

客户知识管理的数据挖掘方法研究

王 峥¹, 王彦庆²

(1. 四川大学 华西医院科技部, 成都 610041; 2. 北京交通大学 交通运输学院, 北京 100044)

摘 要: 作为企业信息化的最新发展, 客户知识管理实现了由以产品为中心向以客户为中心的转变, 客户知识成为企业的重要战略资源, 客户知识管理已成为企业提升企业竞争力的有效措施。通过探讨数据挖掘方法在客户知识管理中的应用问题, 利用数据挖掘方法以实现客户知识的获取。首先结合数据挖掘技术及客户知识管理相关理论知识, 阐明了数据挖掘方法在客户知识管理中的应用流程及应用领域, 在此基础上运用聚类分析方法, 建立了基于数据挖掘技术的数据仓库模型。最后, 通过挖掘聚类分析的实践, 明确了在客户知识管理中实施数据挖掘的步骤, 认识和总结出实践中出现的问题, 验证了数据挖掘方法在客户知识管理中的可行性。

关键词: 客户知识管理; 数据挖掘; 客户知识获取

中图分类号: F270 **文献标志码:** A **文章编号:** 1009-1971(2009)05-0113-06

前 言

随着经济全球一体化进程的加快和各企业间竞争的加剧, 企业已经由传统的以产品和规模为中心的粗放型管理模式向以客户为中心、服务至上、实现客户价值和实现企业利润最大化的集约型管理模式转变, 人们越来越意识到对客户知识的有效管理和开发决定着企业未来的发展。特别是施乐调研中心的报告表明, 一个非常满意的客户的购买力是满意客户购买力的 6 倍, 客户的满意度和忠诚度将直接影响企业的销售和成本。识别高价值客户, 并以个性化服务提高其满意度和忠诚度, 是获得市场竞争优势的有效途径, 是建立个性化客户服务模式的前提^[1]。由于竞争的加剧, 企业必须在客户管理方面下工夫。那么识别各种客户知识, 就显得非常重要。因此企业界和理论界开始思索如何在企业内引入一种创新型的机制, 从内外部获取信息并进行深入地分析, 从而创造新的知识并利用这些知识来寻找解决问题的方法, 于是客户知识管理 (Customer Knowledge Management, CKM) 的研究

开始成为人们普遍关注的问题^[2-3]。

近年来, 信息技术的出现改变了市场行为以及公司管理客户信息的途径。新的信息技术工具使得获取大量客户数据成为可能, 给利用客户数据获取竞争优势带来了机遇和挑战。一方面, 许多企业已经意识到大型数据库中的信息是支持不同组织决策的关键, 因为这些数据库中的客户知识对于市场功能至关重要。另一方面, 日趋激烈的竞争和客户可选择性的增加给市场决策者创造了新的压力, 需要对客户知识进行妥善处理。客户知识管理需要企业根据客户的实际偏好与他们进行交流, 定制他们的产品。大家逐渐意识到只有对客户需要和发现潜在客户等相关知识有了真正的理解才能实现企业的有效管理。这使得客户知识管理的研究变得非常有价值, 而数据挖掘工具可以帮助发现隐藏的知识, 更好地了解客户。

一、基于数据挖掘的客户知识获取与共享

数据挖掘 (Data Mining, DM) 是指从大量的、

收稿日期: 2009-01-17

作者简介: 王峥 (1975-), 男, 重庆人, 助理研究员, 从事知识管理研究; 王彦庆 (1966-), 男, 黑龙江哈尔滨人, 博士研究生, 从事交通运输规划管理与数据挖掘研究。

不完全的、有噪声的、模糊的、随机的数据中,提取隐含在其中的、人们事先不知道的但又是潜在有用的信息和模式的过程^[4]。数据挖掘属于一个多领域学科,它结合了统计学、数据库理论、机器学习等技术,以数据仓库为基础,应用一系列技术从大型数据库或数据仓库的数据中提取人们感兴趣的信息和知识。用这些对决策有潜在价值的知识和信息,为企业经营决策和市场策划提供依据。数据挖掘的一般过程如图 1 所示。

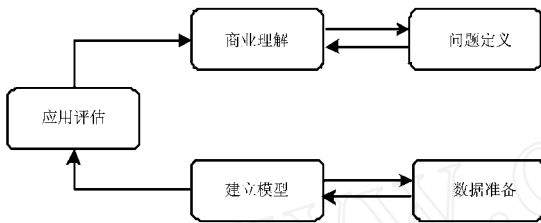


图 1 数据挖掘的流程

知识管理主要是创造、分享及再利用知识,知识管理的重点在于人们如何对待知识,它把企业和个人的知识聚合在一起并进行知识分享^[5]。也就是说,知识管理是企业实现显性知识和隐性知识共享寻找新的途径。从全局角度审视知识管理的整个过程,可以分为五个阶段:一是知识获取,它是指企业或组织进行知识资产的获得与学习;二是知识的存储,是指一旦原始数据被获取,首先面临的困难是确认哪些数据与组织相关并且应该被加入到知识管理系统中;三是知识的解释及转换,是指比较先进的系统能把无关的不同类型数据转变成一个语意相关的可使用的知识,这是真正的知识管理的关键步骤;四是知识的传播使用,是指通过知识的发布与协同达到知识共享;五是知识管理审计,是知识管理实践活动中的一项重要内容,它是指知识管理责任机构或人员按照一定的程序和方法,以知识管理的管理活动和管理的对象资源为主要审计对象,查明问题并提出解决问题的办法,以改善管理素质,提高管理水平和效率为目的,从而促进知识管理的绩效提高而进行的一种审计和知识管理相结合的活动。

知识管理既着眼于获得显性知识,更着眼于获得隐性知识。那么如何来获得这些关于客户的知识呢?关于客户的知识主要来自于交易中

的结构化数据,这些数据在日常的交易中都已经存储,它们就成为获取客户知识的数据源。而要得到对企业决策有所帮助的关于客户的知识,就需要运用现代的信息技术。首先通过运用 ETL (Extract Transform Loading),即数据提取,转换和加载工具,使得从不同渠道得来的数据有统一的格式,并且去除冗余的、错误的数据,存储在数据仓库中。然后通过运用在线分析处理和数据挖掘技术对其进行知识的挖掘和发现,最终形成可用于决策的客户知识。迈克尔 (Michael) 等提出了关于获得客户拥有的知识的三阶段模型:客户知识披露、客户知识分类、客户知识同等待^[6]。

一是客户知识披露。客户知识并不是完全显性的,它需要有人去披露。在客户知识管理中,销售人员不仅要提供关于产品的产地、实用性等基本的信息,还要成为一个专心的听众,试着理解客户的需求。除了提供信息外,销售人员还可以从客户中获得知识:关于产品或服务的选择参数,能够吸引他们的属性,将来产品和服务的行业趋势。销售人员能够收集的知识可能就是客户购买这个产品的原因,比如功能、价格等。客户知识披露是一个持续的过程,各阶段陆续地披露,从而不断地增加客户知识量。

二是客户知识分类。为了能够帮助客户作决策,要对客户知识根据不同的方面进行分类,这些方面有产品特征、功能属性等。通过分类,可以给客户提供个性化的服务,使得客户下次进行重复购买。久而久之,就有了很高的转移成本,最终将客户锁定,也就会使得客户的终身价值最大化。在此机制中,主要有以下决定因素:客户关于产品和企业的知识、客户需求信息、适合于客户需求的信息类型和交流可用的时间。一旦这些因素被决定,销售员就可以提出他为客户鉴定的各类知识。这使得企业可以和客户建立长期的伙伴关系。

三是客户知识同等待。客户知识同等待就是使得客户和销售人员之间达成共识。通过前面两个阶段,客户和销售人员之间都已获得相互间的知识,但是再怎么交流,双方也不可能达到完全了解,从而就需要通过第三阶段即知识同等待,使得双方达成共识。

客户知识管理并不是在获得客户知识后就

停止, 不仅由于上述那样的客户知识是零散的、没有整理的, 主要是知识要通过共享才能体现其最大的价值, 才能为企业带来最大的效益, 因此还要经历以下两个阶段: 一是知识成文化; 二是知识共享。企业众多的销售人员将收集的客户知识经过知识成文化阶段, 对客户知识进行整理、整合, 并最终输入到知识库中, 以便企业的其

他人员使用。而客户知识进入知识库后, 就进入对知识的应用阶段。不同的部门、不同的层次对知识有不同的应用, 从而客户知识的共享要基于知识库, 并且建立一套共享的机制, 使得收集的知识在需要时能发挥其巨大的作用。企业中客户知识的获取过程如图 2 所示。

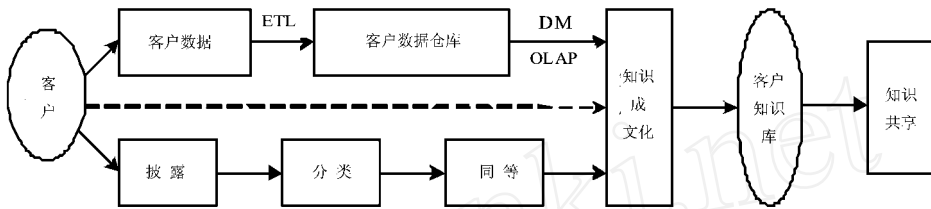


图 2 客户知识的获取过程

二、数据挖掘在客户知识管理中的流程设计

数据挖掘使数据库技术不仅能对过去的数据进行查询和遍历, 并且能够找出过去数据之间的潜在联系, 从而促进信息的传递^[7]。在 CKM 中实施数据挖掘的步骤一般可以包括以下的七

个方面, 分别为: 定义问题、数据选择、数据预处理、建立数据挖掘模型、实施数据挖掘、结果分析和评估、知识的同化。在客户知识管理系统中数据挖掘中具体流程如图 3 所示。

首先是主题定义。在数据挖掘工作开展之前, 必须明确数据挖掘所要解决的问题和达到的预定目标。也只有在目标明确定义的前提下, 才能准备有用的数据, 数据挖掘的工作才有方向和意义。根据用户的需求, 本文定义的主题是客户

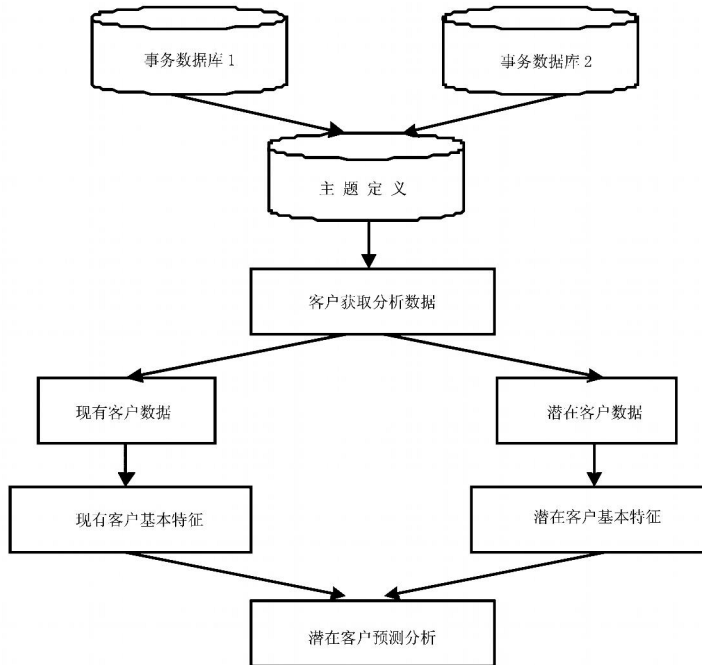


图 3 客户知识库数据挖掘系统流程图

知识的获取。客户知识获取包括发现那些对公司和产品不了解的客户,他们可能是你的产品的潜在消费者,也可能是以前接受你的竞争对手服务的客户。我们把这些客户称为潜在客户,客户知识获取的目的就是把这些潜在客户信息转变为客户知识^[8]。

其次是数据选择与预处理。从主题定义中,可以看到运用数据挖掘知识找出一个好的预测模型是最关键的一步,只有预测出来的结果有较高的准确率,明确哪些客户信息有较大可能转变为客户知识从而真正获得客户,才能有针对性地制定相应的策略,得到较为满意的效果。要得到一个好的模型的前提是高质量的数据基础,所以首先要明确与挖掘任务的相关数据。在 CKM 的数据挖掘系统中,主要用到三种类型的数据,分别是:描述性数据,记录客户的基本信息,如年龄、性别、家庭住址、个人收入等等;促销数据,记录对客户进行了哪些促销活动;交易数据,记录企业和客户相互作用的所有数据。客户知识获取分析中要用到的数据包括现有客户的基本信息和交易情况,以及潜在客户的基本信息和对于一些市场活动的反应行为数据。

再者是建立数据挖掘模型。虽然数据挖掘可以建立在联机事务处理数据库上,但数据库对数据的组织更有利于提高数据挖掘的效率。数据仓库中完备的、经过组织的大量数据为数据挖掘提供了很好的基础。一个健全、有效、完整的数据仓库系统是进行数据分析的前提和必要条件^[9]。数据仓库中,可以对数据进行多维化处理,构成多维数据模型,用多维分析的方法从多角度、多层次对数据进行统计分析。多维数据模型包括事实、维和粒度。事实是分析的目标数据;维是事实信息的属性,也是考察事实的角度;粒度是维划分的单位,也就是明细的程度,可以用超立方体结构来表示事实、维和粒度的关系。所以在数据仓库系统中统计分析就统一归结为从不同的角度、不同的层次来观察分析数据。

还有是实施数据挖掘以及结果分析与评估。数据挖掘模型建立并验证之后,就可以对所得到的经过预处理的数据进行挖掘了,并且在实施数据挖掘的过程中可能需要重新完善挖掘算法。解释并评估结果,可使用可视化和知识表示技术,向用户提供挖掘的知识。利用比较直观的图

形表达方式来加深用户对数据含义的理解,用形象直观的图像来指引检索过程,加快检索速度。同时将结果应用于实践中加以评估和验证^[10]。评估模型结果的多种方法中,最可能产生评价过高的指标就是精确性。假设有个提议仅仅有 1%的人响应。模型预测“没有人会响应”,那么这个预测 99%就是正确的,但是这个模型是 100%无效的。另外一个经常使用的指标是“提升多少”,用它来衡量使用模型后的改进有多大,但是它没有考虑成本和收入^[11]。

三、案例分析

数据仓库中最常用的模型是星型模型。星型模型由事实数据表和链接到该事实数据表的多个维度表组成,结构比较简单,运行效率较高。所以本文主要采用星型模型建构数据仓库^[12],以 Microsoft visual studio net 为开发平台,建立一家酒店客户知识获取模型。

首先是从酒店的联机事务处理数据库中收集、整理原始数据,组建酒店客户的数据仓库系统,在客户数据仓库中抽取适当的字段组成客户分析数据库,为客户细分和建立潜在客户行为模型提供数据源。这其中关键是获得长期客户有价值的信息,并将其储存到酒店数据库中。因为过去是对将来的最好的预测,虽然这条定律并不总有效,但至少是一个有指导意义的猜测。我们将现有客户的特征和消费行为作为挖掘潜在客户的一个参考,结合对潜在客户的二元反应行为建模,从而达到低成本的客户知识获取的目的。

其次对现有客户进行细分,参照一般客户价值发现过程,以最近一次消费、消费频率、消费金额三项指标建立 RFM 模型,找出某一类客户的共同特征,为潜在客户的获取提供参考,如图 4 所示。进行客户细分通常都有一定的目标,根据目标选择合适的细分标准是 CKM 系统中客户细分的原则。本课题中进行客户细分的主要目的是从现有客户的消费行为模式中找出客户的某些重要特征,为潜在客户的特征提供一个参考,以加强对潜在客户获取的决策力度。我们采用 RFM 模型进行客户群的细分。根据酒店用户的要求,分别把 RFM 三个要素划分为三个等级,结

合这三个指标,可以把客户分成九个等级,也就是聚类成九个簇。在前面建立的数据仓库模型中已经有了每个客户的 RFM 数据,利用聚类方

法中的 k - means 算法对客户数据进行聚类,就能得到现有客户细分的结果。

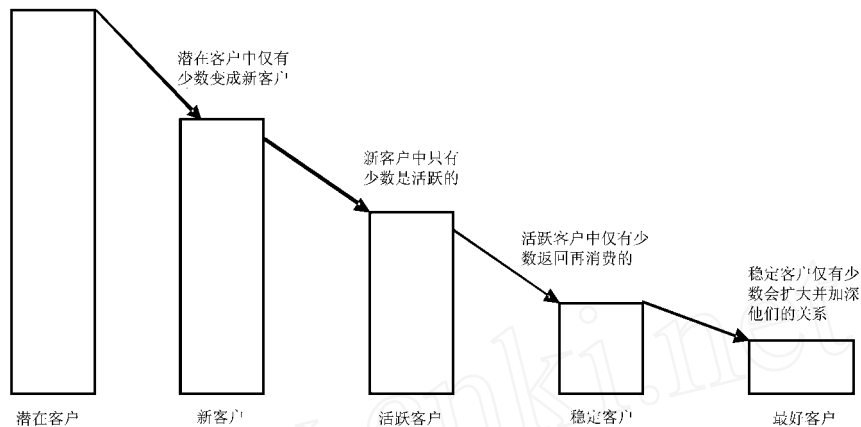


图 4 一般客户价值发现过程

再者对于大量的潜在客户数据,建立潜在客户的二元反应行为预测模型。在早期的数据收集工作中,我们建立了酒店数据仓库,记录了很多潜在客户的信息,我们抽出一部分客户进行营销反应测试,这样就得到了分类用的训练集。一次市场活动后得到的客户反馈叫做“反应行为模式”,反应行为模式可以用来定义出不同类型的客户反应行为,并且可以对不同反应的可能性进行分类,以便深入的研究和制作报告。反应行为模式分为两种,一种是二元反应行为模式,在二元反应行为模式中,客户的反应被归纳成两类:是或者否;另一种是分类反应行为模式,分类反应行为模式允许根据所从事的商业领域定义多种反应行为,带有一些随机性。

综上所述,本文采用一种简单的反应分类方法:二元反应行为模式,分类属性为潜在客户对酒店的反应。用 1 表示积极反应,即对该酒店有兴趣,可能会入住;0 表示消极反应,即对于入住该酒店有兴趣。然后根据资深酒店专家的经验,选择以下属性作为测试属性。

职业:销售 (1),非销售 (0)

收入: < = 2000 元 (1) > 2000 元 (0)

地区:华北,华东,东北,西北,其他

年龄:年轻 (< = 25),中年 (20 - 55),老年 (> = 55)

可以看到在这个训练集中,地区属性的取值

较多,这也就是要 D3 算法的启发式函数的原因。用 D3 算法对训练集数据进行学习就可以得到潜在客户的反应行为模型。

结 论

本文首先依据客户知识管理理论,探讨了客户知识的获取过程,并结合数据挖掘方法,提出了客户知识管理中的流程设计方案,并以一所酒店实施客户知识管理为实例,明确了在 CKM 中实施数据挖掘的步骤,建立潜在客户的二元反应行为预测模型,阐述了在酒店 CKM 中实施客户知识获取的方法,为企业决策者提供参考。本文由于受到能力和研究条件的限制,在此只能对客户知识管理的数据挖掘方法作一些基础性的研究,随着竞争的日趋激烈,在客户知识管理中建立真正意义上的数据挖掘方法的应用是一个比较复杂的系统工程,尽管对于数据挖掘的应用是目前国内外研究的热点项目,但从总体上来看,仍然处于初步研究阶段,需要人们继续不断的深入研究和探讨。

参考文献:

- [1] DARRELL R, REICHHELD F. Avoid the Four Perils of CRM [J]. Harvard Business Review, 2002, 80 (2): 101 - 109.

- [2] FREYTAG P V, et al Business to Business Market Segmentation [J]. Industrial Marketing Management, 2001, 30(6): 473 - 486
- [3] SHANKARA V, SMITH A K, RANGASWAMY A. Customer satisfaction and loyalty in online and offline environments[J]. International Journal of Research in Marketing, 2003, 20(2): 153 - 175.
- [4] AGRAWAL R, MIELNSKI T, SWAMI A. Database Mining: A Performance Perspective [J]. IEEE Transactions on Knowledge and Data Engineering 1993, 5(6): 914 - 925.
- [5] 汪战平,柯青. 客户知识管理概念研究 [J]. 情报科学, 2004, (1): 12 - 14.
- [6] MICHAEL J, et al Knowledge Management and Data Mining for Marketing [J]. Decision Support Systems, 2001, 31(1): 127 - 137
- [7] 黄晓斌,邓爱贞. 现代信息管理的深化——数据挖掘和知识发现的发展趋势 [J]. 现代图书情报技术, 2003, (4): 4 - 8
- [8] 曾海颖. 客户关系管理中的数据挖掘 [D]. 南京:南京航空航天大学, 2003.
- [9] 汤杰,潘尔顺. 数据挖掘在 CRM 中的应用 [J]. 电子商务世界, 2003, (5): 12 - 15.
- [10] BULT J R, WANSBEEK T. Optimal Selection for Direct Mail [J]. Marketing Science, 1997, 14: 9 - 13.
- [11] 张哲,常桂然,黄小原. 数据挖掘技术在 CRM 中的应用 [J]. 中国管理科学, 2003, 11: 53 - 58.
- [12] IKUIRO NONAKA, H ROTAKA TAKEUCHI The Knowledge - creating Company [M]. New York: Oxford University Press, 1995: 345 - 356

A Method of Data Mining of Customer Knowledge Management

WANG Zheng¹, WANG Yan-qing²

(1. West China Center of Medical Sciences, Sichuan University, Chengdu 610041, China;

2. School of Traffic and Transportation Beijing Jiaotong University, Beijing 100044, China)

Abstract: Customer knowledge management has changed from a product - centered model to the client - centered one, which is the latest development of industrial information. And customer knowledge has turned out to be an important strategic resource and its knowledge management has become an effective measure to increase enterprise competence. This paper emphasizes the practical application of data mining method in the customer knowledge management and proposes that the data mining method can be used to obtain client knowledge. By combining the related theories of client knowledge management of data mining technique, this paper firstly illustrates the application flow and field of data mining method in the client data management, then uses the cluster analysis method to establish data mining technology - based data warehouse data model. Finally, the paper defines the steps of implementing data mining, cognizes and summarizes the existent problems, and validates the feasibility that data mining can be used in the customer knowledge management, through a practice of excavation cluster analysis in a case.

Key words: customer knowledge management; data mining; customer knowledge acquisition

[责任编辑 陈 朴]