

无形的市场

——知识本体与网络消费研究

唐一之 著

湖南师范大学出版社

无形的市场

——知识本体与网络消费研究

唐一之 著

湖南师范大学出版社

图书在版编目 (CIP) 数据

无形的市场——知识本体与网络消费研究 / 唐一之著. —长沙: 湖南师范大学出版社, 2008. 10

ISBN 978 - 7 - 81081 - 978 - 7

I. 无… II. 唐… III. 知识—本体—关系—电子商务—市场营销学—研究 IV. G302. F713. 36

中国版本图书馆 CIP 数据核字 (2008) 第 153445 号

无形的市场——知识本体与网络消费研究

唐一之 著

◇责任编辑: 柳 丰

◇责任校对: 蒋旭东

◇出版发行: 湖南师范大学出版社

地址/长沙市岳麓山 邮编/410081

电话/0731. 8853867 8872751 传真/0731. 8872636

网址/http: //press. hunnu. edu. cn

◇印刷: 长沙市宏发印刷厂

◇开本: 787 × 1092 1/16 开

◇印张: 8. 25

◇字数: 190 千字

◇版次: 2009 年 1 月第 1 版 2009 年 1 月第 1 次印刷

◇书号: ISBN 978 - 7 - 81081 - 978 - 7

◇定价: 16. 00 元

目 录

绪 论	1
第一章 网络消费理论基础	13
1.1 网络消费内涵分析	13
1.1.1 网络消费的界定	13
1.1.2 网络消费特点分析	14
1.1.3 网络消费方式的内在结构	17
1.1.4 网络消费行为的特征研究	18
1.1.5 电子商务与网络消费的关系研究	20
1.1.6 网络消费方式对电子商务的效能分析	21
1.2 网络消费行为影响因素分析	22
1.2.1 消费者行为和影响因素	22
1.2.2 网络消费和消费者行为	23
1.2.3 传统的和网络的影响消费者购物因素的比较	25
1.3 基于现代计算技术的消费者行为研究	25
1.3.1 RFM 模型	25
1.3.2 消费者价值分析	27
1.3.3 RFM 方法评价	28
1.4 KOB-NCB 构架及属性分析	28
1.4.1 KOB-NCB 层次模型	29
1.4.2 KOB-NCB 过程模型	31
第二章 KOB-NCB 基础数据的知识挖掘研究	33
2.1 KOB-NCB 对网络消费痕迹的数据挖掘研究	33
2.2 KOB-NCB 的用户发现理论	34
2.2.1 用户行为的 WEB 日志挖掘	34
2.2.2 用户发现所采用的几种挖掘算法	37
2.3 KOB-NCB 架构对商品和网页的描述	38
2.3.1 描述方法研究	38
2.3.2 网络产品和服务概念 - 属性矩阵(CA 矩阵)	39
2.3.3 网页语义相似度	39
2.3.4 挖掘算法	40
第三章 KOB-NCB 架构知识本体构建	43
3.1 网络消费行为知识表示方法	43

3.2 本体基础理论	44
3.2.1 本体的定义	44
3.2.2 本体的描述语言 OWL	45
3.2.3 本体的建模元语	47
3.2.4 本体的类型	48
3.3 知识本体构建	48
3.3.1 知识本体的构建方法	48
3.3.2 知识本体的构建工具	49
3.3.3 知识本体构建过程	51
3.4 概念抽取与领域概念关系	52
3.4.1 领域语义词典生成	52
3.4.2 基于 WEB 页面网络消费概念抽取方法	53
3.4.3 基于 HowNet 的网络消费领域概念关系	55
3.5 知识本体实例	60
3.5.1 知识本体的领域确定	60
3.5.2 知识本体类获取方法	61
3.5.3 知识本体模型	63
第四章 KOB-OCB 架构知识本体映射	65
4.1 知识本体映射方法	65
4.1.1 本体的异构性	65
4.1.2 本体映射的定义	65
4.1.3 本体映射的方法分类	66
4.2 知识本体映射过程	67
4.2.1 实体语义相似性的多匹配策略(Mulit-matching Strategy)算法	67
4.2.2 数据源与局部本体之间的映射	67
4.2.3 全局本体与各局部本体的映射	68
4.3 知识本体映射规则	68
4.3.1 相似度的定义	68
4.3.2 映射发现策略	69
4.3.3 本体映射系统的特点	72
第五章 基于 KOB-OCB 架构的需求与个性化服务	74
5.1 原理	74
5.1.1 基于知识的推荐技术的发展	74
5.1.2 基于知识的网络消费个性化服务的基本原理	75
5.2 知识表示方法	78
5.2.1 Recommend Onto 本体模型	78
5.2.2 规则和案例知识的表示方法	82

5.3 知识获取方法	83
5.3.1 知识获取方法分析	83
5.3.2 商品知识的获取方法	84
5.3.3 用户知识的获取方法	90
5.4 推理机制与推荐结果的产生	99
5.4.1 系统推理机制	99
5.4.2 推荐结果的产生	110
5.5 策略的自适应选择	111
第六章 总结与展望	116
6.1 全文总结	116
6.2 未来研究展望	116
参考文献	117

绪 论

一、选题背景和意义

互联网的普及、网络技术和信息技术的发展给社会提供了一个全新的活动空间。在这个由互联网组成的虚拟空间中,人类的消费也呈现出质的飞跃。这便是被学术界广泛研究的网络消费。

网络消费具有消费内容丰富,消费特征个性化,消费与享受一体化、全球化和非实物性等与传统消费模式不同等特点。随着网络技术的日新月异,网络本身也逐渐完成着在全球的铺设和延伸。数量巨大的、多样化的信息像潮水般涌进人们的感性世界。人类智慧所能涉及的全部内容都能以信息的方式展现在网络世界中。人们不仅可以对已有的网络信息坐享其成,而且还可以凭借创意活动为自己或他人创设、提供信息。科技的发展,促使网络消费呈现出极具个性化色彩的显著特征,信息的供应者和消费者之间的个体互动关系更加密切。从单纯的消费角度而言,面对纷至沓来的众多信息,消费者完全根据个人的兴趣、爱好选择自己最钟情的内容。从信息的供应者角度而言,只有根据消费者的个体需求强化信息的针对性,才可能达到预想的效果。于是,从信息的搜集、加工到发布都鲜明地烙上了个性化的印痕。不仅如此,消费者和供应者的角色与传统意义上的角色内涵也有了极大的不同。在网络经济中,双方融合在一起,共同扮演着“产消者”这一崭新的角色。在网络技术和信息加工技术的促动下,信息高速公路伸向全球的每一个角落,人们的消费全球化了。任何人都可以在“信息高速公路网络中和任何国家、任何地方的人直接沟通,能够在全球内实现知识共享”^[1]。人们不仅可以享受本国的各项创造成果,也可以享受其他国家的各种成果。网络技术的进一步发展,将会极大地加速这一趋势。网络消费不同于传统消费,网络消费的对象是信息。因而,这种消费是一种非实物形态的消费。首先,消费者免费或付费便可以享受对网络信息的浏览,甚至下载;其次,消费者对信息的消费仍停留在非实物阶段,不管消费者选中哪一种产品或服务;再次,这种非实物性消费体现着一种观念性的投资型智能消费。

从 1991 年 Internet 开始出现商业化,互联网伴随着技术的日益成熟和功能的逐步强大,以其简单易用飞入寻常百姓家。目前互联网已开发出了以下日常功能,它们也构成了网络消费的主要平台^[2]:①电子邮件即 E-MAIL,是依靠网络的通信手段实现文件信息的传输,可以瞬间传到全球的任一角落。②万维网 WWW 信息浏览,可以接受文本、图形、音频、视频等多种格式的信息文件。③文字和文件传输,将文件从一台计算机传送到另一台计算机,而不受这两台计算机所处位置、联接方式以及所用操作系统的约束。④远程登录系统,其主要功能是让用户本地的计算机与远程计算机相连,成功登录后,就可利用远程计算机的资源进行远程计算机允许的所有操作,就像在本地使用计算机一样。⑤电子公告板系统 BBS 和新闻讨论组 USENET。随着新的一类互联网技术 WEB2.0 的应用,由 WEB1.0 单纯通过网络浏览器浏览 html 网页模式向内容更丰富、联系性更强、工具性更强的 WEB2.0 互联网模式的发展已经成为互联网新的发展趋势。新的 WEB2.0 技术主要包括:博客(BLOG)、RSS、百科全书(Wiki)、网摘、社会网络(SNS)、P2P、即时信息(IM)等。因此,可以看出,随着技术的不断进步,互联网庞大的信息量正在使用户面临“信息过载”的问题。而且,新技术的应用也使得网络消

费的种类变得越来越丰富。

如此庞大的信息量使得 WEB 用户面临“信息过载”(Information Overload)的问题,如何方便快捷地从中获取所需信息是众多 WEB 用户关注的焦点,也成为信息技术的一个研究重点^[3]。WEB 信息如此快速地膨胀导致传统的信息检索(IR)技术面临巨大的压力。传统信息推荐相关的检索技术只能对结构化的关系型数据提供管理和查询手段,无法有效处理大量的非结构化信息,如 WEB 页面、图片、音频、视频等(这类信息已占整个信息量的 80%以上)。而 WEB 搜索推荐技术的出现为解决此问题提供了一个较好的解决方案,此技术现已成为检索 WEB 信息最主要的手段,其典型实例就是我们日常所使用的搜索引擎。另外,随着电子商务的发展,电子商务网站在为客户提供越来越多选择的同时,其结构也变得更加复杂。一方面,客户面对大量的商品信息束手无策,经常迷失其中;另一方面,商家也会逐渐失去与消费者的直接联系^[4]。而电子商务商品推荐系统通过模拟商店销售人员向客户提供商品推荐,帮助客户找到所需商品,从而顺利完成购买过程,因此可以有效保留客户,提高电子商务系统的销售;商家也可以通过推荐系统不断保持与客户的联系,重建客户关系。目前几乎所有的大型电子商务系统,如 Amazon、CDNOW、eBay、当当网上书店等都不同程度地使用了电子商务推荐技术。

据 CNNIC 发布的统计报告显示,截至 2006 年 12 月 31 日,我国上网用户达到 1.37 亿,上网人数占中国总人口数的比例已经超过了 10%。调查机构 AC 尼尔森最新研究显示,63% 的中国网民曾在网上购物。随着网民的急剧增长,中国正成为网上购物的大市场^[5]。可见,中国的网络消费有非常大的用户群,网上购物市场巨大,网络消费的发展很有潜力。因此,了解消费者的网络消费现状及特点,有利于企业更好地针对顾客需要及时调整营销策略来满足消费者在网上消费的需求,促进 B2C 电子商务的发展。此外,互联网信息量的巨大和服务种类、方式的繁多为网络消费者提供了更多的选择领域,同时也带来了另一个问题,那就是消费者如何更有效率地使用互联网,即为消费者提供个性化的网络消费服务推荐。目前这一研究领域,正成为学术界和产业界的研究热点。所以本论著的选题具有重要的理论意义与实践价值。

二、国内外相关研究的现状综述

1. 国内外关于网络消费的研究

关于网络消费的具体研究,目前学术界所涉足的主要是网络消费的“新经济”特征、信息消费的历史地位以及网络消费的伦理规范等理论与实践问题。

(1) 网络消费的“新经济”特征

人的生存形式实际上是在一定生活空间中呈现出的具体消费方式,而生活空间的扩展又是以生活工具的进步为基础的。因此,互联网这一生活工具的出现必然导致网络消费的兴起。对于网络消费的研究,信息经济学是在信息消费的概念框架内进行的。从较早的文献看,信息消费并未成为信息经济学的重要概念,而仅是在信息商品和信息市场研究中兼有论述。如在信息商品理论中,认为信息的消费与信息成本无关,也与信息使用数量和范畴相互独立,这是信息作为商品区别于其他物质商品的显著经济特征之一^[6]。实际上,实物消费、劳务消费、信息消费这三种消费形式从一开始就并存,只是在最初的经济生活中,实物消费占据主要地位,另两种处于依附地位。工业革命后第三产业兴起,劳务消费也随之壮大并摆脱了

对实物消费的依附地位。而从 20 世纪中叶开始的信息革命使信息产业从第三产业中独立出来,相应地,信息消费也从劳务消费中分离出来而成为独立的一类消费形式。随着多媒体网络技术的发展,信息消费又开始具有了数字化、无形化、网络化、虚拟化、全球化、自由化的新特点^[6]。此时,消费行为在很大程度上得到了改进,消费者通过计算机终端设备所发生的交互式消费行为将占据主导地位^[7]。由于信息爆炸,没有控制和组织的信息不再是一种资源,它反倒成为消费者的敌人^[8],因此,可以选择的信息越多,便越难做出选择^[9]。可见,信息消费要求消费者具有一定的信息消费(能)力,即必须具备与信息消费相适应的知识、经验、理解力和可能占有的条件。信息消费力决定信息消费的对象水平^[10]。将信息消费或网络消费视为一种必然趋势,是学术界居主流的理论观点。法国学者让·雅克·塞尔旺·施赖贝尔认为:工业社会的信息化就如同农业社会的工业化一样,不仅根本改变生产方式和消费方式,而且根本改变生活方式^[11]。爱尔兰专家利亚姆·班农对信息社会的消费生活进行了细致的分析,并预测购物活动等消费行为将受到信息技术越来越大的影响^[12]。

当今世界正在由“物质经济”向“知识经济”过渡,信息消费和网络消费的到来是不可避免的。实际上,这种趋势在发达国家已经初见端倪。学术界较一致的看法是,网络消费不同于传统消费,而有其自身的“新经济”特征。亚当·斯密关于市场的观点是:每一个买者知道每一个卖者的价格,每一个卖者知道每一个买者想付多少钱。斯蒂格勒(G.J.Stigler)在其著名论文《信息经济学》^[13]中,认为这种“完全信息假设”是不存在的;尔后,阿罗(K.J.Arrow)以其一系列论文^[13],维克里(W.Vickery)和米尔利斯(J.Mirrless)以其“不对称信息下的激励理论”使信息经济学先后三次捧得诺贝尔经济学奖。网络经济的兴起,使许多人看到了解决信息瓶颈在技术上的可能性和可预期的美好前景。他们认为,网络“会导致一个没有摩擦”、头脑较为清醒的资本主义新世界,市场信息极为丰富,交易费用很低。这将是消费者的天堂^[14]。由于有了“更加完美的市场”,专家们认为网络经济将表现出“有区别地生产”和“有个性地消费”的新经济特征。在这里,技术可被用于实现更高程度的个人化,更大程度地满足个人需要——这就是个人市场原理的全部内容。

此时,“个人化、个体化和个人市场这些观念逐渐深入人心,而提高效率、降低成本和节约资金这些观念正在悄然逝去”^[15]。姜奇平称之为“直接经济”。他认为:“工业社会生产方式的本质性特征,是迂回生产,通过拉长生产和消费之间的中间链条来取得价值回报;而信息社会生产方式的本质性特征,是缩短生产和消费之间的中间环节,通过快速贴近目标来获取价值。”数字化网络对时空的这种“合力”所产生的生产力,正好为削减不必要的中间环节提供了现实可靠的基础^[16]。

专家们认为,网络消费相对于传统消费而言将对消费者更加有利。哈格尔三世和阿姆斯特朗对微观经济学中典型的供求曲线进行分析后认为,网络消费中市场价格将更靠近供应曲线,即经济活动中的剩余价值将更多地转移到消费者手中^[17]。不仅如此,由于高技术因素所导致的“学习效应”的普遍存在,网络消费还可能步入一个“报酬递增的新商业世界”^[18]。

我国的网络消费启动较晚,但具有强劲、快速的发展特点,这反映了“后发”国家在“示范效应”作用下不断膨胀的信息需求。中、日、韩三国学者所做的一项研究表明,信息化进程已经开始改变中国城市居民的消费生活,使其正向“信息人”转化,对商品性能的消费也逐渐转变为对符号的消费^{[19][20]}。有研究表明,中国城市居民信息化消费方式的兴起与发展,1998 年是启动之年,2000 年初见规模,而 2005 年则得到了较大的发展^[21]。

著名电子学专家罗沛霖院士指出：“产业革命时期技术革命直接作用于人的物质、经济生活，从而间接作用于精神、文化生活，而信息技术革命则直接作用于人的精神、文化生活，并以更鲜明的姿态间接作用于物质、经济生活。”可以预期，网络消费在未来社会将如同今天的衣、食、住、行一样成为民众的一种基本需要。为此，要特别关注互联网发展中网络消费的兴起与演进过程，以构建网络消费的理论模型，进而揭示其发展变化的微观机理和宏观规律。

(2) 信息消费的历史地位

对于网络消费的研究，信息经济学是在“信息消费”的概念框架内进行的。从较早期的文献看，信息消费并未成为信息经济学的重要概念，而仅是在信息商品和信息市场研究中兼有论述。而从 20 世纪中叶开始的信息革命使信息产业从第三产业中独立出来，相应地，随着多媒体网络技术的发展，信息消费也从实物消费、劳务消费中分离出来而成为独立的一类消费形式^[23]。

(3) 网络消费的道德规范

也许是由于实践的需要，学术界对网络消费问题的讨论较集中于“网络伦理”和“网络规范”领域，较早的研究者有温伯格、帕克、皮尔斯、维特尔、约翰逊等人。虽然有人认为使用网络是一种自由的权利，但国外研究者普遍的看法是，网络不仅仅是一个简单的网络，它更是一个由成千成万的个人组成的网络“社会”，为此国外一些计算机和网络组织制定了一些初步的伦理规则，比较著名的有美国计算机伦理协会的“十条戒律”、南加利福尼亚大学网络伦理声明等^[24]。美国网络伦理研究者巴格(R.N.Barger)认为，在当今伦理困境正在变得越来越复杂的计算机世界中，寻找一种简单的、每一个人都赞同的标准的道德规范的希望是比较渺茫的，然而这并不意味着做这种努力是无用的，对具有不同哲学世界观的人来讲，同意相同的标准——尽管出于不同的理由——还是有可能的。他进一步指出，以前的技术被取代了，但基本的行为准则依然可以运用于目前的技术革命。据此，巴格提出了三条基本原则：第一，一致同意的原则，如诚实、公正和真实等；第二，把这些原则运用到对不道德行为的禁止；第三，通过惩罚并且通过对遵守规则行为的积极鼓励来加强对不道德行为的禁止^[24]。可以预期，被人类社会实践所证明了的合理的伦理道德准则，其内核应该得以保留，改变的只是它在网络时代的具体形式。

本类研究中，除了上述一些经典的里程碑形式的研究成果外，国内还有大量对信息化进程中诸多实践问题的分析。乌家培、张维迎、谢康、靖继鹏、刘建、陈仁俊、朱幼平、金建、左美云、何明升等人均做出了一定的贡献，其中，较为集中的是关于信息产业问题的研究^[25]。

目前，学术界对网络经济的研究已经逐步展开并不断深入，但网络消费的探讨仍是网络经济研究中的一个薄弱点，并且都集中在对信息产业的讨论。新的经济现象必定需要理论来支撑，但理论界还没有对网络消费的理论引起足够的重视，相关的研究成果还不多见。所以，从网络消费知识本体的角度研究网络消费的相关理论和实践问题具有重要的理论意义与实践意义。

2. 关于网络消费知识的相关研究

当前，对于许多在因特网上从事网络销售的企业而言，如何创新和发展电子商务的经营与管理模式，系统性地构建客户关系，提升顾客满意度及忠诚度，更有效地吸引新的顾客，赢取利润以保持企业竞争优势，是急需解决的重大问题。

网络环境下的客户关系管理的理论研究多源于美国等电子商务发达国家，其概念和理

论能否移植到其他国家和文化环境仍有待检验。在基于 B2C 电子商务的消费心理管理方面,国内的实证研究并不多见,很多研究只是基于初步调查的经验总结和感性判断。因此,借鉴国外已有的研究成果,在中国文化背景下科学地研究网络消费心理及消费忠诚度的驱动因素,总结网络背景下消费者心理行为的一般性规律,并用计算机语言来表述和应用,对电子商务企业制定正确的营销策略和促进网络消费市场的繁荣发展具有重大意义。

但是由于网络消费知识表示的不统一,网络消费知识存储的分布性等问题使得网络消费知识之间的协同和相互操作变得困难,进而导致 WEB 上丰富的网络消费资源得不到有效利用,不能很好地为电子商务的发展提供支持。本研究以经营 B2C 电子商务的网络信息资源为研究对象,拟以实体环境下的关系质量模式(Crosby et al, 1990)为基础,结合消费价值理论(Sheth et al, 1991)和交易成本理论(Williamson, 1985)来探讨电子商务中网络消费心理及消费忠诚度的驱动因素。本研究将电子商务消费价值提炼为三个维度,即功能性价值、程序性价值和社会性价值。其中,功能性价值由价格优势、购物便利性、产品质量、信息质量、服务失误补救等具体因素构成;程序性价值由网站设计、操作便利性、个性化服务、购物娱乐性、网络安全性、隐私保护和网络互动性等因素构成;社会性价值由网站声誉、社会逃避价值、B2C 关系价值、C2C 关系价值等因素构成。这种消费价值的分类方法为网络消费知识的本体结构提供了很好的借鉴。

根据 Jupiter Research 一份最新的报告《2005—2010 美国网络零售预测》研究结果指出,在 2006 年之后,网络销售成长的关键,不再是获取新消费,而是想办法增加现有消费的花费。因为网络零售业者发现要改变非网络消费者的习惯及想法不是一件容易的事,新的网络消费者愈来愈难开发。因此,网络零售业者未来将愈来愈依赖现有的网络消费者,维持目前的消费并想办法提高消费花费比开发新消费更为重要。因此,网络零售业者面临的挑战不再仅仅是价格竞争,而是更广泛的营销竞争,想办法提高消费购物满意度和忠诚度,消费忠诚度才是电子商务成功的关键。对于购物网站的经营者来说,使用优良的软硬件设施,固然可以提高网站的经营效率,但真正要创造出网站本身特有的价值,则必须取得消费的信心和忠诚,也就是说,着眼于消费关系研究而不是技术研究,才是电子商务成功的关键。因此,对于电子商务企业来讲,必须向消费提供和传递良好的服务质量和卓越的消费价值,通过提高网络消费者的满意度和信任,才能建立和维系消费忠诚度。这是电子商务企业必须充分了解网络消费者心理的重要原因。

目前理论上和现实中面临的主要问题就是网络营销商的商品推荐和网络消费的消费动机与期望不能有效匹配,买卖双方存在严重信息不对称。如何有效利用网络技术和信息技术,系统性地建构消费关系,使电子商务企业更有效率地实现消费满意及忠诚度的提升与企业利润的赢取,创新电子商务的经营与管理模式,是因特网背景下管理科学最重要的研究课题之一。而构建网络消费心理知识本体就是对获取网络消费满意度和信任的有效路径的有益探索,目前国内外的相关学者在这方面的研究成果较少。最近几年相关的研究成果主要体现在以下几个方面:

(1) 网络消费心理的相关研究

Jack A Lesser et al (1998)^[26]测试了与三个购买相关的心理变量模型:假设发展表述模型、假设部署模型和假设危险模型,并对这三个模型进行了整合。Giese et al (2000)^[27]基于文献研究和消费者观点,提出一个消费者满意的定义框架。Steven S. Posavac et al (2002)^[28]研究了消

费者选择购买的机会对购买倾向的影响效果。J. Goodman et al(2003)^[29]提出消费者抱怨提供了很多有价值的信息,产生问题的经历导致的不满可以引起持续的负面的口头信息传播行为。罗子明(2003)^[30]指出,消费者心理研究已经成为我国心理学研究领域中国市场规模最大、涉及人数最多的一个分支,它未来的发展方向包括专业著作的成熟、抽样技术的完善、消费者研究数据库的系统化与商业化。制约电子商务发展的心理因素包括传统购物观念的束缚、价格预期心理得不到满足、个人隐私权受到威胁、对网上支付机制缺乏信任感、对虚拟的购物环境缺乏安全感、对低效配送缺乏保障感(李忠艳,2004)^[31]。肖崇好(2004)^[32]论述了网络环境的特点及其对网络用户心理的影响。柯青(2004)^[33]阐述了网络消费者的购买动机包括需求动机和心理动机。李东进等(2005)^[34]通过实证分析发现搜寻动机和自我效能感直接影响 WOM 信息依赖性,搜寻动机与购买卷入有正相关,与创新性有负相关,自我效能感与购买卷入、认知欲望之间有正相关。Katherine W. et al(2006)^[35]就分离式群体指导对消费者心理的影响进行了相关研究。Barbara Loken(2006)^[36]就消费者心理所涉及的分类型、影响及说服问题进行了研究。

(2) 网络经历 (WEB experience) 的相关研究

网络经历要素的识别和分类对网络的互动研究至关重要,它可以帮助营销实践者去认识和理解网络营销的本质与潜在的工具(Constantinides, 2004)^[37]。

Tamimi et al(2003)^[38]将网络消费经历定义为描述网络交易的连续步骤的四个阶段过程。Watchfire Whitepaper Series(2000)^[39]将网络经历定义为消费者对网络公司的总体印象。网络经历包含的要素例如和网络公司的互动以及交易中的搜索、浏览、发现、选择、对比和信息评估等行为。Constantinides(2004)^[37]提出,网络营销人员的可控因素可以分成三个大类:功能因素(Functionality Factors)、心理因素(Psychological Factors)、内容因素(Content Factors);五个小类:效用(Usability)、互动(Interactivity)、诚信(Trust)、美观(Aesthetics)、营销组合(Marketing mix)。这些研究成果对本研究的研究提供了指导。

(3) 网络消费行为的相关研究

在对关于网络消费者行为这一主题大量的研究文献进行分类方面做出重要贡献的是 Cheung et al(2003)^[40],他们提出了两组不可控的因素——消费者特点和环境影响,三组可控因素——产品和服务的特点、媒体特点选择以及合作对象的选择。可控的和不可控的因素的相互作用也深深地影响着网络决策过程(O'Cass and Fenech, 2003)^[41]。

相关的理论成果主要体现在以下几个方面:

第一,从消费者角度分析了影响消费者网上购物的因素。过去几十年的研究成果描述了消费者的购买过程所经历的五个步骤:问题识别、信息搜索、分项评估、购买决策、购买后的行为。营销人员仍然相信 4P 策略是影响消费者行为结果的有效工具(Bettman, 1979; Dibb et al., 2001; Jobber, 2001; Boyd et al., 2002; Kotler, 2003; Brassington and Pettitt, 2003)^[42]。然而这种作为影响工具的组合价值和声誉经常受到质疑(Grönroos, 1994)^[43]。Pui-Mun Lee(2002)^[44]强调了安全和隐私保护的重要性。

Tybout et al(1994)^[45]对消费者的信息利用及其决策影响进行了检验,提出了“中心与使用边缘”的概念。Magnus Soderlund(1998)^[46]通过口传信息、对供应商的反应以及品牌的忠诚等三个维度对消费满意和消费行为之间的关系进行了研究。Susan Sproule et al(2000)^[47]利用决策支持系统领域的营销和自学习模型构建了一个软件代理系统的设计框架来为电子商务购

买的应用研究提供支持。Marios Koufaris(2002)^[48]认为网络消费者不仅是购买者,而且还是计算机使用者。Ogden et al (2004)^[49]从主文化、次文化和文化传入维度研究了美国种族文化对消费者购买决策的影响,提出了微观文化的影响机制的研究模型。Pimentel et al (2004)^[50]通过消费者与购买前持续行为相关的感情承诺为主线,研究了消费者的卷入模型。

国内学者在相关领域也进行了大量的研究,提出了影响网络消费者购物行为的因素包括人口统计因素、网络的可靠性和安全性、心理、价格、进入网站的便捷性、产品的独特性以及企业形象、网页设计风格、电子商务企业特性等(王希希,2001;伍丽君,2001;黎志成等,2002)^[51],并且还探讨了消费忠诚、网络道德、消费者隐私权保护与电子商务经营者利益的平衡等问题(彭军锋,2003;马荣贵,2002;吴满意,2002)^[52]。

第二,从网络消费者的整个购物过程和相关的路径关系进行了分析。当代的大量研究关注于网络购买和决策过程模型(Miles et al., 2000; Liu and Arnett, 2000; Cockburn and McKenzie, 2001; Liao and Cheung, 2001; McKnight et al., 2002; Joines et al., 2003; O'Cass and Fenech, 2003)^{[53][54][55][56]}。Irene S.Y.Kwan 提出了一个针对网络消费者角色的语义网络心理认知模型^[57]。国内学者黎志成、刘枚莲(2002)^[56]从理论上建立了一个电子商务环境下的消费者行为模型。

第三,从网络消费的测度与方式角度进行分析。例如,何明升(2002)^[58]提出了测度网络消费的四种具体方法,即支出法、系数法、结构法、指数法等。网络消费方式是由网络消费行为、网络消费意识、网络消费资料三个结构要素组成的稳定结构,三个结构要素之间存在着互为依存的功能耦合关系。

第四,从企业网站建立的角度进行分析。例如 Jaeki Song(2001), Solomon Negash(2001)等人从纯粹的网站设计元素角度分析了影响网络消费者的网站元素。

关于在线消费行为的研究是基于传统环境下消费行为规律的,网络环境下消费者行为的研究有两大支流:一是研究网络环境下消费者感知网站质量和感知价值的决定因素,二是对网络环境下消费满意与其行为后果之间关系的本质探讨。但从总体上讲,因特网在商业中的应用改变了消费者行为和企业经营模式,尽管以网络环境下交易的巨大商业潜力吸引了很多研究者探讨在线交易的消费者行为,但总体上这个领域的研究还处于起步阶段,电子商务企业的消费忠诚研究并没有得到与实体商店同等程度的关注 (Balabanis, Reynolds & Simintiras, 2006)^[59]。

(4) 本体论及其相关研究

本体论是近来人工智能领域研究的一个热点,被视为是建立智能系统(或狭义的知识表示系统)的世界观,是对应用领域概念化的说明(M Cannataro, et al, 2002)^[60],通过对本体论的研究可以在电子目录得到一致的、形式化的支持知识处理的信息模型,解决语义方面的问题(M Cannataro, D Talia, 2003)^[61]。

本体已经成为人工智能、知识表示、语义 WEB、数据集成、信息检索等研究领域的热门课题。本体技术能实现电子商务知识共享和重用,故可以用来解决语义层次上协同电子商务知识的共享和交换。在分布式的电子商务环境中,同一种商品知识或客户知识由不同的本体表示而导致差异性,因此本体映射在企业之间的知识集成、协同电子商务的知识共享等方面十分重要(H.Zhuge, 2004)^[62]。

现有研究成果已提出的本体映射方法,按本体定义模型进行分类,可以分为:基于语法的方法;基于概念实例的方法,典型的如华盛顿大学的 GLUE 系统;基于概念定义的方法;基

于概念结构的方法。这里每种方法参照了本体定义的不同部分。本体的映射类型有概念-概念、属性-概念、属性-属性、情境和约束等(H.Zhuge, 2004)^[63]。国外已存在的本体映射工程如 InfoSleuth 的参考本体、斯坦福大学的本体代数、AIFB 的形式概念分析、ECAI2000 的方法、ISI 的 OntoMorph、KRAFT 的本体聚类。

本体映射最重要的过程就是发现语义关联。语义关联包括语义相似度和语义相关度的研究,它们是两个不同的概念,语义相似度反映两个概念相互替换的程度,可用于本体之间相似度计算以达到集成相同商品知识的目标。

但是目前,本体映射仍然需要领域专家的干预,存在着网络环境大规模电子商务应用本体映射的精度与效率问题,而且由于语言、文化等方面的差异,本体的自动映射还是一个研究难点(WU Zhaohui, 2003)^[64]。

综上所述,现有的网络消费心理知识具有分散性、异构性、缺乏语义支持和个性化服务的特点,不能满足网络环境下大规模的电子商务交易和供应链集成对商品知识的需求。因此,随着网络消费心理知识本体的相关研究的发展,将促进我国电子商务消费市场买卖双方信息的共享和重用,这正是我国当前建立统一规范的电子商务市场迫切需要解决的理论问题。所以研究基于本体的网络消费心理知识的语义表示和服务,具有重要的理论研究和实际应用价值。

另外,本研究探讨了利用面向网络消费心理知识应用语义本体的技术来解决网络知识共享的理论和方法,也推动了本体技术在解决现实问题中的应用。

3.国内外关于个性化服务系统的研究

目前,国内外研究领域主要从 WEB 搜索推荐、站内导航推荐和电子商务推荐三方面进行相关 WEB 信息推荐技术的研究。

(1)WEB 搜索推荐

WEB 搜索推荐是指从海量且无序的互联网页面中,依据用户搜索查询,定位相关信息并向用户推荐。这方面主要是关于互联网搜索、搜索引擎相关技术的研究。WEB 搜索推荐主要实现方式是采用人工或自动的方式去搜寻 WWW 上的信息,并将网络信息进行主题分类或建立索引,再把索引的内容存放到索引数据库中,通过查询语法为用户返回匹配资源,以帮助用户在网络中搜寻到所需的信息。所涉及技术有页面内容的分析、网页超链接分析、页面存储和检索等。

随着 WWW 上信息的急剧增长,用户想从如此巨大的信息资源上获取其真正感兴趣的内容变得愈来愈困难,基于搜索引擎技术的推荐服务系统正好解决这个问题,通过它用户可以在网络上快速定位、检索感兴趣的信息。WEB 搜索引擎是一种基于 WEB 的信息检索工具,用来检索 WEB 网站、WEB 网页、新闻组、图片、声音等 WEB 文件。现实运行的搜索推荐系统有 Google、Yahoo、MSN、百度、搜狐等,大大便利了普通用户对网络信息的查询。这些系统实际上应称为专业的 WWW 服务器,也可以说是互联网上提供信息检索服务的网站。另有一些系统另辟蹊径,通过对用户兴趣建模来实现信息的搜索推荐服务,例如:WEBWatcher 系统采用跟踪用户浏览 WEB 站点的行为、访问路径方法,分析用户的访问模式,将用户可能感兴趣的页面在线推荐给用户^[65];SiteHelper 系统采用分析每一个用户已经访问过的 WEB 页面,学习用户的兴趣模型,通过分析页面被浏览的时间和频率来识别页面的受欢迎程度,从页面中提取关键词形成用户关键词表,提供给系统并以相关反馈的形式实现推荐^[66];Foot-

prints 系统则利用可视化技术,为用户显式提供 WEB 频繁访问路径。

目前,搜索引擎已成为向 WEB 用户推荐信息的一个重要手段,但是随着互联网的发展,它的诸多缺陷逐渐显露出来,主要表现为以下几个方面:

①非语义化查询。现有的搜索引擎对信息内容的分析采用词形匹配的方式,这就造成在很多情况下,用户很难选择匹配的词或词组来真实表达其所真正需要检索的内容。这也是众多用户在使用搜索引擎时最常遇到的问题,即找到了结果,但不是自己所需要的内容。

②个性化问题。用户查询信息往往是带有自己特别的需求,就算使用同样的检索词,也有不同的个人兴趣偏好。而现在的搜索推荐技术对于任何用户均采用无差别的推荐模式,这更谈不上不同知识层次、不同专业、不同地理位置的个性化推荐服务。

③权威偏见问题。在 WEB 上有一些权威性的资源,它们常常占据了搜索推荐列表(搜索引擎返回结果列表)的前几位,而同时一些新诞生的但是又相对重要的信息却因为加入时间较晚而被排在搜索引擎推荐列表末尾的位置,这就是当前搜索研究领域所面临的“权威偏见问题”或“评分偏见”问题。这是由于当前 WEB 信息快速、动态、增量更新所造成的。

(2) 站内导航推荐

这是网络挖掘的一个重要应用。通过挖掘用户访问网站的模式,使网站设计者可以清楚地了解自己网站页面的整体访问情况,为改善网站的逻辑组织结构提供决策支持,同时又可以明确自己网站内不同用户的兴趣,建立自适应网站或个性化网站。目前采用的主要技术为利用 WEB 使用挖掘技术对网站用户的访问行为进行分析,以获得描述用户访问网站模式的内在规律,为网站的建设者改进网站结构和内容提供决策支持,从而实现在 WEB 站点内为用户提供导航服务。这方面的研究技术主要包括对网站日志的挖掘、网站页面结构的优化分析等。

存储在 WEB 站点上的用户访问记录中包含着极有价值的用户使用、兴趣等相关信息,从中可发现新的知识、实现网站的有效管理,从而达到更有效地向用户提供信息服务,即网站站内导航推荐的目的。但是,由于目前大多数站点的设计强烈依赖设计者的经验,不免产生难以发现和定位的设计失误。因此,就迫切需要使用一定的技术对站点设计进行优化,以发现其中存在的问题。目前的站点优化技术中,基于 WEB 使用挖掘技术的站点优化是一个重要的研究方向。

早在 1996 年国外学者 H.Mannila, M.S.Chen 和 T. Yan 等就分别将数据挖掘的方法应用到用户日志分析处理上。Mannila^[67]首先把用户访问页面视为事件,尝试从用户访问日志中寻找用户访问网站的周期;Chen^[68]提出了最大前向参考模型,并使用这种模型将用户访问会话分解为单个的事务(Transaction),再基于对事务的处理挖掘用户访问模式;而 T. Yan 通过用户当前访问习惯动态推荐相应页面。Perkowitz^[69]在人机界面的领域提出了自适应 WEB 站点(Adaptive WEB site)的概念。1998 年 Han^[70]把 WEB 服务器日志集成到资源数据一体结构中,以在线分析处理(OLAP)来分析日志数据。1999 年 J.Borges^[71]引入超链接概率原理,根据访问的条件概率判断用户频繁访问路径。

我国在此方面的研究起步较晚,从 1999 年才逐渐开始,包括:周斌采用 E-OEM 模型,又用 5 个用户访问模式做训练数据集,进行了关联规则挖掘;金阳^[72]用有序概念格增量的发现模式;董一鸿^[73]引入竞争神经网络的方法进行 WEB 日志挖掘。此外,在一些国外专业研究 WEB 数据挖掘的网站上面已经出现了用户访问日志分析工具,如 WEBtrender, Fastanalysis

等。这些工具主要是统计每一个页面用户访问的频次,以及用户访问页面的时间分布情况,其中比较成功的 WEBtrender 软件已经可以统计出所有类型的错误访问和有关关系页面。

WEB 日志内存储着海量的用户记录,这导致原有的数据挖掘算法难以直接适用,造成了现在 WEB 日志挖掘方法效果不佳的情况。其中难点主要集中在以下几个方面:

①由于互联网传输协议 HTTP 的无状态性,客户端、代理服务器端缓存的存在,使得用户访问日志分别存在于服务器、代理服务器和客户端。而对于静态 WEB 网站,服务器端的日志容易取得,客户端和代理服务器用户访问日志不容易取得,从而造成分析结果的偏差。

②可扩展性问题。WEB 上信息不断更新,用户的访问兴趣也同样不断发生着改变,而现有的大多数 WEB 使用挖掘方法均不具备动态调整的能力。如何能在保持原有提取用户爱好不受影响的情况下实现新兴趣的动态增加,则成为当前快速多变的 WEB 环境中进行日志挖掘所面临的一个严峻的挑战。

③现有 WEB 使用模式挖掘的结果通常是关联规则、序列模式。但随着人们使用 WEB 方式的增加,WEB 访问模式变得愈来愈复杂,新的信息表示方式和新的使用模式不断涌现。因此,需要引入新的学习机制以捕获用户使用模式中的新出现的规律。

(3) 电子商务推荐技术

电子商务推荐技术的研究主要包括两个方面:基于内容的过滤推荐技术和协同过滤推荐技术。基于内容的过滤推荐利用资源内容与用户间兴趣的相似性来过滤信息,达到信息推荐的目的。协同过滤推荐的方法指利用访问历史记录中已有用户的访问信息同新用户访问兴趣之间的相似性,将相似旧用户的访问习惯提供给新用户,从而达到电子商务推荐的目的。而协同过滤推荐是当前最成功的推荐技术,也是本论著信息推荐研究所要讨论的一个重点。

在基于内容的推荐技术研究方面,NewWeeder 等新闻过滤文本推荐系统采用文本中的单词作为文本的特征,通过采用的学习算法,判定树、神经网络的方法学习客户已评价过的商品的特征来获得对客户兴趣的描述,被称之为“商品与商品的相关性推荐”。该技术存在的最大的缺点是必须分析资源的内容信息,因此对音乐、图像、视频等信息无能为力,无法分析信息的质量,无法提供新颖的推荐。

由于基于内容的推荐技术的自身局限性,协同过滤推荐技术成为当前效果最好且广泛使用的技术。协同过滤的最大优点是不需要分析对象的特征属性,对推荐对象没有特殊要求,能处理非结构化的复杂对象。著名的推荐系统有 GroupLens/Net Percep。协同过滤推荐技术又可分为两类:①基于记忆的协同过滤,其先用统计的方法得到具有相似兴趣爱好的邻居客户,再基于邻居进行计算,所以该方法也称为基于客户的协同过滤或基于邻居的协同过滤;②基于模型的协同过滤,先用历史数据得到一个模型,再用此模型进行预测。目前电子商务推荐的研究多集中在对协同过滤推荐算法的改进上。如苑守慈^[74]提出了一种基于语义本体论的耦合聚类算法(OBPC),应用于异质移动产品的推荐上,不仅解决了现有协同过滤算法存在的局限性,即只能应用于具有历史项评分记录的产品推荐,而且通过设计出产品的本体,提高了推荐和搜索质量;余力^[75]等人在假设客户多兴趣具有相似性的前提下,提出了客户多兴趣下的个性化推荐算法。但是协同过滤的方法在实际运用中也逐渐暴露出一些需要解决的问题:

①数据规模大。电子商务中涉及的数据量非常庞大,而像最近邻居这样的算法效率会随着用户和产品的数目增多而下降,如何改进算法使之能更适应大规模计算是研究的重点之一。

②数据的稀疏性(Sparsity)。电子商务网站通常拥有大量商品,而每个用户购买或作评价的只是其中很小的一部分,通常不到 1%。这种情况带来的问题是用户间相似度计算的不准确,邻居用户不可靠。数据的稀疏性直接导致了算法准确率的降低^[76]。

③评分偏见问题(Bias)。每个用户均有着自己对事物的评价标准,这就造成用户的评分数据依各自的爱好而不同。例如,有的用户对所有的项目(产品)均标以高的分数,而另外的用户对事物评价严格,对每个项目(产品)标分均偏低。这样的现象会导致用户评分数据的不可比性。

④噪音问题(Noise)。有的用户进行评分时,并没有严格按照自己的喜好,随意为项目(商品)标记分值。另外,也有商家为了竞争,恶意破坏竞争对手的推荐系统,人为地制造假数据,形成对正常分析的干扰。

三、研究内容与研究方法

1. 论文研究的主要内容

网络消费作为互联网出现之后新型的消费方式,有其自身的一些特点。同时与传统消费相比,网络消费的规律与传统的经典微观经济学理论描述有所不同。本论著在总结国内外网络消费理论相关研究的基础上,探讨了网络消费的自身特点和与传统消费方式的区别。此外,中国的互联网发展进入了一个新的快速发展期,作为网民人数全球第二的国家,对目前国内网络消费现状的研究很有必要。本论著从分析《互联网调查报告》,得出了中国互联网经济条件下网络消费的一些现实情况,并总结了一些问题。

近年来,Internet 正以令人难以置信的速度在飞速发展,越来越多的机构、团体和个人在 Internet 上发布和查找信息。虽然 Internet 上蕴涵着丰富的数据,但由于 WEB 是无结构、动态的,并且 WEB 页面复杂程度远远超过了文本论著档,人们要想找到自己想要的信息犹如大海捞针。解决这些问题的一个途径,就是将传统的数据挖掘技术和 WEB 结合起来,进行 WEB 知识挖掘。本论著研究了基于 WEB 知识挖掘的内容和方法,并着重阐述了基于语义网的 WEB 挖掘。

本体是知识表示的一种形式,它能将领域知识表示成挖掘算法能够理解的形式。通过建立某个领域的本体,以这个本体为出发点去引导数据挖掘过程,可以提高获取知识的效率和质量^[77]。应用于信息检索上,可以使检索的信息更能满足用户的需求,克服目前检索方法造成的信息冗余或信息丢失^{[78][79]}。因此,我们引入了本体论,研究了网络消费本体的构建方法、过程和工具,并构建了基于 WEB 的网络消费行为知识本体。

电子商务推荐服务是提高电子商务网站竞争力、满足客户个性化需求的必然趋势,现有电子商务推荐技术和系统实质上只适用于单个电子商务网站,远远不能满足网络条件下大规模电子商务推荐的需求,并缺乏有效的推荐结果解释模型和评估体系。实时性差、推荐范围狭窄与网络信息资源丰富的矛盾、推荐质量远远落后于用户期望等都是亟待解决的问题。本论著从网络消费的基础理论入手,以网络消费知识为基础,以基于知识的网络消费个性化服务系统中要素智能和协同工作为目标,深入研究基于网络消费行为知识的挖掘、表示和推理的基本原理和方法,在此基础上探讨基于网络消费行为知识本体的构建,最后研究个性化服务推荐结果产生的方法以及基于知识的网络消费个性化服务系统推荐策略的自适应性理论与方法。

本文主要研究内容如下: