_1, p_2, \dots, p_9) \quad \text{其中: } p_i = \sum_{j=1}^n a_j r_{ji}$$

为了使综合结果的优劣程度更易于区分,引进分数集 $F = (F_1, F_2, \dots, F_9)^T = (90, 80, 70, \dots, 10)^T$ 其中 F_j 表示 J 级信用等级的分数。因此,客户信用等级综合评价值为: $Z = PF = \sum_{j=1}^9 p_j f_j$ 求得客户信用等级综合评价后,再对照表 2 找出相应的等级评语,即为信用等级最终评定结果。

表 2 客户信用等级评语集

等级评语	AAA	AA	A	BBB	BB	B	C
综合评价值	(90,100)	(80,90)	(70,80)	(60,70)	(50,60)	(40,50)	... (10,20)

4 结束语

在以客户为中心的时代,数据仓库有近乎实时的查询反应和强大的数据分析能力。基于数据仓库技术的客户信用管理系统不仅满足了实时性的要求,而且具备强大的决策分析能力。一个有质量、有广度、有深度的客户信息分析系统可以有效地帮助企业识别潜在的信用风险,为企业及时有效的决策支持,从而实现企业信用风险的规避。

参考文献

- 1 夏敏仁、林汉川,企业信用评级[M],上海:上海财经大学出版社,2006.
- 2 王方华、陈洁,数据库营销[M],上海:上海交通大学出版社,2006.
- 3 朱荣恩、丁浩梁,企业信用管理[M],北京:中国时代经济出版社,2005.
- 4 陈京民,数据仓库原理、设计 and 应用[M],北京:北京水利水电出版社,2004.