

河北省重点学科技术经济及管理 资助出版

# 数据挖掘 在 电子商务领域中的应用

SHUJU WAJUE ZAI DIANZI SHANGWU LINGYUZHONG DE YINGYONG

张永礼 丁超 安海岗 马伟 著



冶金工业出版社  
Metallurgical Industry Press

河北省重点学科技术经济及管理

河北省人力资源社会保障科研合作课题 (JRSHZ-2015-01032)

河北省社会科学发展研究青年课题 (2015041229)

资助出版

河北省社科联社会发展项目 (2015031259)

河北省高等教育教学改革研究与实践项目 (2015GJJG122)

# 数据挖掘在 电子商务领域中的应用

张永礼 丁超 安海岗 马伟 著

北京

冶金工业出版社

2015

## 内 容 提 要

本书重点介绍了数据挖掘技术在电子商务领域的典型应用,分为理论篇和实证篇两大部分,理论篇包括第1~5章,实证篇包括第6~9章,分别研究了客户分类、网店评价、网站评价、互联网在线时长分析等问题。

本书可作为从事电子商务的企事业单位信息管理部门的管理者、信息分析人员、数据统计人员、市场营销人员、研究与开发人员的参考资料,也可作为高等院校信息管理类、数据分析类等相关专业的教材和参考书。

### 图书在版编目(CIP)数据

数据挖掘在电子商务领域中的应用/张永礼等著. —北京:  
冶金工业出版社, 2015. 11

ISBN 978-7-5024-7163-7

I. ①数… II. ①张… III. ①数据采集—应用—电子商务 IV. ①TP274 ②F713.36

中国版本图书馆CIP数据核字(2015)第281415号

出版人 谭学余

地 址 北京市东城区嵩祝院北巷39号 邮编 100009 电话 (010)64027926

网 址 www.cnmp.com.cn 电子信箱 yjcs@cnmp.com.cn

责任编辑 曾 媛 李鑫雨 美术编辑 彭子赫 版式设计 孙跃红

责任校对 郑 娟 责任印制 牛晓波

ISBN 978-7-5024-7163-7

冶金工业出版社出版发行;各地新华书店经销;固安华明印业有限公司印刷

2015年11月第1版,2015年11月第1次印刷

169mm×239mm;12.5印张;244千字;190页

45.00元

冶金工业出版社 投稿电话 (010)64027932 投稿信箱 tougao@cnmp.com.cn

冶金工业出版社营销中心 电话 (010)64044283 传真 (010)64027893

冶金书店 地址 北京市东四西大街46号(100010) 电话 (010)65289081(兼传真)

冶金工业出版社天猫旗舰店 yjgycbs.tmall.com

(本书如有印装质量问题,本社营销中心负责退换)

# 前 言

近年来，电子商务蓬勃兴起，物联网、云计算、大数据、移动互联网等新技术不断涌现，企业、政府和社会加速推进数字化和网络化，这使得越来越多的数据被收集并沉睡在数据库之内，简而言之，“数据爆炸，信息贫乏”。如何对这些海量数据进行分析和挖掘，得到企业想要的信息，进而指导企业做出科学决策，成为摆在企业面前迫切需要解决的问题，基于此，数据挖掘技术应运而生，并在短短几年之间迅猛发展。

数据挖掘（Data Mining）是从大量的、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，但却是潜在有用的信息和知识的过程。电子商务专业区别于传统商务的一点在于，与客户非“面对面”，企业对客户的一切了解都源于对交易过程中生成数据的分析和挖掘。电子商务本身也提供了丰富的数据。因此，数据挖掘对电子商务十分重要。目前数据挖掘技术已广泛应用于电子商务各个领域，如购物篮分析、商品捆绑销售、商品推荐、货架摆放、客户购买预测、公司营销最大利润点预测、商品销售量回归分析、2/8 客户分类、孤立点检测、客户购买行为模式预测、Web 网站访问模式预测、商品分类销售预测等。

本书重点介绍了数据挖掘技术在电子商务领域的典型应用，由 9 章组成，分为理论篇和实证篇两大部分，理论篇包括第 1~5 章，实证篇包括第 6~9 章，分别研究了客户分类、网店评价、网站评价、互联网在线时长分析等问题。

本书在撰写过程中得到了石家庄经济学院电子商务教研室和河北

## II ..... 前 言

省重点学科技术经济及管理的大力帮助和支持，同时得到了以下项目的资助：河北省人力资源社会保障科研合作课题“河北省农村剩余劳动力转移就业影响因素实证研究”（JRSHZ-2015-01032）、河北省社会科学发展改革研究青年课题“河北省农业机械化影响因素及发展路径研究”（2015041229）、河北省社科联社会发展项目“新常态下河北发展低碳经济与产业结构调整良性互动机制研究”（2015031259）、河北省高等教育教学改革研究与实践项目“电子商务专业‘三主体’应用型人才培养模式研究”（2015GJJG122），特此表示感谢！

本书由张永礼、丁超、安海岗、马伟著。张永礼完成了本书第1~4章、第6章、第8章，第9章内容的撰写，丁超、安海岗完成了其中部分章节数据收集和模型建立的工作，第5章和第7章理论由马伟撰写。由于作者水平所限，书中不足之处，欢迎各位读者批评指正！

张永礼

2015年6月

# 目 录

1 绪论 .....	1
1.1 研究背景 .....	1
1.2 研究意义 .....	3
1.2.1 客户分类 .....	3
1.2.2 网店评价 .....	4
1.3 国内外研究综述 .....	5
1.3.1 客户分类 .....	5
1.3.2 网店评价 .....	8
1.4 本书主要内容 .....	9
1.5 本书创新之处 .....	10
参考文献 .....	11
2 数据挖掘理论 .....	13
2.1 数据挖掘的定义 .....	13
2.1.1 数据挖掘与数据仓库 .....	13
2.1.2 数据挖掘与 OLAP 及商务智能 .....	14
2.2 数据挖掘任务及体系结构 .....	15
2.2.1 数据挖掘任务 .....	15
2.2.2 数据挖掘的体系结构 .....	17
2.3 数据挖掘过程及过程模型 .....	17
2.3.1 数据挖掘的过程 .....	17
2.3.2 数据挖掘过程模型 .....	19
2.4 数据挖掘算法 .....	20
2.4.1 关联规则 .....	20
2.4.2 聚类算法 .....	23
2.4.3 决策树算法 .....	25
2.4.4 神经网络算法 .....	28
2.5 数据挖掘在 CRM 的应用 .....	30
2.5.1 营销 .....	30

## IV ..... 目 录

2.5.2	销售 .....	31
2.5.3	客户服务 .....	31
2.5.4	客户保持 .....	31
2.5.5	风险评估和欺诈识别 .....	32
2.6	数据挖掘的工具 .....	32
2.6.1	Intelligent Miner .....	32
2.6.2	Enterprise Miner .....	33
2.6.3	SPSS Clementine .....	33
2.6.4	Microsoft Business Intelligence Development Studio .....	33
	参考文献 .....	35
<b>3</b>	<b>分形理论 .....</b>	<b>36</b>
3.1	分形理论简介 .....	36
3.1.1	简单分形 .....	36
3.1.2	多重分形 .....	37
3.2	发展历程 .....	38
3.3	分形理论的应用 .....	39
	参考文献 .....	40
<b>4</b>	<b>客户分类理论 .....</b>	<b>41</b>
4.1	客户价值内涵 .....	41
4.1.1	客户让渡价值理论 .....	41
4.1.2	客户生命周期价值理论 .....	43
4.2	基于客户行为的客户分类 .....	46
4.2.1	RFM 模型 .....	46
4.2.2	修正的 RFM 模型 .....	47
4.3	基于客户生命周期的客户分类 .....	48
4.3.1	忠诚度阶梯分类法 .....	48
4.3.2	依据客户关系的不同阶段进行客户分类 .....	48
4.4	基于客户价值的客户分类 .....	49
4.4.1	定量分类方法研究 .....	49
4.4.2	定性分类方法研究 .....	51
4.5	客户忠诚度 .....	54
4.5.1	客户忠诚度概念 .....	54
4.5.2	客户忠诚度维度结构 .....	55

参考文献 .....	57
<b>5 虚拟商店理论 .....</b>	<b>59</b>
5.1 虚拟商店 .....	59
5.1.1 虚拟商店的概念和内涵 .....	59
5.1.2 虚拟商店的特点 .....	60
5.1.3 虚拟商店销售的创新点 .....	61
5.1.4 虚拟商店的基础环境 .....	62
5.1.5 虚拟商店的客户培养 .....	63
5.2 网络经济 .....	64
5.2.1 网络经济的内涵及其形成 .....	64
5.2.2 网络经济的特征与影响因素 .....	65
5.3 绩效评价理论基础 .....	67
5.3.1 资源理论 .....	67
5.3.2 生命周期理论 .....	69
5.3.3 战略适应理论 .....	70
5.3.4 激励理论 .....	72
参考文献 .....	74
<b>6 数据挖掘在客户分类中的应用 .....</b>	<b>75</b>
6.1 引言 .....	75
6.1.1 传统的客户分类理论 .....	75
6.1.2 基于客户行为的客户分类 .....	76
6.1.3 基于客户生命周期的客户分类 .....	76
6.1.4 基于客户生命周期价值的客户分类 .....	76
6.2 模型构建 .....	77
6.2.1 当前价值的计算 .....	77
6.2.2 潜在价值的计算 .....	77
6.2.3 客户忠诚度的计量 .....	78
6.2.4 数据挖掘流程 .....	80
6.3 实证研究 .....	81
6.3.1 数据来源 .....	81
6.3.2 模型建立 .....	83
6.3.3 模型训练 .....	84
6.3.4 模型预测 .....	88



## VI 目 录

6.4 客户分类结果 .....	92
6.4.1 客户现有价值的评价 .....	92
6.4.2 客户潜在价值的评价 .....	93
6.4.3 客户忠诚度的评价 .....	95
6.4.4 客户分类结果 .....	95
6.5 市场策略 .....	96
参考文献 .....	97
<b>7 分形理论在虚拟商店评价中的应用 .....</b>	<b>98</b>
7.1 引言 .....	98
7.2 相关理论概述 .....	99
7.2.1 绩效评价的内涵 .....	99
7.2.2 虚拟商店绩效评价要素 .....	101
7.2.3 虚拟商店绩效评价的目的 .....	102
7.2.4 虚拟商店绩效评价的影响因素 .....	103
7.2.5 虚拟商店绩效评价指标选取原则 .....	107
7.3 虚拟商店评价指标的构建 .....	108
7.3.1 虚拟商店绩效评价内容 .....	108
7.3.2 虚拟商店绩效评价指标体系 .....	114
7.4 虚拟商店绩效分形评价模型 .....	116
7.4.1 评价方法的确定 .....	116
7.4.2 模型设计 .....	116
7.4.3 子指标数据的标准化 .....	117
7.4.4 子指标的相关性分析 .....	117
7.4.5 分形评价 .....	118
7.5 实证分析 .....	119
7.5.1 案例背景 .....	119
7.5.2 MATLAB 软件简介 .....	121
7.5.3 实证结果 .....	122
参考文献 .....	131
<b>8 基于熵—灰色关联度电子商务网站评价研究 .....</b>	<b>132</b>
8.1 引言 .....	132
8.2 电子商务网站评价指标体系的构建 .....	133
8.2.1 内容指标 .....	133

8.2.2 商业指标 .....	134
8.3 电子商务网站评价方法 .....	134
8.3.1 熵权法 .....	134
8.3.2 灰色关联度分析 .....	135
8.4 实证研究 .....	136
8.5 结论 .....	138
参考文献 .....	138
<b>9 基于生存分析的互联网用户在线时间实证研究 .....</b>	<b>140</b>
9.1 生存分析 .....	140
9.1.1 基本函数 .....	140
9.1.2 分析方法 .....	141
9.2 数据来源与处理 .....	142
9.2.1 数据来源 .....	142
9.2.2 流失条件界定与生存时间计算 .....	142
9.3 应用实例 .....	142
9.3.1 生存时间分布 .....	142
9.3.2 分组对比分析 .....	143
9.4 影响因素分析 .....	150
9.5 结论 .....	151
参考文献 .....	151
<b>附录 A 客户忠诚度预测测试数据集 .....</b>	<b>153</b>
<b>附录 B 网民/专家调查问卷 .....</b>	<b>188</b>



# 绪 论

## 1.1 研究背景

数据库知识发现, KDD (Knowledge Discovery in Database) 一词首先出现在 1989 年 8 月美国底特律召开的第 11 届国际人工智能会议 (The 11th International Joint Conference on AI)。1999 年亚太地区在北京召开的第三届 PAKDD 会议收到 158 篇论文, 空前激烈。IEEE 的 Knowledge and Data Engineering 会刊率先在 1993 年出版 KDD 技术专刊, 并行计算、计算机网络和信息工程等其他领域的国际学会、学刊也把数据挖掘和知识发现列为专题和专列讨论, 甚至到了脍炙人口的程度。到目前为止, 美国人工智能协会已经主办很多次 KDD 国际研讨会, 规模由原来的专题讨论会发展到国际学术大会, 研究重点逐渐从发现方法的研究转向实际的系统应用, 注重发现多种策略和技术的集成, 以及多种学科之间的渗透。

数据挖掘技术一开始就是面向应用的, 它不仅是面向特定数据库的简单检索、查询调用, 而且要对这些数据进行微观、中观及宏观的统计分析、综合、推理, 以指导实际问题的求解, 企图发现事件间的相互关联甚至用已有的数据对未来的活动进行预测。对于数据挖掘技术的研究, 在国外已经有好多年的历史了。在国外, 数据挖掘技术及相关的决策支持系统发展很快, 已经直接给商业界、公共服务行业等众多行业带来了令人吃惊的利润, 并且有很多学校和科研机构也正投入大量资金进行数据挖掘技术的进一步开发和深入研究。

加拿大 BC 省电话公司要求加拿大 Simon Fraser 大学 KDD 研究所根据其拥有十多年的客户数据, 总结、分析并提出新的电话收费管理方法, 制定既有利于公司又有利于客户的优惠政策。

美国运通公司 (American Express) 使用神经网络检测数以亿计的数据库记录, 辨别个体消费者是如何及在哪里持卡交易的, 得到了每个持卡用户的“购买倾向价值”, 根据这些价值, 美国运通公司将个人持卡者的购买历史与关系销售的商品匹配, 并将这些情况附在月报后面, 这样既节省了费用又提供给持卡者更有价值的分析。

NSRC 是一家位于克里夫兰的市场调研机构, 它介绍了一种数据挖掘工作的情况, 使用了市场调研的成果来找出具有销售潜力的那些消费者中排在最前面的 1% 的消费者, 根据对客户成本分析估计, 这项数据挖掘工作将销售额提高到

501%，将净收入增加了 3587%，这一卓越的成绩之所以取得，由于数据挖掘技术找准了各种消费群体之间的细微差别。

数据挖掘在医学上的应用也很广泛，利用数据挖掘来分析艾滋病的基因，找出 SPN（一种肺癌的前兆症状）的诊断率，分析具有早期乳腺癌 X 光片，达到了较高的准确率，分析肺癌数据库发现了一个有趣的规则，右肺出现肿瘤频率与左肺相比为 3 : 2 等。

目前很多领域数据挖掘都是一个很时髦的词，尤其在证券、银行、保险、零售等领域，数据挖掘所能解决的典型问题是数据库营销（Database Marketing），客户群体划分（Customer Segmentation & Classification），背景分析（Profile Analysis），交叉销售（Cross Selling）等市场分析行为以及客户流失性分析（Churn Analysis），客户信用记分（Credit Scoring），欺诈发现（Fraud Detection）等。在国外市场激烈的环境下，每个市场为自身的生存已经想尽了办法，很多被人工发现的规律早就发现了。

最近几年，国内也有相当多的数据挖掘和知识发现方面的研究成果，许多学术会议上都设有专题进行学术交流。许多科研单位和高等院校竞相开展数据挖掘的基础理论及应用研究，这些单位包括清华大学、中科院计算技术研究所、空军第三研究所、海军装备论证中心等，其中北京系统工程研究所对模糊方法在知识发现中的应用进行了深入研究，北京大学也在开展对数据立方体代数的研究，华中科技大学、复旦大学、浙江大学、中国科技大学、中科院数学研究所、吉林大学等开展了对关联规则开采算法的优化改造，南京大学、四川大学和上海交通大学等探讨研究了非结构化数据的知识发现以及 Web 数据挖掘。

但是国内与国外相比，我国对数据挖掘领域研究仍处于初期阶段，绝大多数工作集中于局部算法设计，有的开始设计软件，但还是处在业务数据转移和建立数据仓库的初级阶段，进行综合的系统集成设计寥寥无几。由于核心技术的欠缺，数据挖掘在国内一些领域只是初步开始应用。虽然在零售业、证券业等行业有所研究，但也只是提出一些应用构思、解决方案，在实现系统方面仍处于初级阶段，还没有对数据进行深一步挖掘、实证研究，所以国内虽然实施了数据挖掘，仍存在一些问题，结果不尽如人意。其原因如下：

- (1) 数据挖掘一定要先确认为什么要投资数据挖掘。
- (2) 要用数据挖掘解决什么问题。
- (3) 如何评价成功与否。

当前，数据挖掘方法已广泛应用于电子商务各个阶段和领域：

(1) 客户获取。即根据性别、收入、交易行为等属性特征把客户细分为具有不同需求和交易习惯的群体，同一群体中的客户在产品需求、交易心理等方面具有相似性，而不同群体间差异则较大。这有助于企业在营销中更加贴近客户需

求。分类和聚类等挖掘方法可以把大量的客户分成不同的类（群体），适合于用来进行客户细分。通过群体细分，CRM 用户可以更好地理解客户，发现群体客户的行为规律。在行为分组完成后，还要进行客户理解、客户行为规律发现和客户组之间的交叉分析。

(2) 重点客户发现。重点客户发现就是找出对企业具有重要意义的客户，主要包括：发现有价值的潜在客户；发现有更多的消费需求的同一客户；发现更多使用的同一种产品或服务；保持客户的忠诚度。根据 20/80（即 20% 的客户贡献 80% 的销售额）以及开发新客户费用是保留老客户费用的 5 倍等营销原则，重点客户发现在 CRM 中具有举足轻重的作用。

(3) 交叉营销。商家与客户之间的商业关系是一种持续的不断发展的关系，通过不断地相互接触和交流，客户得到了更好更贴切的服务质量，商家则因为增加了销售量而获利。交叉营销向已购买商品的客户推荐其他产品和服务。这种策略成功的关键是要确保推销的产品是用户所感兴趣的，有几种挖掘方法都可以应用于此问题，关联规则分析能够发现客户倾向于关联购买哪些商品。聚类分析能够发现对特定产品感兴趣的客户群，神经网络、回归等方法能够预测客户购买该新产品的可能性。

(4) 客户流失分析。分类等技术能够判断具备哪些特性的客户群体最容易流失，建立客户流失预测模型，从而帮助企业对有流失风险的客户提前采取相应营销措施。利用数据挖掘技术，通过挖掘大量的客户信息来构建预测模型，可以较准确地找出易流失客户群，并制定相应的方案，最大程度地保持住老客户。研究证实，数据挖掘技术中的决策树技术（Decision Tree）能够较好地应用在这一方面。

(5) 性能评估。以客户所提供的市场反馈为基础，通过数据仓库的数据清理与集中过程，将客户对市场的反馈自动地输入到数据仓库中，从而进行客户行为跟踪。性能分析、客户行为分析、重点客户发现三者的相互交叠，保证了企业客户关系管理目标的顺利达成和良好客户关系的建立。

## 1.2 研究意义

### 1.2.1 客户分类

随着全球经济一体化进程的加快和竞争的加剧，传统的以生产为中心，以产品和规模为目的的粗放式经营管理模式逐渐被以客户为中心、以服务为目的集约式经营管理模式所取代。客户关系管理（Customer Relationship Management, CRM）正被越来越多的企业所关注。

CRM 成功实施的前提是确定“谁是你的客户”和对客户进行科学有效的细

分。通过客户分类,企业可以更好地识别不同的客户群体,采取差异化营销策略,从而有效地降低成本,同时获得更强、更有利可图的市场渗透<sup>[1]</sup>。

客户分类(Customer Segmentation)是指按照一定的标准将企业现有客户划分为不同的客户群。目前,学术界、企业界广泛认可的客户分类理论是客户价值细分理论<sup>[2]</sup>,但该细分理论存在两大不足:

(1) 客户潜在价值(Customer Potential Value, CPV)的衡量。客户价值细分理论将客户价值分为当前价值(Customer Current Value, CCV)和潜在价值两部分,但在衡量客户潜在价值的时候,往往涉及大量主观感知成分,需要采用问卷调查、直觉判断等手段获得,其度量难以付诸实践,也很难做到客观准确。

(2) 客户忠诚度的忽略。客户价值细分理论在衡量客户价值的同时,往往忽略客户忠诚度对客户价值的影响。一个忠诚度低的客户,即使他拥有很高的当前价值和潜在价值,其总体价值也相对较低,企业如果对其进行重点投入就会带来损失,因为高的客户转换率会使企业的营销努力付之东流。因此仅利用客户当前价值和客户潜在价值两个维度对客户价值进行预测并进行客户分类存在一定的局限性。

随着社会的信息化,企业在日常的经营活动中越来越方便、越来越多地获得了有关客户的大量信息,但却很少对这些数据进行再提炼加工和深入挖掘,获得隐藏其中的规律或信息,企业正逐步陷入“数据丰富,知识贫乏”的尴尬境地。面对如此巨大的数据资源,人们迫切需要一种新技术和自动工具可以帮助我们科学地进行各种决策,数据挖掘技术就是这一类技术。

数据挖掘也被称作KDD(Knowledge Discovery in Database),即数据库中的知识发现,是一种决策支持过程,它主要基于AI(人工智能)、机器学习、统计学等技术,高度自动化地分析企业原有的数据,做出归纳性的推理,从中挖掘出潜在的模式。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、做出正确的决策。

数据挖掘的主要方法包括关联分析、时序模式、分类、聚类、偏差分析及预测等,它们可以应用到以客户为中心的企业决策分析及管理的不同领域和阶段。

数据挖掘可以根据企业大量的客户信息把客户分成不同的类,确定每一类客户的特征,进而调整企业经营策略,有针对性地为客户提供服务,提高客户满意度,维持好优质客户,提高企业的竞争力。

### 1.2.2 网店评价

以计算机和网络技术为主导的高新科技的发展,促进了互联网和虚拟经济的大发展,从而极大地改变了人们购物和结算的方式,一些有别于传统购物方式的

商场应运而生,如虚拟商店、虚拟购物中心、虚拟购物街等。人们不用再像以前那样走很远的路去购物,只需要坐在自家的电脑桌前,用鼠标点击一些虚拟商店的网页便可以找到自己喜爱的商品,然后经过网上结算就可享受到送货上门的服务。随着信息技术和网络技术的不断发展,我们可以预见,在不久的将来,虚拟商店将风靡全球,并且逐步补充传统的店面式经营,从而成为人们购物的另一种方式。

人们在进行网上购物的时候也存在许多顾虑,比如网上结算的安全性、虚拟商店的诚信问题、所购物品的送达时间、货物的实际质量如何等。这些问题也直接影响了消费者对虚拟商店的印象,从而也影响着虚拟商店的业绩,如何评价虚拟商店的绩效已是当务之急。管理大师彼得·德鲁克曾说过:“如果你不能评价,你就无法管理”,可见评价对管理的重要性,尤其是绩效评价更受到管理者的青睐。虚拟商店是有别于一般店面式商场的无店铺经营商店,这种无店铺经营的特点给它的评价带来了一定的难度。目前很多学者对连锁经营、专卖店经营等进行了绩效评价方面的研究,有学者对商务网站绩效评价进行了研究,也有学者对电子商务系统绩效评价进行了研究,但对虚拟商店相关问题的研究还很少,尤其是虚拟商店绩效评价方面的研究更少。在此背景下,本章以探讨虚拟商店的内涵和特点及其相关理论为基础,并构建虚拟商店的绩效评价指标体系,用分形评价法对虚拟商店进行绩效评价。

在进行企业绩效评价方面,有很多学者进行了讨论,但是具体到虚拟商店绩效评价方面,笔者发现还没有人研究过。虚拟商店作为一个新生事物,也势必会成为一个后起之秀,所以对其进行绩效评价是很有意义的。本章考虑到虚拟商店的独特性,在构建评价指标体系的时候从多个角度进行探讨,这些指标为以后的虚拟商店管理者在进行绩效评价的时候可以起到借鉴和参考的作用。同时,本章也引进了一种新的评价方法——分形评价法,将分形、分维评价引入到虚拟商店绩效评价中来,也是一种新的尝试,可以为后来的研究者提供一种参考。

总之,本研究可以帮助想建立虚拟商店的企业或个人对虚拟商店的理论有一个充分的认识,帮助他们在建立虚拟商店的时候能够很好地把握住关键;同时也能使已经运营的虚拟商店在进行自我认识和评价的时候能够有更清楚、透彻和全方位的把握,准确找出自己的优点与不足,从而能更好地发展壮大自己。

## 1.3 国内外研究综述

### 1.3.1 客户分类

Smith Wendell 于 1956 年在探讨市场细分和产品差异这两种不同的营销策略时首先提出客户分类。他认为“客户分类是基于某一时期市场中个体需求的不同

特点而做出的产品决策，而产品差异策略则仅定位于市场竞争者，不考虑需求的复杂性。”<sup>[3]</sup>

从国内外相关文献来看，目前的客户分类是在传统市场细分研究的基础上所进行的更为深入的研究，大体是从客户、企业以及两者相结合这三个角度展开的，其中客户角度和企业角度分别基于客户让渡价值和生命周期价值理论。

### 1.3.1.1 客户价值的内涵

目前关于客户价值的内涵主要分为两大类：客户让渡价值理论和客户生命周期价值理论。

客户让渡价值的概念由著名市场营销学权威菲利普·科特勒（Kotler）博士于1995年在他的著作《营销管理》（第八版）首次提出（这也标志着客户价值时代的来临）。他认为客户让渡价值（Customer Delivered Value）是指总客户价值与总客户成本之差。总客户价值（Total Customer Value）是指客户期望从某一特定产品或服务中获得的一组利益，它包括产品价值、服务价值、人员价值和形象价值等。而总客户成本（Total Customer Cost）是指客户在评估、获得和使用该产品或服务时所引起的预计费用，它包括货币成本、时间成本、精力成本、体力成本等。

纵观有关客户生命周期价值的文献，发现当前的研究对其定义有各种不同表述，先后有许多文献定义过客户生命周期价值。Barbara Jackson（1985）将客户生命周期价值定义为客户当前以及将来所产生的货币利益的净现值；Roberts 和 Berger（1989）、Barbara Jackson（1994）定义客户生命周期价值为客户将来在降低企业经营费用以及增加利润上所带来的收益的净现值；Bitran 和 Mondschein（1996）认为客户生命周期价值是客户在整个生命周期内所产生的净利润的折现值。可以看出在客户生命周期价值的具体涵义上，一种观点是将收益定义为利润流，一种观点是将收益定义为客户在企业降低经营费用和增加利润上的收益，这两种看法其实并无太多的异议，后一种定义将经营费用单列出来的目的只是为了方便探讨这类企业投资对客户生命周期价值的影响作用。关于客户生命周期价值中时间的界定上有较多的偏差，一种看法是将其视为从当前到客户关系解体时的剩余生命周期（Remaining Life）时间段，另外一种看法是将其视为从客户关系的开始直至客户关系解体的全生命周期（Life Cycle）。如何统一在时间上对于客户生命周期价值的认识，Courtheoux（1995）对这一问题作了区分，他用了另外一个概念——客户长期价值（Customer Longtime Value），来区别客户全生命周期价值（或客户终身价值），认为客户长期价值是指客户在未来为企业创造的预期价值，客户终生价值是指客户在全生命周期内为企业创造的价值。按照这一定义，他认为客户终身价值只适合于描述新客户的未来价值，对于企业的老客户，使用客户长期价值更为适宜。从总体来看，客户生命周期价值的定义比较一致，



上述文献实质上是相同的，基本上认为客户生命周期价值是客户在整个生命周期中各个交易时段上的利润的净现值的和，这也是大多数学者所持的观点<sup>[4~6]</sup>。

### 1.3.1.2 基于客户的细分研究

Wilkie 和 Cohen 最早按照不同的层次将细分变量分为五种：个人总体特征描述变量（如性别、年龄、职业、收入等）、心理图示、需要的价值、品牌感知和购买行为。Schiffman 按照地理、人口、心理、社会文化、使用情境、利益以及混合细分变量进行归纳。Halev 则认为在传统市场细分中，地理区域、人口统计和销量细分变量占据了统治地位<sup>[7,8]</sup>。从以上学者对传统市场细分变量的总结不难看出，它们实际上可以归属于三类——环境细分、心理细分和行为细分。

### 1.3.1.3 基于企业的细分研究

客户导向的细分方法是围绕客户各方面差异展开的，目的是实现差异化营销策略。由于差异化必须付出相应的成本代价，过分关注客户需求而忽视企业利益的细分则恰恰违背了市场细分的初衷——更好地集中有限资源为某一客户群体提供差异化服务。于是，相当一部分学者就转向从企业角度出发研究细分方法，其成果集中体现在价值细分上。价值细分的思想就是以客户价值为细分变量，根据客户价值大小将所有客户分为具有不同价值的客户群体。

价值细分最早出现在数据库营销中，其中最典型的的就是 Jin Sellers 和 Arthur Hughes 提出的 RFM 客户分类方法<sup>[9]</sup>。该方法是按照上次购买至今的时期（Recency）、购买频率（Frequency）和购买金额（Monetary）三个要素乘积的大小，对所有客户的交易数据进行排序，前面的 20% 是最有价值的客户，后面 20% 是企业应该避免的低价值客户，中间 60% 的客户是需要向上迁移（Migrate Up）的客户。RMF 细分法的缺点是分析过程复杂，需要花费很多时间，细分后得到的客户群体过多，难以针对每个细分客户群体制订有效的营销策略，并且购买频率与购买金额之间存在多重共线性。为了解决这些缺陷，Marcus 提出用购买次数和平均购买额构造客户价值矩阵，从而将现有客户划分为乐于消费型、最好的客户、不确定型客户和经常性客户四类<sup>[10]</sup>。

随着客户价值的深入研究，一些学者提出应该用客户终身价值（Customer Lifetime Value）来衡量客户对企业的利润贡献，因为它不仅能体现客户的当前价值，还能反映其潜在价值。在此基础上，国内外众多学者都提出了基于客户终身价值的客户分类方法<sup>[11,12]</sup>，即将客户当前价值和客户潜在价值作为客户价值细分两个具体维度，每个维度分成高、低两档，由此可以构造客户价值矩阵，从而根据每类客户价值的大小提出相应的客户保持策略。此外，他们还提出客户潜在价值的预测模型，并通过实证研究证明了模型的有效性。

### 1.3.1.4 将客户和企业相结合的细分研究

从客户的角度进行客户分类满足了不同客户的差异化需求，而利润或价值是