

# 基于数据挖掘的客户细分框架模型

方安儒, 叶强, 鲁奇, 李一军

(哈尔滨工业大学管理学院, 哈尔滨 150001)

**摘要:** 数据挖掘技术在客户关系管理领域的应用较广泛, 能提高客户细分能力。针对目前客户细分研究缺乏统一研究框架的问题, 分析现有的客户关系管理系统构架及其与客户细分的集成关系, 对客户细分问题进行构架性研究, 提出一种基于数据挖掘的客户细分框架模型, 包括空间逻辑模型和数据-功能-方法模型。

**关键词:** 客户细分; 框架模型; 数据挖掘

## Customer Segmentation Framework Model Based on Data Mining

FANG An-ru, YE Qiang, LU Qi, LI Yi-jun

(School of Management, Harbin Institute of Technology, Harbin 150001)

**【Abstract】** Data mining techniques are broadly used in customer relationship management, which enhance the capacity of customer segmentation to a new level. Since there is no unified framework model for customer segmentation by now, this paper analyzes the recent relation of customer relationship management structure and customer segmentation, and proposes a customer segmentation framework model based on data mining, which includes space model and DFM(Data, Function, Method) model.

**【Key words】** customer segmentation; framework model; data mining

### 1 概述

在经济发展过程中, 企业管理理念经历了从以产品为中心到以市场为中心, 再到以客户为中心的转变历程。20世纪90年代初, Gartner Group 最先提出了完全以客户为中心的管理理念及信息系统——客户关系管理。随着管理理论与管理手段的发展, 深入的客户数据分析将在客户关系管理中发挥越来越重要的作用, 客户细分则成为一种基础性的分析手段, 对包括操作层、战术层以及战略层在内的客户管理提供全面的信息支持。客户细分是企业明确的战略、业务模式和特定的市场中, 根据客户的属性、行为、需求、偏好以及价值等因素对客户进行分类, 并提供针对的产品、服务和营销模式的过程。美国著名的管理咨询公司 Bain & Company 对来自世界各地的 451 位高级管理人员的一项调查表明, 客户细分已经成为全球使用率最高的十大管理工具之一。

传统的客户细分方法一般是基于经验的分类方法或基于统计的简单划分方法, 无法满足诸如对哪些客户的潜在价值更高、哪些客户的资信程度更高等复杂分析的需求。面对电子商务所产生的海量客户数据, 传统方法也显得力不从心。

以数据挖掘为核心的商务智能技术的出现, 为海量数据下的复杂客户细分问题提供了新的解决方法。文献[1]以面向分类的 BP 神经网络为基础, 提出基于 RFM 分析的合并反应模型, 用以解决海量数据下的客户细分问题; Natter 等提出基于 K 平均聚类的客户细分方法; 文献[2]在对网上企业客户的细分问题的研究中, 采用自组织映射神经网络方法进行了聚类分析; 文献[3]提出基于 Hopfield-Kagmar 神经网络的客户细分方法; 文献[4]提出利用分对数混合模型, 根据客户多种购买行为数据对客户进行了细分; 马辉民等针对特定领域问题提出基于客户份额的 K 平均聚类客户细分方法。

上述研究者普遍采用了 2 类数据挖掘技术——分类技术(classification)与聚类技术(clustering)。然而, 由于目前尚未形成统一的基于数据挖掘的客户细分构架, 它们的地位和作用还不够明确, 这 2 类技术在研究和应用中常常被细分(segmentation)一词所混淆。目前针对客户细分的研究也还较为分散, 缺乏统一的研究框架。本文针对基于数据挖掘技术的客户细分进行框架性研究, 初步建立了基于数据挖掘的客户细分框架模型。

### 2 客户细分与 CRM 系统的集成构架分析

#### 2.1 操作型 CRM 系统

最初的客户关系管理系统是偏重客户关系管理业务流程的, 称为操作型 CRM 系统, 其结构如图 1 所示。

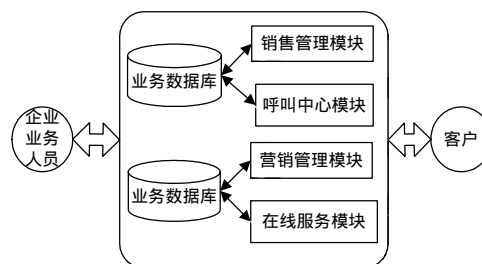


图 1 操作型 CRM 系统结构

**基金项目:** 国家自然科学基金资助项目“面向非确定性客户细分的商务智能理论与方法研究”(70501009); 国家自然科学基金资助项目“面向电子商务的中文客户评论情感分析研究”(70771032)

**作者简介:** 方安儒(1962 - ), 男, 博士, 主研方向: 电子商务; 叶强, 教授; 鲁奇, 博士研究生; 李一军, 教授

**收稿日期:** 2009-01-23 **E-mail:** yeqiang2006@gmail.com

作为企业业务人员的客户管理中心及客户接触中心，操作型客户关系管理系统主要执行面向作业级的客户管理应用，其主要功能模块包括销售管理模块、呼叫中心模块、营销管理模块及在线服务模块，由一个或多个业务数据库支持。在操作型的客户关系管理系统中，由于客户分析功能不是其主要功能，因此一般不包含复杂的客户细分模块。有些操作型的客户关系管理系统可以对客户进行简单的分类管理，此时所使用的一般仅是基于客户属性特征的简单客户分类。

## 2.2 分析型 CRM 系统

分析型 CRM 系统主要包括数据存储模块与分析手段模块 2 个部分，其结构见图 2。

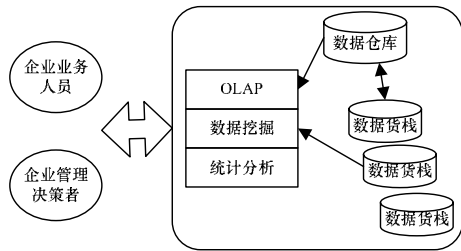


图 2 分析型 CRM 系统结构

分析型 CRM 通过采集企业内部与客户及交易有关的数据和各种相关的外部数据，建立企业数据仓库，在此基础上，运用多种分析手段，形成对企业各种级别决策具有支持作用的知识，并按照规范的接口，供企业的其他信息系统使用。分析型 CRM 能够对客户数据进行智能化分析，其主要技术内容包括数据仓库、在线分析处理(OLAP)、数据挖掘、统计分析等。其中，数据仓库将集中存储客户的数据与信息，为一切分析活动提供基础数据，是分析型客户关系管理系统的核心支撑部件；统计分析与在线分析处理模块将实现用户驱动的初步数据分析，是分析型客户关系管理系统的基本组成部分；数据挖掘模块则将实现数据驱动的自动数据分析与信息发现，这一模块将执行各种复杂的数据分析功能。

在分析型客户关系管理系统的各种分析功能中，客户细分作为一种重要的分析功能及客户管理手段，正日益受到企业及软件开发商的重视。客户细分模块或客户聚类与分类模块也成为大多数分析型客户关系管理所具有的一个基本功能模块。

## 2.3 综合型 CRM 系统

目前的许多客户关系管理系统同时具有分析性与操作型系统的功能，称为综合型 CRM 管理系统，其结构见图 3。

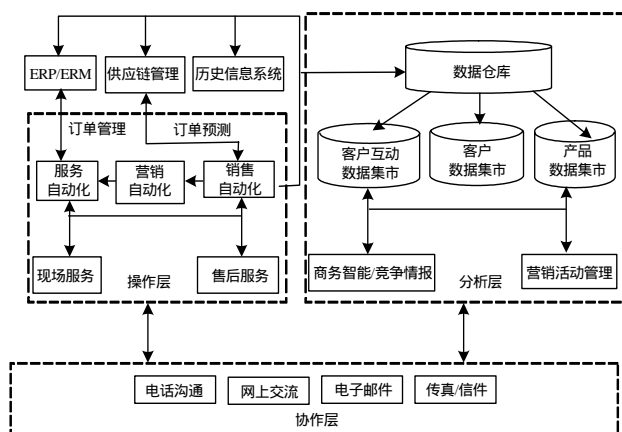


图 3 综合型 CRM 系统结构

综合型 CRM 系统相当于分析型与操作型客户关系管理系统的结合，按其功能结构划分可以分为 3 个层次，即操作层、分析层及协作层。在该系统中，客户细分模块作为一个客户分析工具，位于系统的分析层中。客户细分对客户的分类结果将为操作层及协作层提供信息支持。

## 3 客户细分的逻辑模型

### 3.1 客户细分的空间逻辑模型

在制定、选择营销策略或进行有关客户关系管理的决策时，管理者经常用到一些描述客户的抽象的概念，如“高收入客户”、“低收入客户”、“新潮客户”、“保守客户”等。有些概念的内涵比较明确，外延也较容易确定，如“高收入客户”；而更多概念的内涵常不能结构化描述，其外延的确定也受到许多主客观因素的影响，如“新潮客户”。客户细分的主要任务就是确定这些概念与客户间的对应关系。

客户数据中包含若干离散客户属性与连续客户属性，以每一客户属性作为一个维度，每个客户作为空间的一点，则企业客户数据库中的所有客户可构成一个多维空间，称为客户的属性空间。

设  $A = \{A_1, A_2, \dots, A_n\}$  是一组描述客户特征和行为的属性，这些属性可以是连续属性，也可以是离散属性，它们构成了客户的一个  $m$  维属性空间  $A^m$ ，每一个属性又称为一个属性维。每个客户在这  $m$  个属性维上的取值决定了客户在属性空间的位置，设  $c \in C$ ， $c$  在属性  $A_i$  上的取值记为  $c[A_i]$ 。

设  $g$  是一个描述客户的抽象概念， $f(g)$  是符合概念  $g$  所描述的客户的集合，即概念的外延。对于一组概念  $g_{i1}, g_{i2}, \dots, g_{ik}$ ，若  $f(g_{i1}), f(g_{i2}), \dots, f(g_{ik})$  在任一确定时刻都是  $k$  个互不相交的集合，并且  $c = f(g_{i1}) \cup f(g_{i2}) \cup \dots \cup f(g_{ik})$ ，则称  $G_i = g_{i1} \wedge g_{i2} \wedge \dots \wedge g_{ik}$  为一个概念维度， $g_{i1}, g_{i2}, \dots, g_{ik}$  称为概念维度  $G_i$  上的概念。对于  $\forall c \in C$ ，若  $c \in f(g_{ij})$ ， $1 \leq j \leq k$ ，则记为  $\varphi(c, G_i) = g_{ij}$ ，表示  $c$  在  $G_i$  这个概念维度上属于  $g_{ij}$  这个概念概括的范围。

如在客户价值这个概念维度上，可以有“价值客户”、“潜在价值客户”和“无价值客户”3 个概念，这 3 个概念的外延概括了所有的客户。

设  $G_1, G_2, \dots, G_n$  为  $n$  个不相同的概念维度，称这  $n$  个概念维度构成了客户的一个  $n$  维概念空间  $G^n$ 。

设  $B \subseteq G$ ， $R_B \subseteq C \otimes C$  是  $C$  上的一个二元关系， $C \otimes C$  表示  $C$  的笛卡尔乘积， $R_B$  定义如下：

$$R_B = \{(c_1, c_2) \mid c_1, c_2 \in C, \forall b \in B, \varphi(c_1, b) = \varphi(c_2, b)\}$$

显然  $R_B$  是一个等价关系，由  $R_B$  可以将属性空间中的客户分为若干个等价类，每个等价类称为一个概念类。这一组概念类称为由  $B$  确定的概念类。

客户细分是建立客户属性空间  $A^m$  与概念空间  $G^n$  映射关系  $A^m \rightarrow G^n$  的过程，本文称这种客户细分模型为客户细分的空间逻辑模型，如图 4 所示。

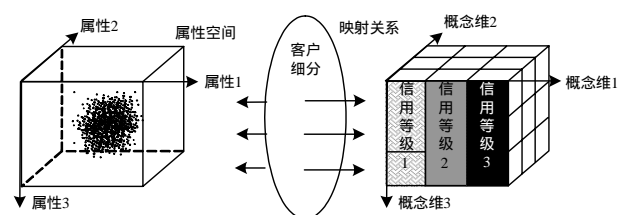


图 4 客户细分的空间逻辑模型

### 3.2 客户细分的数据-功能-方法模型

为在客户细分研究中建立统一的功能结构模型，作为今后客户细分研究工作的基础，本研究探索建立了一个标准的客户细分功能结构模型——客户细分的DFM模型。

#### 3.2.1 DFM模型总体结构

本文提出的基于数据挖掘的客户细分结构模型总体结构见图5。

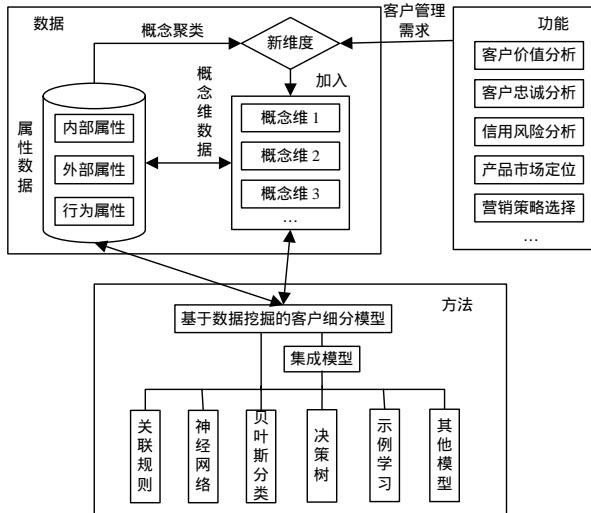


图5 客户细分的DFM模型

该模型包括数据、功能及方法3个部分，为此本研究将该模型命名为客户细分的DFM模型，即数据(Data)、功能(Function)、方法(Method)模型。

数据部分主要由客户数据存储和概念维数据构成。其中，客户数据存储包括企业客户的全部内在属性、外在属性及行为属性数据。概念维数据包含概念空间中与客户分析相关的各种概念描述。客户数据存储与概念维间通过2种方式建立相互关联，一小部分客户数据通过外生方式与概念维建立关联，这部分数据将成为数据挖掘的学习样本源；此外多数客户数据则将通过数据挖掘模型建立与概念维间的映射关系。

方法部分是该客户细分模型的核心部分，该方法部分将包括多种数据挖掘方法以及这些方法的集成。数据挖掘将通过对客户数据存储中数据的学习算法，建立客户数据与概念维的映射关系。

功能部分包括客户价值分析、客户忠诚分析、信用风险分析、产品市场定位、营销策略选择等客户管理功能，这些管理功能的基础是客户与概念间映射关系的建立。随着管理

(上接第235页)

## 7 结束语

本文利用SDG对化工流程工艺的内在机理进行建模获取规则，建立故障诊断系统。知识库采用Access数据库，规则提取和人机交互界面采用VC编程实现。并将其应用到原油电脱盐系统的故障诊断中，有较好的实用价值。该知识获取方法及知识库构建方法可推广到其他流程工业故障诊断知识库的建立过程，为故障诊断提供有力辅助，最终达到减少事故发生，提高设备可靠性，减少或解决环境污染，提高企业经济效益与社会效益的目的。

### 参考文献

[1] 刘铭, 时昕, 姚燕南. 基于数据库的电力设备故障诊断模糊

实践的发展, 客户关系管理过程将产生新的概念维度, 这些新的概念维度将被加入到概念维数据中, 并通过数据挖掘建立起与客户属性空间的映射关系。

#### 3.2.2 概念维数据的维护

概念维数据中的概念维数由客户管理的具体需要而确定, 并将随着时间的推移而不断增加。新概念维度的产生主要有2个来源: 一个是来自于客户关系管理的需求, 另一个是来源于对客户属性数据的概念聚类。

将客户按管理需要进行分类, 既是客户细分的目的, 也是客户细分问题中导致新概念维产生的直接原因。例如, 某企业原本没有对客户按风险进行分类的需要, 然而, 由于市场条件的变化, 客户信用风险逐渐成为影响企业经营策略的一个重要因素, 因此便产生的对客户按信用风险进行分类的客户细分需求。假设根据管理需求, 应将客户按A, B, C3级进行风险管理, 那么在其客户细分系统的概念维数据中将加入由A, B, C3个概念组成的信用风险这一概念维。

另外, 聚类分析也是一个重要的概念维来源。聚类分析既可以产生新的概念维度, 又可以为客户属性空间中的客户与该概念维度建立直接关联, 因此, 也是客户细分中的一种重要的分析工具。

## 4 结束语

本文分析了现有的客户关系管理系统构架及其与客户细分的集成关系, 提出客户细分的空间逻辑模型和基于数据挖掘的客户细分数据-功能-方法模型——DFM模型。下一步研究方向是将不同研究成果进一步置于该分析框架下进行对比分析。

### 参考文献

[1] Suh E H, Noh K C, Suh C K. Customer List Segmentation Using the Combined Response Model[J]. Expert Systems with Applications, 1999, 17(2): 89-97.  
 [2] Alfredo V. A Methodology for the Characterization of Business-to-Consumer E-commerce[D]. Liverpool, UK: John Moores University, 2000.  
 [3] Boone D S, Michelle R. Retail Segmentation Using Artificial Neural Networks[J]. International Journal of Research in Marketing, 2002, 19(3): 287-301.  
 [4] Heilman C M, Douglas B. Segmenting Consumers Using Multiple-category Purchase Data[J]. International Journal of Research in Marketing, 2002, 19(3): 225-252.

编辑 金胡考

专家系统的设计与实现[J]. 计算机工程, 2001, 27(3): 75-77.

[2] 曹文君. 知识库系统原理及其应用[M]. 上海: 复旦大学出版社, 1995.  
 [3] 张贝克, 夏涛, 吴重光. 集成化SDG建模、推理与信息处理软件平台[J]. 系统仿真学报, 2003, 15(10): 1360-1363.  
 [4] 李洁, 沈士团, 陈星. 航电设备故障诊断专家系统研究与实现[J]. 计算机工程, 2007, 33(14): 163-165.  
 [5] 李鸣, 李辉, 高娜, 等. 某武器装备故障诊断专家系统的设计与实现[J]. 计算机工程, 2004, 30(23): 121-122, 170.

编辑 索书志

