

数据仓库技术在 CRM 领域的应用

Research of Data Warehouse in CRM

方 力 蒋 杰 齐治昌

(国防科技大学计算机科学与技术学院研究生队 长沙410073)

Abstract This paper introduces the knowledge of CRM and data warehouse. It also makes some research for the related pivotal technologies.

Keywords Customer Relationship Management (CRM), Data Warehouse, On-line Analytical Processing (OLAP)

一、客户关系管理的基本概念

客户关系管理 CRM (Customer Relationship Management) 是目前业界非常流行的术语。作为管理企业的上帝——客户的系统,客户关系管理系统已成为目前全球最炙手可热的市场之一。

总的说来,CRM 是一种旨在改善企业与客户之间关系的新型管理机制,它实施于企业的市场营销、销售、服务、技术支持等与客户有关的领域,其目标是一方面通过提供更快速和周到的优质服务吸引和保持更多的客户;另一方面通过对业务流程的全面管理降低企业的成本。它的出现使企业能够真正全面观察其外部的客户,从而提高企业的利润。

CRM 既是一种概念,也是一套管理软件和技术。利用 CRM 系统可以实现客户服务的协同管理。一旦一个客户进入企业的视线,有关他的各种信息就进入到 CRM 系统内部,企业能搜集、追踪分析每一个客户的信息并且观察和分析客户行为对企业利益的影响,使企业与客户的关系及企业赢利得到最优化。

由此可见,CRM 软件能较好地实现客户智能功能,拥有强健的信息处理和商业分析能力,能够横跨多个应用软件模块和业务单位,对客户进行全面和智能的分析。例如,它能通过各个业务环节收益情况的分析,找出在哪些客户那里最能为企业赚钱;它也能分析客户行为与企业盈利之间的关系,帮助决策者做出明智和及时的商业决定。

要达到以上目标,就要求 CRM 系统应具有一个强大的信息决策系统。该信息决策系统可以及时、准确地收集和分析市场、销售、服务和整个企业的各类信

息。同时,通过获取并分析与客户所有的交往历史,从整个企业的角度认识客户,并为企业的决策者提供直观、准确的分析结果,达到全局性销售预测目的,从而增加获利能力。这一要求最终导致了数据仓库技术的引入。

二、数据仓库知识简介

2.1 数据仓库基本概念

数据仓库概念的形成是被誉为数据仓库之父、现任 PineCone System 公司总裁的 W. H. Inmon 在 1993 年于《建立数据仓库 (Building the Data Warehouse)》一书中正式提出的。他为数据仓库下的定义是“面向主题的、集成的、稳定的和随时间变化的数据集合,主要用于决策制定过程”。这个定义概括了数据仓库最主要的四个特征:

1) 数据仓库是面向主题的 主题是一个较高层次的数据归类标准,它是与传统数据库面向应用相对应的,每一个主题基本对应一个宏观的分析领域。比如,一个保险公司的数据仓库所组织的主题可能为:客户、政策、保险金、索赔等。而按照应用来组织则可能为:汽车保险、生命保险、健康保险、伤亡保险等。可以看出,基于主题组织的数据被划分为各自独立的领域,每个领域有自己的逻辑内涵互不交叉。而基于应用的数据组织则完全不同,它的数据只是为处理具体应用而组织在一起的。而应用是客观世界既定的,它对于数据内容的划分未必适用于分析所需。

2) 数据仓库是集成的 通过比较,可以发现原始数据与分析型数据之间的区别如图1所示。

方 力 硕士研究生,研究方向为软件工程等。 蒋 杰 硕士研究生,研究方向为计算机网络,齐治昌 教授,博士生导师,主要研究方向为软件理论与软件工程等。

操作型数据	分析型数据
细节的	综合的或提炼的
在存取瞬间是准确的	代表过去的数据
可更新的	一般不更新,只追加
一个时刻操作一个单元	一个时刻操作一个集合
面向事务	面向分析
一次操作数据量小	一次操作数据量大
支持日常操作	支持决策需求

图1 原始数据与分析型数据的比较

从图1可以清楚地看到,操作型数据与适合于决策支持系统(DSS)应用的分析型数据之间的差别甚大。由于数据仓库要用到分析型数据,因此在数据进入数据仓库之前,必须要经过加工与集成。首先要统一原始数据的格式,然后还要将原始数据做一个从面向应用到面向主题的大转变。

3)数据仓库是稳定的 数据仓库反映的是历史数据的内容和历史趋势的变化,而不是当前数据。因而,数据经集成进入数据仓库后是很少或根本不更新的。

4)数据仓库是随时间变化的 数据仓库随时间变化的特征表现在以下几个方面:首先,数据仓库内的数据时限要远远长于操作型环境中的数据时限。其次,操作型环境存储的是当前数据,即在存取一刹那正确、有效的数据;而数据仓库中的数据都是历史数据(哪怕只是几分钟前的数据)。最后,数据仓库中的数据都是按照时间顺序追加的,它们都带有“时间”这一属性,标明了该数据的历史时期。

2.2 决策分析的新进展——OLAP

与数据仓库紧密联系在一起的是“联机分析处理”OLAP(On-line Analytical Processing)。它的产生是由于传统的决策分析工具在性能和灵活性上已经不能满足人们对海量数据的分析需求、客观上要求有另外一种直接面向分析人员的、灵活的、符合人们思维习惯的分析工具来替代现有的工具。而面向决策分析的数据仓库的产生为这种工具的出现提供了基础,于是OLAP技术应运而生。

OLAP是基于数据仓库的信息分析处理过程,是数据仓库的用户接口部分。根据OLAP委员会的定义,OLAP是使分析人员、管理人员或执行人员能够从多种角度对从原始数据中转化出来的、能够真正为用户所理解的并真实反映多维数据特性的信息进行快速、一致、交互地存取,从而获得对数据更深入了解的一类软件技术。OLAP的目标是满足决策支持或多维环境特定的查询分析需求。

建立数据仓库的目的是为决策支持提供服务的。数据仓库中存储的数据是面向决策分析目标的、经过提炼、加工后的数据集合,这种数据的存储结构为OLAP

的实施提供了理想的环境;而OLAP作为一种多维查询和分析工具,是数据仓库功能的自然扩展,也是数据仓库中的大容量数据得以有效利用的重要保障。

2.3 数据仓库与CRM的结合

通过以上研究可以发现,作为一套基于大型数据仓库的客户资料分析系统,客户关系管理系统可以引进先进的数据仓库技术和数据挖掘技术。其中,数据仓库综合了企业的各种信息,包括来自多种源系统的数据。它支持多种结构,通过网络进行商业通信和交易处理。数据挖掘从大型数据库或者数据仓库中发现并提取所需信息或知识,帮助分析人员寻找各种数据之间的关联,寻找其中的规律性,从而提供有效的决策支持。CRM系统建立在企业原有的数据系统之上,通过建立科学的数据模型,进行数据挖掘,向企业提供客户的各种信息资料,以便企业进行决策支持。

总之,从技术角度来讲,数据仓库技术适于和市场营销相结合,用以建立一个基于数据仓库技术的CRM解决方案。

三、基于数据仓库技术的CRM解决方案

下面,我们将结合客户关系管理周期来说明如何建立一个基于数据仓库技术的CRM解决方案。

3.1 客户关系管理周期

客户关系管理周期包括评估阶段(Assess)、计划阶段(planning)、执行阶段(Execute)。在这个周期中,评估额是由在这个营销过程中获取的知识所组成,计划阶段则包括营销过程中最有创新性的一部分,而执行阶段被划归为与客户交互这一范畴。评估阶段运用内部数据、外部人口统计学、心理学和其他数据的综合信息开发出一个针对特定用户行为的模型,这个阶段是最具有技术性的阶段,也是周期中最关键的阶段。通过前面的介绍不难发现,由于数据仓库是面向主题的、集成的、稳定的、不同时间的数据集合,所以可以用以支持经营策略制定过程中的决策制定过程。采用数据仓库技术,可以对已有的数据和从企业外部渠道获得的数据进行转换、综合处理。OLAP可以作为分析数据要求的工具,用来记录、分析和挖掘数据中隐藏的信息,针对特定问题进行联机数据访问和分析。而作为一个高度自动化决策支持过程,数据挖掘则用以分析原有数据、潜在的信息,预测客户的行为,帮助企业的决策者调整市场策略,减少风险,作出正确判断和决策。

在计划阶段中,如何制定营销战略和策略来最好地达到客户所要求的目标是要解决的主要问题。这个阶段较少依赖于技术。通常来说,计划阶段是营销中最具有创造力的阶段,由工具和框架来支持。

执行阶段把所有的知识运用于实践。从IT业的前景来看,有必要知道要采集什么样的数据和怎样运用、

才能建立正确的数据标准和数据定义,能够使后来使用这些数据变得简单化和有意义。

3.2 客户关系管理系统的结构

CRM 应用软件应该是多种功能组件、先进技术与多种渠道的融合。功能组件包括销售应用软件(销售队伍自动化软件,销售配置和基于 Web 的自助销售)、营销自动化(SFA),以及客户服务和支持应用软件。实现的渠道包括 Web、呼叫中心和移动设备。

通过以上分析可以知道,完整的客户关系管理方案由客户信息中心和客户服务模块以及呼叫中心平台组成。其中的客户信息中心(或称为客户信息管理系统,包括客户信息采集、客户信息分析以及跟客户服务接口)作为客户关系管理方案的核心,负责完成完整客户档案库的建立和维护。数据仓库技术是它的关键所在。只有当所有在评估阶段所需的数据形成数据仓库后,这些具体的信息才被发送到呼叫中心、SFA 和决策机构。

3.3 CRM 中有关数据仓库的实现步骤

1) 建立数据模型 由于企业汇聚了客户的多种数据信息,并来自不同的数据库系统,所以在它构建客户关系管理系统时,不可避免地会遇上如何管理这些浩如烟海的数据,以及如何从中提取有用的信息的问题;而数据仓库的最大优点在于它能把企业网络中不同信息岛上的商业数据集中到一起,存储在一个单一的集成的数据库中,并提供各种手段对数据进行统计、分析。

建立数据仓库的首要任务是明确客户的需求,为数据仓库中存储的数据建立模型。企业在长期经营中存储了大量的客户信息,各自分散在不同的数据库中。通过数据模型,可以得到企业客户完整而清晰的描述信息,通过分析客户具体指标,逐渐量化来构造数据模型,进一步发掘数据构建性能卓越的数据仓库,建立起整个 CRM 的框架体系。

2) 设计和创建 CRM 数据仓库 数据仓库是数据的中心仓库,包括来自多种源系统的数据。它支持多种结构,通过网络进行商业通信和交易处理。在数据模型构造成功后,将进入数据仓库系统的开发阶段,实现大型数据库与市场营销的结合。

考虑到技术实现上的相似性,可以把客户信息数据库当作面向客户信息主题的一个数据仓库。建立数据仓库的一般思路是:首先建立一个企业级的中央的唯一的数据库,选用适当的管理工具,实现各种来源数据的自动采集、清洗、转换、汇总,然后通过各种分析工具(如多维分析和数据挖掘),有效地将数据转化为灵活的、对决策有用的信息,最后将信息(指 CRM 信息)共享给有关业务决策部门(这里指通过 CALL CENTER 平台对客户交互服务的客户服务模块或市场销售人员),便于他们对客户进行“一对一”个性化服

务。

作为一个数据仓库,它一般包括数据采集、数据存储和维护、数据分析、信息访问四个步骤,下图是数据仓库系统的示意图。图中包括数据仓库层、工具层及它们之间的关系,展示了各部分是如何有机集成的,其中,“维”是数据仓库中的一个新概念,它提供了测量值(如销售量、投资额、收入等)的上下文关系。例如,销售量与特定的产品、地区和销售时间有关,这些相关的维唯一确定了销售量这个测量值。而其中的“产品维”可以是“PC—电脑—信息产业”,“地区维”可以是“长沙—湖南—中国—亚太地区”,“时间维”一般是“日—月—季度—年”,时间维是一个有特殊意义的维,它对决策中的趋势分析有着重要的意义。利用多维数据可以从下不同角度方便地对这些测量值进行分析比较,有利于发现或找到事物的内在规律或问题产生的原因。

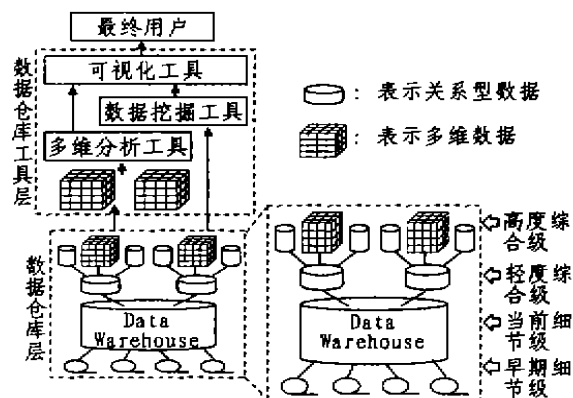


图2 数据仓库系统示意图

总结 数据仓库是使企业信息集成起来的最有力的方式。数据仓库技术与其他软件有机结合,可以有效地进行企业运作中的客户关系管理。基于数据仓库技术的 CRM 系统通过分析各种数据之间的关联,衡量客户的需求、忠诚度、满意度、赢利能力、潜在价值、信用度和风险度等指标,为企业管理层提供正确的决策支持,提升其竞争能力和赢利能力。

目前国内基于数据仓库技术的 CRM 解决方案的应用基本处于示范阶段,软、硬件都有关键技术需要攻克。在市场竞争日益激烈的今天,数据仓库技术已经作为企业竞争制胜的关键因素正被逐渐地应用在众多领域。在可以预见的将来,数据仓库技术将成为决定企业胜负的重要因素之一。

参考文献

- 1 陈文伟,等.数据仓库与决策支持系统.计算机世界专题
- 2 李纪华,等.决策支持工具的新进展——联机分析处理.计算机世界专题