

B2C 电商企业下顾客购物信息定价模式研究

摘要

随着精准营销崭露头角，其实现基础——顾客购物信息的热度也水涨船高，且顾客购物信息交易平台在世界各地如雨后春笋般涌现，定价策略的制定对于其交易尤为重要。

在梳理数据与信息商品定价文献的基础上，通过结合顾客购物信息商品特点，提出了顾客购物信息价值模型：首先使用德尔菲法和访问调查法来获取专家意见，结合熵法和 AHP 共同确定顾客不同指标的价值权重，其次基于信息规则和高斯函数引入稀缺性指数、信息关联价值和信息即时性价值，继而得到单独顾客购物信息价值模型。在此基础上通过 logistic 增长函数来表示网络效应并引入数据饱和的概念，得到动态的、包含多个单独购物信息的数据包价值模型。

当 B2C 电商平台将其数据包作为商品出售时，依据 Philip Kotler 的定价理论：首先确定定价方法为成本加成定价法，其次通过需求弹性函数和市场生命周期来估计市场需求，接下来考虑影响定价的其他因素（市场管控政策、隐私政策和外部效应），最后说明如何执行与控制价格。

在关于顾客购物信息定价文献较少的情况下，通过此定价模式的研究，在一定程度上弥补了学术空白，并可以为交易平台定价提供借鉴，具有一定的理论和现实意义。

关键词：顾客购物信息；数据商品；定价模式

Pricing personal information of E-consumers

Abstract

With the emergence of precision marketing on B2C e-commerce platform, its foundation of realization, personal information of E-consumers has become a hot data commodity. At the same time, personal information trading platform is springing up all over the world, and the pricing strategy is particularly important for the trading. However, in China, a mature pricing mechanism has not yet been formed.

In this paper, on the basis of sorting out relevant research literature on data and information product pricing at home and abroad, combining with the status quo of customer shopping information trading and existing pricing strategies. The customer shopping information value model under the B2C e-commerce platform based on the cost-plus pricing method was proposed, and the Delphi method and the interview survey method were innovatively used to obtain expert opinions. Based on this, entropy method and AHP are combined to determine the value weight of different customer indicators, and then the scarcity index, information correlation value and information immediacy value are introduced based on information rules and Gaussian function. On this basis, the concept of data saturation was introduced by combining the logistic growth function, so as to calculate the total number of samples needed when the K value was reached in different fields, and then the value model of packets containing multiple personal shopping information was obtained dynamically.

When B2C e-commerce platform sells customers' shopping information packet as data commodity, its price should not be determined only according to its value. Combined with Philip Kotler's pricing process, this paper divides the enterprise pricing process from a strategic perspective into five steps: First, the pricing method is cost plus pricing. Then the market demand is estimated by the demand elasticity function and the market life cycle. Next, combined with the customer shopping information value model, other factors affecting pricing (market control policies, privacy policies and externalities) are briefly described. Finally, how to implement and control the price is explained.

In this paper, the customer shopping information pricing model under the B2C e-commerce platform is proposed, which makes up the research gap to a certain extent and has certain academic and theoretical significance. At the same time, the pricing model proposed in this paper can provide reference for setting the pricing mechanism of shopping information trading platform for domestic and foreign customers, which has certain practical and practical significance.

Key words: Customer shopping information, Data commodity, Pricing model.

目录

摘要	I
Abstract	II
第 1 章 绪论	1
1.1 选题背景	1
1.2 选题意义	2
1.3 国内外研究现状综述	2
1.3.1 国内研究现状综述	2
1.3.2 国外研究现状综述	2
1.3.3 研究评述	3
1.4 研究内容和方法	6
1.4.1 相关概念界定	6
1.4.2 研究内容	6
1.4.3 研究方法	7
1.5 研究技术路线	8
1.6 本章小结	8
第 2 章 顾客购物信息市场分析	9
2.1 顾客购物信息市场参与主体	9
2.1.1 顾客购物信息交易平台	9
2.1.2 顾客购物信息提供方	9
2.1.3 顾客购物信息买方	9
2.2 顾客购物信息交易流程	10
2.3 顾客购物信息商品特性	11
2.4 顾客购物信息交易现状	12
2.5 本章小结	14
第 3 章 B2C 电商企业下顾客购物信息价值模型	15
3.1 单个顾客购物信息价值模型	15
3.1.1 基于熵法与 AHP 的权重确定	15
3.1.3 顾客购物信息的总价值	19
3.2 顾客购物信息数据包价值	20
3.2.1 网络效应	20
3.2.2 数据饱和	20
3.2.3 数据包总价值	21
3.3 本章小结	21
第 4 章 数据包交易定价策略	22
4.1 成本加成定价法	22
4.2 数据包价值的需求弹性	22
4.2.1 市场生命周期	22
4.2.2 需求弹性函数	22
4.3 影响定价的其他因素	23

4.3.1 宏观环境分析	23
4.3.2 外部效应	23
4.4 价格的确定与控制	24
4.4.1 价格的确定	24
4.4.2 价格的执行与控制	24
4.5 本章小结	25
参考文献	26
致谢	错误!未定义书签。

第1章 绪论

1.1 选题背景

信息技术的不断发展为 B2C 电子商务领域带来了新的发展机遇。客户生成的数据已成为 B2C 电子商务领域的重要工具之一。近年来，中国的电子商务进入了快速发展时期，电子商务的交易量和规模迅速扩大，如图 1-1 所示。另一方面，B2C 电子商务的庞大用户和商品基础要求电子商务公司根据用户需求和潜在需求建立精确的营销体系，以提高其核心竞争力。

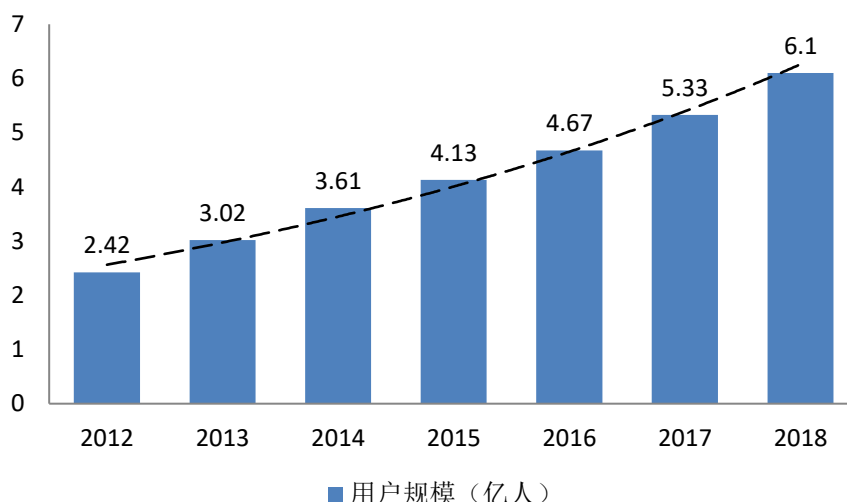


图 1-1 中国网络购物市场用户规模图

数据时代的精准营销模式，可以就消费者展开初步的市场细分，细分的过程主要是为用户建立消费者画像，依据其消费痛点对其进行精准营销或提供个性化服务，接下来依据此顾客的反应对其画像进行持续不断的更正，并根据对顾客购物信息的深切分析来找出并满足消费者的痛点。而精准营销则少不了其数据基础——反映消费者消费能力、消费偏好的顾客购物信息作为支撑。因此，顾客购物信息交易成为了创新的商业模式，推动了精准营销时代的来临。反之，在精准营销时代，顾客购物信息成为企业的重要资产，顾客购物信息从企业内部的独享走向企业间的共享。由于缺乏规范的顾客购物信息共享渠道和统一的交易规范，顾客购物信息交易平台的出现是时代的趋势。

由于我国顾客购物信息交易还处于起步阶段，各类顾客购物信息交易平台刚刚建立，顾客购物信息交易的流程规范尚不成熟，顾客购物信息定价也没有一个清晰的规则。目前顾客购物信息交易市场定价混乱，大多是卖方单方面决定信息商品的定价，这就导致有些信息商品定价过高，买方处于被动地位，参与正规顾客购物信息交易的企业较少，顾客购物信息市场的活跃度不高。反而是利用黑客技术和 cookies 等违法手段获取数据并通过不正当渠道交易的现象时有发生。另外，由于缺乏统一的定价机制，各家顾客购物信息交易平台定价规则不一，导致顾客购物信息市场定价混乱，严重影响了交易秩序。基于上述背

景，制定统一科学的顾客购物信息商品交易定价策略势在必行，既可以规范交易双方的行为，活跃顾客购物信息交易市场，还有助于顾客购物信息交易行业持续健康发展。

1.2 选题意义

数据交易过程就是买卖双方博弈的过程，本文所提出的定价模型，探讨不同交易环境下的最优定价策略，具有理论和现实意义如下：

（1）对现有理论的进一步探索和补充。现有理论研究关于数据市场研究较多，但是关于具体顾客购物信息和定价策略的探讨涉及较少，本文从相关研究出发，结合顾客购物信息的特点和真实交易过程中的供需波动得到一个较为切实的顾客购物信息定价模式。

（2）关注商品实际价值，活跃顾客购物信息交易市场。现有的顾客购物信息交易平台基本都是卖方市场，定价大多由卖方单方面确定，买方话语权较小。本文提出的定价模型考虑到商品的实际价值以及买卖双方需求，使交易双方参与到交易过程中来，公平公正，扩大交易量，进而活跃顾客购物信息交易市场。

（3）规范顾客购物信息交易。在互联网交易日益繁荣的时代，一个稳定的价格机制不仅可以对交易平台的卖方起到一定的约束作用，同时可以提高买方在交易中的主动权，获得更好的预期收益，形成良性循环。

1.3 国内外研究现状综述

因为国内外对于研究购物信息定价的文献微乎其微，而购物信息本就是信息的一种，也属于消费者产生的数据。而信息本就是已被处理成某种形式的数据，即从数据演化而来。因此笔者收集整理和分析了关于数据定价和信息定价的文献。

1.3.1 国内研究现状综述

关于数据商业模式方面，李骥宇梳理了国内外的数据交易模式，主要分为：原始数据交易模式、经过分析处理后的数据交易模式、数据决策方案交易模式、数据中间商交易模式^[1]。关于数据交易虽然是近一两年才有所发展，但我国学者也已有初步探索，高丹和向阳提出阻碍国内数据交易的关键在于数据价值链缺失和企业对于数据流通的恐惧和回避心理，因此他提出从数据商品化、政府推进数据开放、消除安全顾虑三个方面入手来解决^[2]。

刘朝阳对数据交易问题进行了全面、深入的研究，发现数据定价机制是一个关键性的、共性的基础问题：由于缺乏统一且符合标准的交易渠道和交易规范，加之当前的数据交易平台组成来源十分复杂，导致数据交易总体看来集中度还不高，交易规模也并不大^[3]。张晓玉基于讨价还价理论对数据商品交易价格进行了讨论，提出了“一对一”讨价还价博弈模型^[4]。黎春兰提出按使用付费定价模型(Pay-per-use Pricing Model)和预订定价模型(Subscription Pricing Model)^[5]

1.3.2 国外研究现状综述

A.Mowshowitz 提出了“千层盆”模型，他认为原始的信息产品在生产出来的最初有一

个基价，也叫价格的核心层，此后对该原始信息的每一道加工工序都将增加它的一层交易价格，而由于每一道加工过程对信息产生的增值效果是不同的，需要根据效果加权计算价格，因此信息产品的价格等于它的基价加上每一次加工的加权价格^[6]。Jean Tirole 认为传统的定价方法和供求分析不适用于信息产品，要根据消费者的支付意愿来对产品进行区别定价，或者在不同的市场、不同的时间段、不同的购买数量定为不同价格^[7]。Carl Shapiro 和 Hal Varian(2000)提出，商家可通过不同版本的差别化定价来区分不同的买家类型，不同类型买家对信息产品的支付意愿和效用追求不一样，因此设计出吸引不同类别买家的差别版本，来引导买家的自我选择^[8]。

Muschalle A 提出免费定价策略、按使用付费定价策略、数据包定价策略、固定费用关税策略、二部收费制策略和免费增值策略^[9]；M.Balazinsk 和 N.G.Mankiw 认为当前定价模式存在 4 个缺点，即当前定价模型允许套利、模型假设所有数据集是等价的、数据由买家进行缓存以及数据提供者没有如何设定价格的指导性建议，并提出一种细粒度的数据定价策略^[10-11]。M.Balazinska 提出了数据市场中的视图概念，一个数据实例的视图相当于实例的一个版本，视图也许只包含数据的一个子集或仅仅某些列或粗粒度级别的信息，并根据不同的价格进行销售^[10]。Li.C 提出了一种根据噪音查询应答来进行定价的理论框架，进而划分数据拥有者之间的价格，根据他们隐私丢失情况获得应有补偿，并指出隐私估值也许与数据本身是强烈相关的^[13]。

黎春兰指出：数据交易仍然采用协议定价、拍卖定价、集合定价等定价方式，数据定价方法无法实现人工智能时代自动化定价的目标；再者，由于信息不对称，存在第三方非法套利的情况，严重影响数据交易市场的秩序^[5]；此外，Balazinskam 认为由于用户隐私保护的限制，再加上数据本身极易复制的特性，使得数据在价值上具有极高且不可恢复的可变成本和极低的固定成本^[10]，导致了传统定价方法在数据领域不再生效。因此，数据定价方法成为了国内外科研工作者的一个研究重点。

1.3.3 研究评述

综上所述，现在主流的数据（或信息）商品定价模式^[15]如下

（1）基于博弈论的协议定价模式

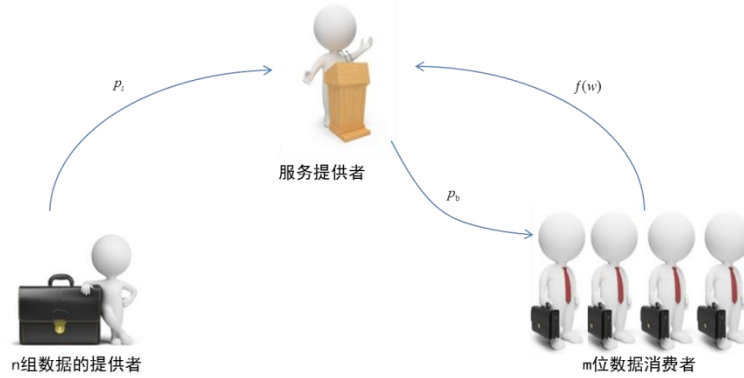
协议定价就是数据拥有者和数据购买者通过协商，对价格达成统一，这也是目前应用最为广泛的数据定价方法^[8]。首先，数据拥有者根据自身对数据的认识，率先为打算出售的数据定价；其次，数据购买者如果认可数据拥有者提出的价格，则二者交易成功，否则，可通过反复磋商的方式进行议价；最后，二者如能达成一致则交易成功，若不能则交易失败^[16]。若存在多名数据购买者，并且数据购买者有独占数据需要的时候，可采用拍卖的方式对数据进行定价，出价最高者获得数据的购买权。根据此种情况，张晓玉^[17]根据博弈论方法运用“一对一”和“一对多”的讨价还价模式，对数据这种特殊商品的价格确定过程进行了详细分析，对数据定价协商过程进行了建模。协议定价是当前实践中应用最为广泛的数据定价方法，其能够较为便捷地对数据进行定价。但是，由于数据买卖双方信息的不对称，

对数据价格认识的不一致^[18]，往往不能准确地评估数据价值，致使数据价格出现偏差，可能会出现非法套利的情况。

（2）基于数据特征的第三方定价模式

当前，国内外数据交易平台普遍采取的一种方法是可信第三方定价。在数据拥有者无法准确针对数据进行定价的情况下，可委托可信第三方进行交易。例如：Azure、Datamarket、上海数据交易中心、贵阳数据交易所等数据交易平台均可根据平台自有的包括数据量、数据完整性、数据时间跨度、数据稀缺性等在内的数据质量评价指标，对数据进行定价。通过第三方定价方法，每个数据集的价格都将根据数据属性因素和数据集的数据量进行计算^[9]。

无独有偶，Niyato 等^[20]结合了经济学中 Stackelberg 模型和机器学习中的分类算法，将数据交易分为数据提供者、服务提供者和数据消费者三个部分，如图 1-2 所示。



当数据量为 n 的时候，数据提供者向数据平台提供的使用数据的价格为 p_s ； M 位数据消费者愿意为数据付款的概率密度为 $f(w)$ ，其中 $t \in [0, T]$ ；服务提供者设定数据提供者提供的“生数据”的单元价格为 p_b ，通过最大化 p_b ，使 p_b 尽可能逼近 p_s 来实现效益最大化。

$$p_b = \arg \max_{p_b} p_b n(p_b),$$

$$s.t. \quad n(p_b) = \arg \max_{n, p_s} (p_s M \int_{p_s}^T f(t) dt - p_b n)$$

然而，采用第三方辅助定价的方法，首先必须保证第三方数据交易平台是完全可靠的。但是，当前国内外数据交易平台背景复杂，政府、企业、个人都参与其中，没有形成一个统一、规范的数据交易平台。对于用户而言，平台缺乏透明度，会导致信息误传和非对称信息的产生^[21]。其次，采用第三方定价模式，交易的数据往往是整个数据集，并没有针对每个数据元组进行定价，如果用户需要的是部分数据，则必然要购买整个数据集，造成了一定程度上的浪费。此外，采用人工标注的方式进行第三方定价，无法针对海量数据集自动生成数据的价格，缺乏时效性。

（3）基于元组的定价模式

所谓元组，指的是待交易数据集中的每条交易数据。而人工定价领域中，被定价的对象往往是一个数据集合，并不能针对数据集合中的每个元组进行定价。目前的定价模式往往假设集合中所有的元组在价格上都是完全一致的。当然在实践过程中由于不同元组中包含隐私信息量的不同，这一假设是完全不成立的。为了解决这个问题，Balazinska 等^[10]提出在元组这一结构粒度上设定数据的价格，并基于公共数据库中数据的顺序建立定价函数。定义价格函数 f 为

$$f : D \rightarrow R^+$$

其中， D 为数据集合， R^+ 为数据价格。包含多个元组的数据集价格，就是各个元组单一价格之和。

在以上研究的基础上，Shen 等^[24]提出了积极分级和反转定价机制。数据属性根据其影响数据质量的程度被分为不同等级，每个数据元组的准确价格都是根据数据属性因素和数据集的整体价格进行计算。将数据元组作为最基本的数据度量组合，结合信息熵、权重、数据引用指数、花费等影响数据价值的因素对每个元组进行定价。但是，没有对选取信息熵、权重、数据引用指数、花费作为影响数据定价的特征的原因进行说明。Shen 等^[24]的具体计算过程如下：

(1) 假定 P_D 为数据集的需求价格， C 为收集、分析数据以及运行数据交易平台的花费，供给价格为 P_S ，则

$$P_S = P_D - C$$

(2) 假定信息熵(q)、权重(w)、数据引用指数(r)的系数分别为 α 、 β 、 γ ，其满足约束：

$$\alpha + \beta + \gamma = 1$$

(3) P_i 为数据集中第 i 个数据元组的价格，其计算公式为

$$P_i = P_S * (\frac{w_i}{w} * \alpha + \frac{q_i}{q} * \beta * \frac{r_j}{r} * \gamma)$$

其中， q_i 、 w_i 、 r_i 分别为第 i 个数据元组的信息熵、权重和数据引用指数，其计算公式为

$$\sum_{i=1}^n \frac{q_i}{q} = 1, \quad \sum_{i=1}^n \frac{w_i}{w} = 1, \quad \sum_{i=1}^n \frac{r_i}{r} = 1$$

采用元组的方式对数据进行定价，虽然能够公式化地计算出数据价格，从技术上表示了元组在一定程度上具有的价值。但是，在现实交易的过程中，用户需求偏好也从一定程度上决定数据的价格，而通过现有元组定价的方式无法表征数据获取难度、数据稀缺性、用户通过数据获得的盈利等数据价值。此外，在数据环境下，数据往往包含用户的隐私，

因此影响数据价格的因素十分复杂，如果仅仅通过几个简单的公式进行计算，不能完全表示数据的价格。

(4) 基于查询的定价模式

待交易的数据往往是存储在结构化或非结构化的数据库中，用户需要购买的数据往往需要从数据中查询获得，因此就诞生了基于查询的定价模式。该定价模式的数学模式允许卖方指定一些视图的价格，允许买方根据自身需要进行任意查询来购买需要的数据，同时模型能够通过指定视图的价格生成其他任意视图的价格。这样查询到的数据价格就是一系列能够组合出该查询的视图中最优的情况。以下为此定价模式的两个先决条件：

(1) 抗套利 (arbitrage-free)：以购买全美国的商业数据为例，美国全国的数据价格应该比分别购买 50 个州的价格之和便宜。

(2) 免贴现 (discount-free)：当确定每一个视图价格时，对于整体数据库而言，应在各视图之和的基础上有一个折扣。这一模型存在一个问题，原始的基于查询的定价模型限制了用户只能以固定的数量或通过预定义的视图购买数据。Tang 等^[25]对数据库中的每个元组分配价格，然后通过生成满足查询结果的最小视图，来定义任意查询的价格。为了保证高效生成查询结果，Tang 等^[25]还提出使用了 MiniCon 算法对用户提出的查询进行修正，在查询结果一致的情况下，对查询过程进行优化，保证生成数据价格过程的时效性。

1.4 研究内容和方法

1.4.1 相关概念界定

(1) B2C 电商企业

目前网络购物的形式主要有 B2B、B2C、C2C 和 G2B 四种。B2B(Business-to-Business 的缩写)即企业和企业之间的互相交易关系；C2C (Customer to Customer 的缩写)即个人与个人之间的互相交易关系；G2B(Government to Business 的缩写)即政府和企业之间的互相交易关系；B2C (Business to Customer 的缩写)即商家和个人之间的互相交易关系，这种购物形式一般是电子商务企业利用互联网平台进行在线销售，消费者利用互联网购物平台运用在线支付的方式进行购物。由于本文研究的是我国网络购物消费者的消费信息定价的问题，所以我们在本文中不探讨不涉及电子商务平台和个人的交易，因此本文把研究对象界定为 B2C 电商企业。

(2) 顾客购物信息

顾客购物信息指商家通过消费者用户注册、填写订单或浏览网页等过程收集到的消费者在购物过程中产生的数据。商家可以通过对顾客购物信息加以分析，了解消费者的背景资料、上网习惯和关注重点等进而针对性的进行营销方式的改变来迎合消费者的喜好。

1.4.2 研究内容

本文研究的主要内容是：基于顾客购物信息交易的特性和现状，探讨顾客购物信息交易

定价策略，从市场营销学中成本加成定价法的角度为顾客购物信息交易市场的定价实践提供理论指导。

（1）顾客购物信息交易与常见定价策略研究

建立定价模式的先决条件是对定价对象有深入的了解，因此本文对顾客购物信息交易进行了基本研究，明确数据交易涉及到的数据需求方、数据提供方、顾客购物信息交易平台三方的职责，探讨常见的数据交易模式，分析顾客购物信息商品的特性，对数据交易现状进行了调查研究。将常用的数据定价模式整理归纳后，建立本文研究的 B2C 电商企业下顾客购物信息定价模型。

（2）顾客购物信息数据包价值研究

首先，基于德尔菲法和访问调查法结合的方法使用熵法和 AHP 法确定顾客不同指标的价值权重，其次，通过成本加成定价法建立个人消费者的购物信息价值模型。然后，基于信息规则和高斯函数引入稀缺性指数、信息关联价值和信息即时性价值。在此基础上通过结合 logistic 增长函数并引入数据饱和的概念，可以计算不同领域达到 K 值时所需样本总量，进而得到动态的包含多个个人购物信息的数据包价值模型。

（3）战略视角下的数据包定价模型

结合 Philip Kotler 的定价流程将基于战略视角下的企业定价流程分为五个步骤：首先确定定价方法为成本加成定价法，然后通过需求弹性函数和市场生命周期估计市场需求，接下来结合上一章的估算商品价值简述影响定价的其他因素（市场管控政策、隐私政策和外部效应），最后对如何执行与控制价格进行了说明。

1.4.3 研究方法

本文以 B2C 电商企业下顾客购物信息定价模式问题为研究对象，结合市场营销学中的定价策略，采用文献研究法、混合研究方法、问卷调查的方法。

（1）文献分析法：

通过搜索并整理大批相关中英文文献资料和相关著作，对研究对象的发展过程和未来趋势进行归纳，做好完整深切的文献综述，同时准确把握研究对象的最新动态，并对相关理论进行了全面学习，使自己具有较高的理论基础，并分析这些资料形成自己的观点，为后面的文章撰写做好比较充足的准备。

（2）混合研究方法：

立足理论研究，将跨学科的经典理论、模型与 B2C 电商企业下顾客购物信息定价模式的现实问题相结合，进行理论分析，并予以构造数学模型。之后结合顾客购物信息数据包价值模型，构建其在 B2C 电商企业下的定价模型，通过对预期收益的确定，得出结论。

（3）访问调查：

基于本文的研究内容以及 B2C 电商企业特性，按照现有研究及科学程序设计一套科学合理的调查方式，先进行小规模의预调查分析，根据数据分析结果选取对象包括

省级“双一流”高校电子商务领域教授、500强电商企业中层和数据公司高层合计9人，采用德尔菲法与访问调查法相结合的方法开展调查。

1.5 研究技术路线

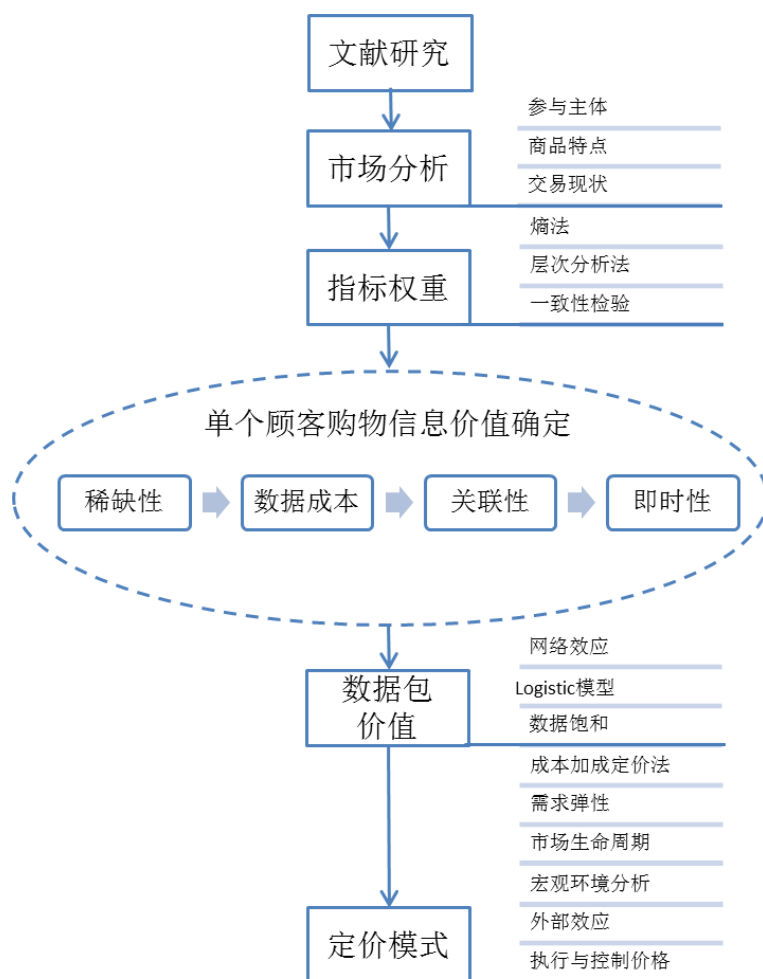


图 1-3 研究技术路线图

1.6 本章小结

本章是本文的绪论部分，主要论述了以下内容：首先，笔者在阅读大量中英文文献后叙述了研究的现实背景和理论背景，并在此基础上总结了国内外研究现状。其次，对本文的研究对象进行了界定。最后，对研究所采用的研究方法与技术路线进行了说明。

第 2 章 顾客购物信息市场分析

2.1 顾客购物信息市场参与主体

2.1.1 顾客购物信息交易平台

顾客购物信息交易平台是指向顾客购物信息需求方、顾客购物信息提供方提供交易服务的场所。顾客购物信息交易平台通过提供顾客购物信息 API 的录入、检索、调用来提供顾客购物信息交易的场所和平台。

顾客购物信息交易平台在顾客购物信息交易中起着非常重要的作用，首先，在交易前，需要对顾客购物信息双方首先进行市场准入审核，进行顾客购物信息确权，买卖双方保证顾客购物信息的所有权合法、可信、不被滥用，对顾客购物信息的使用权、运营权进行交易。然后，要对进行交易的顾客购物信息商品进行审核，防止造假欺诈的、来源不合法的、不安全的顾客购物信息进入市场。最后，交易平台要根据需求设置相应合理的定价机制来撮合交易双方，达成顾客购物信息交易。

2.1.2 顾客购物信息提供方

本文的顾客购物信息提供方指的是 B2C 电商企业。它们是能够自身生产顾客购物信息的公司，比如综合商城、百货商店、垂直商店、复合品牌店等，这类公司运营时产生顾客购物信息，经过脱敏化处理后可以帮助业务相关的公司做出更好的决策，或者整合后对特定人群的消费趋势、市场走向进行预测。

B2C 电商企业对消费者数据的搜集包括两个方面：一个方面是对消费者显性信息的搜集。比如用户在登录网站、购买商品、获取服务时搜集消费者的个人身份信息，如姓名、年龄、性别、地址、电话号码、电子邮箱地址、信用卡号码、职业、个人爱好。另一个方面是对隐性信息的搜集，包括用户的网络行为轨迹、消费习惯、社会关系、兴趣爱好等。比如利用 cookie 来跟踪消费者行为、用日志程序记录消费者的交易情况、利用嵌入式软件标记用户特征和环境、利用开放式或半开放式的社交应用抓取用户关系信息。通过对消费者显性信息和隐性信息的跟踪，生成消费者的个人档案和标签，从而获取消费者购物信息。此外，随着移动互联技术的发展，不仅更容易获取消费者信息，而且对消费者信息的获取不再拘泥于上网数据，还包括了用户的日常生活轨迹信息。比如，很多 B2C 电商平台只有真实的填写用户信息才能够正常使用，虚假的用户信息使得消费者无法使用或影响使用效果，更无法获得个性化的服务；基于实时地理位置的 APP 能够绘制出更立体、更实时的用户行为及消费场景等。利用这些数据能够还原用户的真实生活场景，如日常安排、身体健康状况、个人兴趣爱好等。

2.1.3 顾客购物信息买方

顾客购物信息买方，也即顾客购物信息需求方。目前，需要应用顾客购物信息进行分析预测和辅助决策的领域包括政府公共服务、商业分析、企业管理等方面。可具体归纳为

以下三类：

第一类是企业，各类企业是顾客购物信息买方的主要构成部分。企业进行决策分析时，往往需要依靠大量的顾客购物信息支持，当企业内部顾客购物信息资源不足以支撑时，就需要外部顾客购物信息辅助决策，可以是本行业的顾客购物信息，还可以是跨行业的联合分析，进行用户行为偏好顾客购物信息整合。利用顾客购物信息分析技术提供精准营销与个性化服务，带来的销售收入是相比以往传统营销高得多的。

第二类是政府部门，政府部门可结合社会及网络上的顾客购物信息，了解社会状况和民意，做出更好的公共服务。还有一类是科研机构，科研机构进行研究调查时，往往需要大量的实验和数据做支持，而科研机构本身能够产生的数据并不多，此时在市场上购买成为一个较好的选择，比如预测特定细分消费者集合的购物趋势研究可能需要大量的购买商品记录、消费者信誉记录、消费能力记录等消费者购物信息。

2.2 顾客购物信息交易流程

在顾客购物信息交易平台上发生的交易流程一般如图 2-1 所示。顾客购物信息卖方先将自己的顾客购物信息产品按照规定进行预处理，交由顾客购物信息交易平台审核，审核通过后在顾客购物信息市场上挂牌待售。顾客购物信息买方可以直接进入顾客购物信息交易平台，浏览上面的顾客购物信息产品选择自己需要的进行交易，也可以在顾客购物信息交易平台上发布需求，经由顾客购物信息交易平台对需求进行审核后，等待符合条件的顾客购物信息卖方来应标，完成交易。

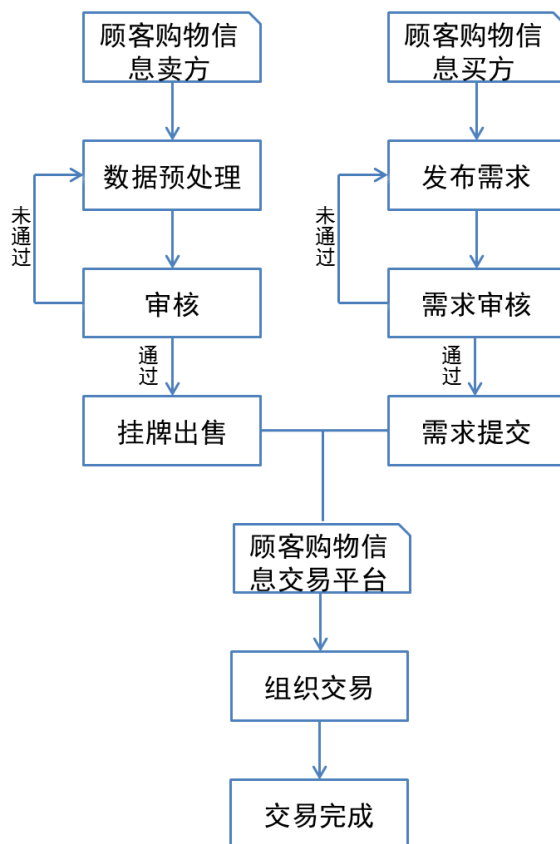


图 2-1 顾客购物信息交易流程图

2.3 顾客购物信息商品特性

不同于可以直接进行交易的传统商品市场，顾客购物信息在进行交易前必须进行清洗、加工。一方面是要进行脱敏处理，去掉顾客购物信息中的商业和个人隐私信息，另一方面是要对顾客购物信息进行统一要求的标准化处理，比如文件格式、传输协议、API 接口标准等，便于顾客购物信息的需求方对顾客购物信息的再加工和利用。经过初步处理后的顾客购物信息成为商品，具有如下两个特性。

(1) 顾客购物信息商品价值的相对性

传统商品交易市场对商品价值计算已经有较为完善的数学模型。但对于顾客购物信息商品来说，因为顾客购物信息商品的使用价值在交易时不能直接显现，其价值表现在对顾客购物信息的分析和挖掘之后，而这种价值取决于不同企业的需求，不同的企业对顾客购物信息的需求和利用方式是不同的，其价值也可能具有天壤之别^[27]。顾客购物信息商品的这种特性使得统一贴上一个价格标签在平台进行出售这种传统商品交易的方式不再合适，而更合理的是通过计算其价值，再加入市场影响情况进行定价。

(2) 顾客购物信息商品的时效性

顾客购物信息商品对时间的敏感系数不低，大多数顾客购物信息是实时变化的，要求顾客购物信息供应商要在一段时间内保持更新，所以需要一种随之变化的定价策略；有的实时性强的顾客购物信息在一段时间后就会贬值，很可能会迅速被更有价值、质量更高的新顾客购物信息替代，这样的顾客购物信息在定价时也要考虑其特性制定适合的策略。顾客购物信息商品的时效性强决定了其不能用传统商品的定价策略，也不能一成不变。

2.4 顾客购物信息交易现状

全国数据交易市场于 2008 年伊始出现，并且快速发展。2015 年是我国数据交易产业发展的黄金期，北京、上海、贵阳、武汉等地纷纷成立数据交易平台，开始建设性运营。与此同时，顾客购物信息成为绝大多数非政府数据交易平台的主打产品。但是目前还没有成熟的顾客购物信息交易机制，顾客购物信息市场也是刚刚起步，顾客购物信息孤岛现象仍然存在，企业所拥有的顾客购物信息资源还是多为企业内部所用，把顾客购物信息作为商品在市场上进行交易的企业较少，因此市场上的顾客购物信息交易还未形成规模。目前顾客购物信息交易面临的主要问题和挑战有：

（1）顾客购物信息商品标准化

市场上顾客购物信息商品由于来源渠道不同，种类繁多，导致顾客购物信息的格式不一，标准化程度很低。要把顾客购物信息变成商品进行交易，首先要解决标准化问题，标准化的商品才能实行标准的定价策略，按照标准的销售模式进行销售，顾客购物信息产业才能持续健康发展。但顾客购物信息商品又和工业化的商品不一样，工业化的商品主要是实物商品，是通过标准化的原料和制作工艺，统一流程生产出来的。顾客购物信息商品的标准化问题主要体现在以下两个方面：

首先是关于文件格式的标准，顾客购物信息交易市场上文件格式多种多样，顾客购物信息进行交易后的二次加工利用存在障碍，不利于顾客购物信息的流通。顾客购物信息交易市场应该统一顾客购物信息商品的文件格式。一般来说，每种顾客购物信息商品应该提供机读格式 XML/JSON/CSV 和人类可读格式 X TXT/PDF/CSV/XLS/XLSX 等多种选择，以便让顾客购物信息需求方依据其使用目的，自行选择合适的顾客购物信息集格式进行下载。

其次是关于传输协议的标准，顾客购物信息提供方的顾客购物信息要经过顾客购物信息封装，转化成可以在网络上传输的信号，发送到网络上。顾客购物信息传输双方在传输过程中必须遵守统一传输规则，使顾客购物信息交易顺利完成。一般常见的传输协议有 TCP/IP、NetBEUI、DHCP、FTP、TFTP 等。顾客购物信息买方可通过卖方提供的 API 接口实现顾客购物信息查询或传输。顾客购物信息交易市场规定所有的顾客购物信息都按统一的格式和传输协议输出后，才能保证顾客购物信息的流通和使用畅通。

（2） 顾客购物信息商品定价及资产估值

当下的顾客购物信息市场处于卖方市场，卖方掌握定价的主动权，没有考虑买方的购买能力和发言权，缺乏第三方对顾客购物信息商品的科学有效的评估，不利于顾客购物信息市场规模的扩大。由于顾客购物信息商品的价值因人而异，难以评估，市场上还未出现行之有效的价值评估量表。顾客购物信息定价和资产估值是很具挑战性的前沿的研究方向，目前还尚未出现成熟的定价模型。市场上给顾客购物信息定价一般是依据两个方面，一是顾客购物信息的效用，二是顾客购物信息的稀缺性。顾客购物信息的效用是指量化顾客购物信息商品的贡献，预测其将会给用户带来的效果，是从分析结果来逆推顾客购物信息的渊源^[28]。

（3） 顾客购物信息交易机制缺乏

顾客购物信息交易市场的活跃依赖于其成熟的交易机制，目前各顾客购物信息交易平台处于起步探索阶段，各种交易机制尚不完善。对于顾客购物信息商品来说，因其易复制、易传播、估值困难等区别于普通商品的特性，不能完全参照金融交易所和商品交易所的定价模式。传统交易所的竞价模式一般是连续竞价和集合竞价，涉及的是多对多的关系，而顾客购物信息交易一般是一对一或者一对多的关系^[29]。不同类型的顾客购物信息需要设计不同的交易机制，比如有的顾客购物信息是一次性交易完成，有的顾客购物信息需要卖方一段时期的实时供应，有的顾客购物信息容易采集同类型顾客购物信息多，有的顾客购物信息不易采集并且同类型顾客购物信息很少。顾客购物信息交易需要平台的集中撮合，既要考虑到买方的利益，又要考虑到卖方的利用，同时还需要能够活跃市场，维护市场的正常秩序。

（4） 顾客购物信息隐私保护及安全问题

在顾客购物信息交易的整个生命周期都需要注意顾客购物信息隐私和安全问题，顾客购物信息采集、交易平台、顾客购物信息使用都要有相应的措施来监管^[28]。首先是顾客购物信息采集阶段，不能侵犯个人和商业隐私，首先要明确可以加工的顾客购物信息范围，何种顾客购物信息不可以被采集，然后要对采集到的顾客购物信息进行脱敏化处理后可以进入市场进行交易。

顾客购物信息通过交易平台进行交易时，交易平台要注意对顾客购物信息的安全保护，要对顾客购物信息进行加密后再进行买卖，防止顾客购物信息随意被人窃取，还可对顾客购物信息进行安全等级划分，顾客购物信息的存放也要注意顾客购物信息库的加密。交易完成后，顾客购物信息买方或顾客购物信息的使用者，要严格遵守顾客购物信息使用规范，不可将顾客购物信息用于危害公民和国家安全的方面，更不可随意将顾客购物信息泄露给第三方。比如电商网站将用户的购物兴趣标签顾客购物信息出售后，用户接到大量的营销电话的骚扰，这就属于侵犯用户个人信息隐私^[27]。

2.5 本章小结

本文所讨论的顾客购物信息交易是指发生在交易平台上的，顾客购物信息提供方和顾客购物信息需求方之间对顾客购物信息商品进行买卖的交易。顾客购物信息提供方需要在不侵犯个体隐私和安全的前提下，采集顾客购物信息后对顾客购物信息进行脱敏处理，去除个人身份属性^[26]，接下来才可以进行交易。顾客购物信息需求方根据企业对于顾客购物信息的需求，可以在顾客购物信息交易平台上发布需求，也可以购买交易平台上顾客购物信息提供方已有的顾客购物信息商品。在和提供方达成价格一致要求后，可在交易平台的撮合下完成交易。

本章首先对顾客购物信息交易的各方面进行了概述：对顾客购物信息交易涉及到的三个参与主体：顾客购物信息交易平台、顾客购物信息卖方（提供方）、顾客购物信息买方（需求方）进行了详细介绍，然后探讨了 B2C 电商企业下顾客购物信息交易流程。其次，对顾客购物信息商品与传统商品相比特性进行了分析。最后，基于研究的现实背景，通过文献综述对顾客购物信息交易现状进行了梳理，并根据现有研究提出了交易的四点主要问题和挑战，选取第三个问题“顾客购物信息商品定价及资产估值”作为本文的主要研究问题，将放在三四章进行建模分析。

第 3 章 B2C 电商企业下顾客购物信息价值模型

3.1 单个顾客购物信息价值模型

对于的单个顾客购物信息的综合价值确定，考虑三个主要组成部分。

第一个部分是预期价值，对于同一个消费者，其顾客购物信息的各个指标的价值对于 B2C 电商平台来说是不同的（例如顾客的年龄和收入的价值是不同的）。笔者通过熵法与 AHP 确定指标权重后，依据指标稀缺性，结合成本导向定价法进行计算。

第二个部分是信息的关联价值。顾客购物信息价值不仅取决于信息的完整性，还取决于关联性。（例如单独顾客的年龄和收入产生的价值少于这两个指标的组合价值）。笔者建立了高斯混合模型来研究相关值与信息相关值之间的关系^[32]。

第三个部分是信息的即时性价值。顾客数据的价值随着时间衰减。另一方面，一段序列长度很长的数据应该具有附加价值。

3.1.1 基于熵法与 AHP 的权重确定

（1）选择的特征和参数

顾客购物信息记录了顾客从产生购物想法到完成交易过程中的一系列数据，其中包括很多无法计算和采集的信息。为了简化他们，在参考了卡拉曼的文章^[30]和 Whinston 的文章^[31]（这是被引用最多的人口统计学文献和电子商务文献）后，建立了 11 个最具有代表性和概括性的指标：姓名、年龄、性别、受教育程度、收入、职业类型、商品偏好、购物车商品购买率、电子商务平台偏好、登录次数、购买次数。

（2）熵法与 AHP 的结合

因为各个指标的价值对于 B2C 电商平台来说是不同的，因此，首先应该为以上指标设定科学合理的权重，以代表不同因素的价值。对数据进行了无量纲化后，为了科学合理地确定每个指标的权重，选择层次分析法（AHP）与熵方法相结合的方法。AHP 是科学赋权领域最常用的方法，它通过层次结构将复杂的多指标评估问题作为一个系统，将总目标分为多个子目标。虽然它结合了经验，知识等因素，但仍具有主观性太强的缺点，而这个缺点恰好可以被熵法弥补。熵法由 Shannon 提出，最初是热力学概念。应用熵可以测量索引系统中索引数据中包含的信息量，并确定每个索引的权重。熵方法充分挖掘了原始数据中包含的信息，结果更客观，但不能充分反映知识和经验，有时权重可能不符合实际重要性。因此，我们结合这两种方法来获得综合考虑权重向量的主观和客观指标的权重结果，即求得两种方法的结果的加权和：

$$\omega_{\text{总}} = \alpha \omega_{\text{AHP}} + (1 - \alpha) \omega_{\text{熵}} \quad (0 \leq \alpha \leq 1)$$

（3）专家打分

数据来源采用德尔菲法与访问调查法相结合的方法。德尔菲法本质上是一种反馈匿名函询法。其大致流程是：在对所要预测的问题征得专家的意见之后，进行整理、归纳、统计，

再匿名反馈给各专家，再次征求意见，再集中，再反馈，直至得到一致的意见。参与对象包括省级“双一流”高校电子商务领域教授、500 强电商企业中层和数据公司高层合计 9 人。在对访问调查法所获取数据求得几何平均值后，结合德尔菲法结果得到专家打分表。

（4）一致性检验

由于多阶判断的复杂性，往往使得判断矩阵中某些数值具有前后矛盾的可能性，即各判断矩阵并不能保证完全协调一致。当判断矩阵不能保证具有完全一致性时，相应判断矩阵的特征根也将发生变化，于是就可以用判断矩阵特征根的变化来检验判断的一致性程度。^[36]在层次分析法中，令判断矩阵最大的特征值为 λ_{\max} ，阶数为 n，则判断矩阵的一致性检验的指标记为：

$$CI = \frac{\lambda_{\max} - n}{n - 1}$$

CI 的值越大，判断矩阵的一致性越差。当阶数大于 2 时，判断矩阵的一致性指标 CI 与同阶平均随机一致性指标 RI 之比称为随机一致性比率，其中 RI 的值由表 10 确定，CR 的计算公式为：

$$CR = \frac{CI}{RI}$$

通过计算得到 $CR < 0.1$ ，因此可认为判断矩阵具有满意的一致性。

（5）指标权重

通过 AHP、熵法和两者的组合得到 11 个指标的权重如表 3-1 所示：

表 3-1 指标权重表

AHP		熵法		结合熵法与 AHP 的最终权重	
姓名	0.02	姓名	0.00	姓名	0.01
年龄	0.04	年龄	0.00	年龄	0.02
性别	0.03	性别	0.00	性别	0.02
受教育程度	0.07	受教育程度	0.11	受教育程度	0.09
收入	0.11	收入	0.47	收入	0.29
职业类型	0.09	职业类型	0.07	职业类型	0.08
商品偏好	0.09	商品偏好	0.05	商品偏好	0.07
购物车商品购买率	0.10	购物车商品购买率	0.01	购物车商品购买率	0.05
电子商务平台偏好	0.14	电子商务平台偏好	0.08	电子商务平台偏好	0.11
登录次数	0.15	登录次数	0.03	登录次数	0.09
购买次数	0.16	购买次数	0.18	购买次数	0.17

将这 11 个指标的权重用特征向量表示为：

$\omega_{i \in \{1,11\}} = (0.010, 0.018, 0.016, 0.093, 0.290, 0.079, 0.070, 0.053, 0.111, 0.090, 0.170)$

3. 1. 2 单个顾客购物信息价值模型构建

(1) 稀缺性指数

B2C 电商平台很难获取一个消费者的全部的 11 个消费者数据，因此，大部分的消费者的数据是不完整的。基于此，认为数据越稀缺，反映出其价值越高。这里引入资源的稀缺性指数：

$$b_i = \frac{1}{11} - \frac{Num_i}{\sum_{i=1}^{11} Num_i}$$

其中，Num_i 代表在一整个顾客购物信息包中 i 数据的总量。

(2) 数据成本

产品成本是指企业生产产品的各种成本，由固定成本和可变成本组成。交易平台收集、组织和分析数据，然后形成最终的数据产品与客户进行交易。因此，数据产品的成本由固定成本和边际成本组成。由于数据产品的固定成本较低，当顾客购物信息量较大时会被忽略。因此，生产数据的成本主要是由在 B2C 电商平台平台上生产、存储和共享数据的成本决定。

[37]

(3) 系数权重设置

如上所述，有许多因素会影响数据元组的值。设计一个覆盖所有影响因素的通用模型几

乎是不可能的。有两个最重要的因素：价值权重和稀缺性指数。设价值系数为 β ，稀缺性系数为 θ ，让他们满足以下约束：

$$\beta + \theta = 1$$

因此，第 j 个消费者的第 i 个指标的价值为

$$W_{ij} = C_j g(\omega_i \beta + b_i \theta)$$

C_j 为收集第 j 个消费者购物信息时产生的成本，进而得到第 j 个消费者的购物信息价值为

$$W_j = \sum_{i=1}^{11} C_j g(\omega_i \beta + b_i \theta)$$

（4）信息关联价值

对于提供的顾客购物信息，应进一步挖其数据价值。根据实际经验，顾客购物信息关联度越高，可以挖掘出越精准的消费需求，因此信息的价值越大。我们定义“信息完整性”来衡量信息的相关程度，“信息增值率”为信息关联价值。根据研究，两者之间的关系是“S 形”正相关，因此建立高斯概率密度函数用于准确量化的其关系。^[32]

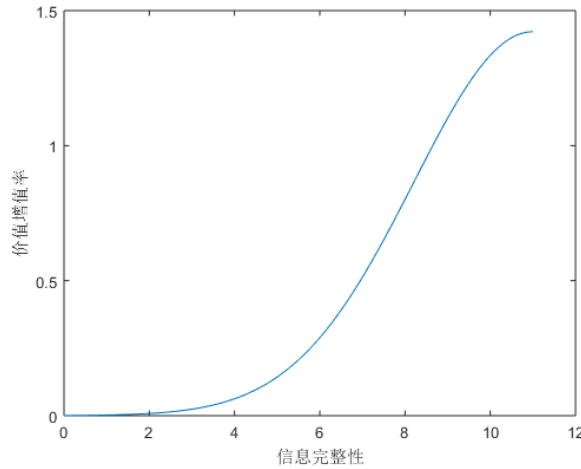


图 3-1 高斯函数曲线

信息完整性与信息增值率之间的关系与上面的高斯函数曲线趋势一致，如图 3-1 所示。

$$D(\sigma) = ae^{-\frac{(\delta-\mu)^2}{2\sigma^2}}$$

a, μ 和 σ 是参数，使用最大似然法^[33]来估计参数得到

$$\delta = 2.8, \quad \mu = 11, \quad a = 1.42$$

信息相关价值根据收入预期值计算，具体步骤如下：

$$D_j = W_j gD(\sigma)$$

(5) 信息即时性价值

顾客购物信息的即时性包括数据采集的时间节点和传输的时间节点，在数据源头采集数据后立即存储并立即加工，就是即时数据，而经过一段时间之后再传输或交易，则数据的即时性就稍差。数据的及时性是分析消费趋势和挖掘潜在客户的保障。

众所周知，顾客数据价值随着时间衰减。另一方面，一序列长度很长的数据应该具有附加价值。同样，第 j 个消费者的数据总量也会非线性地影响其价值。因此，可以得到信息的即时性价值：

$$V_j = e^{\varepsilon T - \tau t} gNum_i^\mu$$

其中，动态元素 T 表示数据的序列长度（/天）， t 代表数据的新鲜度（自生成数据以来的天数）。参数 ε, τ 和 μ 为影响指数和多项式的实数值，C.Shapiro 和 H.R.Varian 通过基于信息规则计算^[38]给出 $\varepsilon=1.28 \times 10^{-3}$ $\tau=9.50 \times 10^{-4}$ 和 $\mu=1.05$ 。

3.1.3 顾客购物信息的总价值

$$TW_j = V_j (W_j + D_j) = e^{\varepsilon T - \tau t} gNum_i^\mu \sum_{i=1}^{10} C_j g(\omega_i \beta + b_i \theta) g(1 + ae^{-\frac{(\delta - \mu)^2}{2\sigma^2}})$$

其中：

TW_j : 第 j 个顾客的购物信息总价值	W_j : 期望价值
D_j : 相关价值	C_j : 获取成本
ω_i : 指标价值权重	b_i : 指标稀缺性
δ : 信息完整性指标	T : 信息的序列长度
t : 数据的新鲜度	

3. 2 顾客购物信息数据包价值

在对个人购物信息价值进行准确计算后，我们需要建立一个包含多个个人购物信息的数据包的定价模型。但是，由于不同消费者的不同特征，不同的细分群体成员受到不同因素的影响，成员之间可能存在交叉影响。在模型中，我们基于这种交叉影响，引入网络效应的概念。

在此基础上通过结合 logistic 模型并引入数据饱和的概念，可以计算不同领域达到 K 值时所需样本总量，进而得到动态的数据包价值模型。

3. 2. 1 网络效应

对于单个消费者，其数据只能用于面对此消费者的精准营销，但是一个具有很多消费者数据的数据包还可用于消费者消费趋势、研究细分消费者特征等领域。我们称这个增值过程为消费者购物信息的网络效应。由于网络效应是一个随着消费者购物信息数据量的增长不断提高的过程，我们利用 Verhust 提出的 logistic 模型来表示社交网络效应。

$$P_j = \frac{KP_0e^{dj}}{K + P_0(e^{dj} - 1)}$$

其中可以用极大似然估计法求得 P₀=1, K=2.58

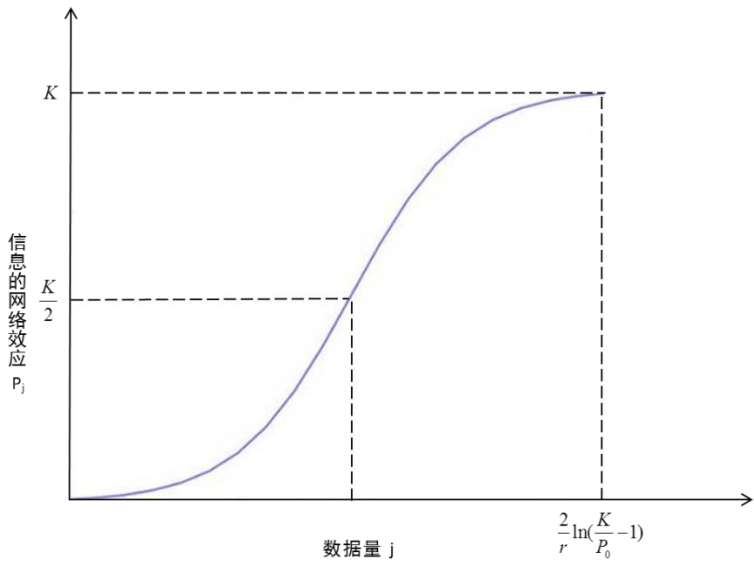


图 3-2 logistic 模型图示

3. 2. 2 数据饱和

数据饱和是来源于学术调研的一个概念。其最基础的理念是收集足够多的有意义数据来确定关键的问题及作出相应推荐方案。^[34]一旦提炼数据并确定了主要的问题后，现有数据仍无法确定主要的问题或样本带来的数据无法通过一致性检验，那么需要收集更多的数据。越丰富的数据越容易达到数据饱和，数据饱和后更多的数据无法带来数据的网络效应。一个数据包是否达到数据饱和可以用以下三个条件来进行判定：

- 1. 在有意义的基础上去决定样本量；

2. 达到数据收集的饱和；
3. 达到数据分析的饱和。

市场调研公司 Research by Design 的 Donna Bonde 提供了针对市场调研样本量的基础方针：

$$N = \frac{SF}{E} \pm R$$

其中，

F：调查的范围，这个范围可以是 1 到无限大中的任一个数字。0 指的是没有调研，这是违反用户体验设计的原则的。

C：研究的消费者特征，将每个不同消费者画像的多种代表或用户类型进行全面、细致的覆盖。

E：专业程度。理论上，这个值可以从 1 到无限大。现实中，这个范围应该在 1 到 2 。Victor 推荐每增加 5 年工作经验，这个值就增加 0.1，例如 5 年经验的研究员应该让公式中的专业程度（E）=1.10。

R：资源，指在市场调研时将预算和时间约束考虑在内。数字上，资源的值可以是 N-1 或 N+1 或更多（N 为预期的样本量）。

3.2.3 数据包总价值

$$TP = \begin{cases} \frac{KP_0 e^{dj}}{K + P_0(e^{dj} - 1)} g \sum_{j=1}^n TW_j, & \frac{2}{r} \ln\left(\frac{K}{P_0} - 1\right) \leq \frac{SF}{E} \pm R \\ K g \sum_{j=1}^n TW_j, & \frac{2}{r} \ln\left(\frac{K}{P_0} - 1\right) \geq \frac{SF}{E} \pm R \end{cases}$$

3.3 本章小结

本章提出了 B2C 电商平台下的顾客购物信息价值模型。首先，基于德尔菲法和访问调查法结合的方法使用熵法和 AHP 法确定顾客不同指标的价值权重，其次，通过成本加成定价法建立个人消费者的购物信息价值模型。然后，基于信息规则和高斯函数引入稀缺性指数、信息关联价值和信息即时性价值。在此基础上通过结合 logistic 增长函数并引入数据饱和的概念，可以计算不同领域达到 K 值时所需样本总量，进而得到动态的包含多个个人购物信息的数据包价值模型。

第 4 章 数据包交易定价策略

4.1 成本加成定价法

成本加成定价法是按产品单位成本加上一定比例的利润制定产品价格的方法。也就是 B2C 电商平台可以通过出售自己收集的顾客购物信息获得一定的利润。企业或政府机构想要购买这些信息,他们将支付给 B2C 电商平台一些数据价值以外的利润(根据目前的市场利润,除了奢侈品配件和暴利商品外,商品的利润率约为 10%-30%)。数据包的定价为

$$B = TP \times (1 + r)$$

r:期望利润率

TP: 数据包总价值

4.2 数据包价值的需求弹性

4.2.1 市场生命周期

根据生命周期的概念,将数据包的市场分为四个阶段,如图 4-1 所示

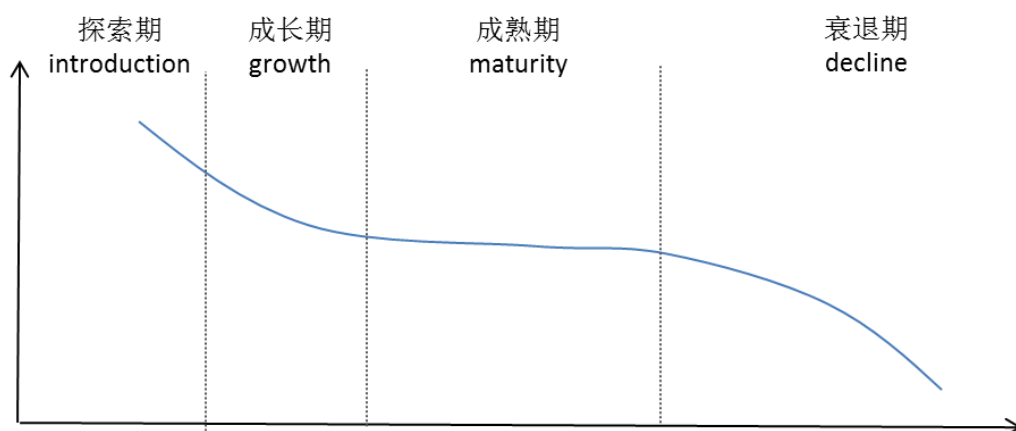


图 4-1 顾客购物信息市场生命周期图示

- 探索期 (introduction): 市场上多采用协议定价、拍卖定价、集合定价等方式进行交易,并存在第三方非法套利的情况,弹性系数很高.
- 成长期 (growth): 在此期间,价值与需求量之间的关系更加灵活. 当数据包价格上涨时,大多数平台都愿意出售,而当价格下跌时,平台会选择保密.
- 成熟期 (maturity): 随着数据包概念的深化成为商品,机构或组织对数据包有固定的需求,弹性系数会更小,甚至是固定价格.
- 衰退期 (decline): 当 B2C 电商平台的数据包被出售多年时,顾客购物信息市场接近饱和,并且可能存在市场容量归零的现象.

4.2.2 需求弹性函数

当数据包成为市场上的商品时,它将受到多种因素(主要包括数据包的可代替性、用途

是否广泛、耐用性)引起的市场波动的影响。这些影响将导致数据包的需求和价格的变化。为了描绘变化的程度和关系,引入经济学中需求价格弹性的概念来解释数据包价值的需求弹性^[35]。

$$E_d = \frac{\partial Q}{\partial B} \frac{B}{Q}$$

一般来说,弹性系数是负的,为了简单起见,我们将其视为正数。 它也分为下面五种情况:

- $E_d = 0$: 完全无弹性 (perfectly inelastic)
- $0 < E_d < 1$: 缺乏弹性、不富弹性, 或无弹性 (relatively inelastic)
- $E_d = 1$: 单位弹性、单一弹性, 或恒一弹性 (unitary)
- $1 < E_d < \infty$: 富有弹性 (relatively elastic)
- $E_d = \infty$: 完全弹性或完全有弹性 (perfectly elastic)

4.3 影响定价的其他因素

4.3.1 宏观环境分析

完成基本模型后,基于国家市场管控政策、隐私政策和经济学的外部效应对模型进行了进一步的优化。

(1) 市场管控政策

国家具有维护社会和谐,保证市场稳定发展的作用。在国家层面,国家不仅仅依靠市场中的供需关系或是产品本身价值来定价,国家将制定相关政策。这些政策实行后,不仅会影响到市场上数据包的供给量,也影响购买者的需求量。而这些变化无疑会对数据包的定价模式产生影响。

(2) 隐私政策

在某些情况下,出于公共安全或国家危机考虑,国家政府可能会出台一些法律或法规来保护某些某一类特殊的消费群体。而这些类型的顾客购物信息的可替代数据包可能会供不应求,进而导致价格上涨。

4.3.2 外部效应

不能忽视这样一个事实,即当数据包广泛应用于电子商务的各个领域时,部分价值或成本可能并非由其所有利益相关者所有。其他人可能同信息收集者一起分享利益或成本。这些价值或成本被称为经济学中的外部效应。

数据包的外部影响是多种多样的,例如,B2C 电商平台滥用消费者信息可能会导致整个社会对行业的不信任和其他数据公司的损失。再例如通过收集消费者购买食品的频率和类别可以实现健康风险管控,提醒消费者定期体检可能会使医疗机构、交通机构获利。

基于以上三点,为了准确构建数据包的定价模型,我们在上面的定价模型中加上修正系数 μ 得到修正后的数据包价格模型:

$$B = TP \times (1+r) + \mu$$

4.4 价格的确定与控制

4.4.1 价格的确定

基于成本加成定价法下的定价模型能够解决信息不对称的问题，更好的兼顾顾客购物信息商品提供方的利益（B2C 电商企业），活跃交易市场，丰富交易机制，促使顾客购物信息交易市场朝着健康稳定的方向发展。此种定价模式可用于一般类顾客购物信息的交易模式，可以大批量成交，顾客购物信息的可获得性和可复制性强，对数据需求方没有限制，可以让更多的人得到数据。综上所述，对于一个包含 j 个顾客购物信息的数据包的参考价格为：

$$B = \begin{cases} \frac{K(1+r)P_0 e^{dj}}{K + P_0(e^{dj} - 1)} g \sum_{j=1}^n TW_j + \mu, & \frac{2}{r} \ln\left(\frac{K}{P_0} - 1\right) \leq \frac{SF}{E} \pm R \\ K g(1+r) g \sum_{j=1}^n TW_j + \mu, & \frac{2}{r} \ln\left(\frac{K}{P_0} - 1\right) \geq \frac{SF}{E} \pm R \end{cases},$$

$$TW_j = V_j(W_j + D_j) = e^{\varepsilon T - \tau t} gNum_i^\mu \sum_{i=1}^{10} C_j g(\omega_i \beta + b_i \theta) g(1 + ae^{-\frac{(\delta - \mu)^2}{2\sigma^2}})$$

其中：

B: 数据包的参考价格	r: 期望利润率
TW_j : 第 j 个顾客的购物信息总价值	C_j : 获取成本
ω_i : 指标价值权重	b_i : 指标稀缺性
δ : 信息完整性指标	T: 信息的序列长度
F: 调查的范围	C: 消费者特征
E: 专业程度	R: 资源
t: 数据的新鲜度	

4.4.2 价格的执行与控制

价格制定后就是执行和控制价格，这是定价循环周期中的最后一个步骤。对于 B2C 电商平台来说，价格的执行主要指价格表的公布及与数据包买方、销售人员、中间商（交易平台）、报价人员、开发票的人员等相关者就价格进行的沟通，以保证价格政策顺利执行。

价格控制就是企业在价格执行过程中仔细分析数据包买方、中间商和竞争者对本企业价格的反应，如果他们的反应与预期的不同，就要采取控制行动^[40]。首先明确导致价格策略失效的原因并采取纠偏措施，如果局部纠偏措施无效，就要仔细分析影响定价的内外部因素，如果因素的变化对公司的战略产生明显的影响，那么就应该对整个价格策略进行调整，如此这般，就进入定价流程的下一个循环。

在交易中，企业应该尽可能提高自己的评价来获得更高的预期收益。具体提高自己的评价的方式有遵守交易平台规范、积极进行交易资质验证、提高数据质量、做好数据标准化的处

理工作、诚信对待每一次数据交易，提高自己的历史交易信用。

最后，交易平台要实行交易透明机制，交易过程公开透明，公开成交价格，形成市场标准化，对以后的顾客购物信息交易也可起到案例示范作用，对同类数据或信息商品提供交易价格参考，促进交易市场健康发展。对积极参与交易的顾客数据需求方和提供方进行奖励，在需求方交易平台中，对数据提供方进行评分，对以后的交易报价进行约束，也可以在交易过程中间接对信用良好或不好的需求者进行鼓励或者惩罚。同时应该对需求者的恶意购买行为进行监管，防止多名顾客数据需求者联合起来进行恶意压价，交易系统后台应该实时监控不同顾客数据需求者的活动轨迹，如果相互关联就视为作弊，降低其评价。对于数据供给，也可以针对性的提出相应的评分机制，来对其行为进行约束。

4.5 本章小结

当 B2C 电商平台将顾客购物信息数据包（以下简称数据包）作为商品出售时，其价格不应该根据其价值进行确定，定价是与公司战略密切相关的管理行为，是支持企业战略目标实现的策略性活动。价格策略应当服从企业的总体战略，企业战略的规范性和动态性，决定了企业的定价决策是个科学的流程，同时也是个往复循环的过程，价格应随企业内外部环境的变化而随时调整。按逻辑顺序，本章结合 Philip Kotler 的定价流程^[39]将基于战略视角下的企业定价流程分为五个步骤：

1. 确定定价方法
2. 估计市场需求
3. 估算商品价值
4. 影响定价的其他因素
5. 价格的确定与执行价格

这五个步骤构成一个闭环，其中的每一个步骤既是上一步的结果，又是下一步的基础。下面将分别对该定价流程的除估算商品价值外的四个步骤进行阐述。

参考文献

- [1]李骥宇.大数据交易模式的探讨[J].移动通信,2016,05:41-44
- [2]高丹,向阳.中国大数据市场的特点与趋势[J].中国工业评论,2016,07:50-55.
- [3]刘朝阳.大数据定价问题分析[J].图书情报知识,2016(1):57-64.
- [4]张晓玉.基于讨价还价博弈的大数据商品交易价格研究[D].辽宁:辽宁科技大学,2016.
- [5]黎春兰,邓仲华,张文萍.云服务的定价策略分析.图书与情报,2013(1):36-41
- [6] Mowshowitz A. On the theory of virtual organization[J], Systems research and behavior a science, 1997, 14(6):373-385.
- [7]Jean. Tirole 泰勒尔.产业组织理论[M].张维迎译.北京:中国人民大学出版社,1997
- [8]卡尔·夏皮罗,哈尔·瓦里安.信息规则—网络经济的策略指导[M].张帆译.北京:中国人民大学出版社,2000.06
- [9] Muschalle A, Stahl F, Lser A, et al. Pricing approaches for data markets//Proceedings of the Workshop Business Intelligence for the Real Time Enterprise.Istanbul,Turkey,2012:129-144
- [10] Balazinska M, Howe B, Suciu D. Data markets in the cloud: An opportunity for the database community. Proceedings of the VLDB Endowment,2011,4(12):1482-1485
- [11] Mankiw N G .Principles of Economics. Ohio, USA: South-Western Cen gage Learning,2012
- [12] Balazinska M, Howe B, Koutris P, et al. A discussion on pricing relational data. Lecture Notes in Computer Science,2013:167-173
- [13] Li C, Li D Y, Miklau G, et al. A theory of pricing private data //Proceedings of the 16th International Conference on Database Theory.Genoa,Italy,2013:33-44
- [15] [1]彭慧波,周亚建.数据定价机制现状及发展趋势[J/OL].北京邮电大学学报, 2019(01): 1-8[2019-04-01]. <https://doi.org/10.13190/j.jbupt.2018-108>.
- [16]陈筱贞.大数据交易定价模式的选择[J].港澳经济,2016(18):3-4.
- [17]张晓玉.基于讨价还价博弈的大数据商品交易价格研究[D].鞍山:辽宁科技大学,2016.
- [18]RIEDERERC, ERRAMILLIV, CHAINTREAU, et al. For sale: your data :by: you[C] //ACM Work shop on Hot Topics in Networks. New York: ACM,2011: 13.
- [19]MUSCHALLEA, STAHLF, LÖSERA, et al. Pricing approaches for data markets [M]//CASTELLANOSM, DAYALU, RUNDENSTEINEREA. Enabling real-time business intelligence. Berlin: Springer, 2012:129-144.
- [20]干春晖,钮继新.网络信息产品市场的定价模式[J].中国工业经济,2003(5):34-41.
- [21]NIYATOD, ALSHEIKHMA, WANG P, et al. Market model and optimal pricing scheme of big data and Internet of things (IoT)[C]// IEEE International Conference on Communications. New York: IEEE Press, 2016:16141810.

- [22]杨琪,龚南宁.我国大数据交易的主要问题及建议[J].大数据,2015,1(2):38-48.
- [23]BALAZINSKAM, HOWEB, DANS. Data markets in the cloud: an opportunity for the data base community [J].Proceedings of the VLDB Endowment, 2011,4:1482-1485.
- [24]SHEN Yun cheng, GUO Bing, SHEN Yan, et al. A pricing model for big personal data [J]. Tsinghua Science and Technology, 2016, 21(5): 482-490.
- [25]TANG Ruiming, WU Huayu, BAO Zhifeng, et al. The price is right [M]//DECKERH, LHOTSKAL, LINKS, et al. Data base and expert systems applications. Berlin: Springer,2013:380-394.
- [26] 王 玮 . 探 索 数 据 交 易 模 式 、 产 业 生 态 及 合 规 需 求 [OL]. <http://www.csdn.net/article/2015-12-15/2826481>, 2015-12-15
- [27] 胡 嘉 琪 . 数 据 交 易 史 话 : 隐 私 、 定 价 、 交 易 与 策 略 初 探 [OL]. <http://www.36dsj.com/archives/37614>,2015-02-01.
- [28]阿里研究院.数据交易面临 5 大挑战:能否避免数据集市沦为数据黑市 [OL].http://www.gygov.gov.cn/art/2016/1/7/art_22021_875745.html?2016-01-07
- [29]徐广斌,牛壮.大数据交易产品及交易机制创新实践[R/EB].上海证券交易所,2015-09.
- [30] Nada Karaman Aksentijević. Demographic Characteristics[J]. Physical ReviewA, 2015, 35(12):5274-5277.
- [31] Kalakota R, Whinston A B. Frontiers of electronic commerce[J]. 1996.
- [32] Sun Guangling, Tang Xianglong. Research and development of semi supervised learning algorithm [J]. layered computer based on Gauss mixture model,2004(01):159-164
- [33] Zhang Rongquan, Du Yuming, Yang Jianyu. A LFM signal maximum likelihood estimation model and a fast algorithm for parameter estimation [J]. Journal of radio wave science,2005 Journal of radio wave science (05):101-105.
- [34]Victor Yocco. Filling Up Your Tank, Or How To Justify User Research Sample Size And Data [J]. Smashing Magazine, March,30,2017.
- [35] Diego S. Price elasticity of demand[J]. Betascript Publishing, 2009, 3(4):1717-1718.
- [36] 郭东硕,程正敏,彭茜.基于 Matlab 的层次分析法及其运用浅析[J].中小企业管理与科技(上旬刊),2010(11):269-270.
- [37] Yuncheng Shen,Bing Guo,Yan Shen,Xuliang Duan,Xiangqian Dong,Hong Zhang.A Pricing Model for Big Personal Data[J].Tsinghua Science and Technology,2016,21(05):482-490.
- [38] C. Shapiro and H. R. Varian, Information Rules: A Strategic Guide to the Network Economy. Harvard Business Press, 1998.
- [39] 菲利普·科特勒:《营销管理一分析、计划、执行和控制》(第九版),上海人民出版社,1999年版。
- [40]田志龙,衣光喜.战略视角下的企业定价流程[J].经济管理,2004(21):54-57.

