

“大数据杀熟”的法律规制研究*

田 蕾,洪 晟

(北京航空航天大学 网络空间安全学院,北京 100083)

摘 要: 随着大数据技术发展,“大数据杀熟”现象日益严重,侵害了消费者的诸多合法权益。针对该现象,在充分研究了其成因和目前规制困境的基础上,提出了完善相关法律规制的建议,使大数据能够更好地发挥其应有价值,有助于在互联网平台利益与消费者合法权益保护之间找到平衡点。

关键词: 大数据杀熟;大数据;法律;消费者权益

中图分类号: TP393.08;D922.11;D922.17 文献标识码: A DOI: 10.20044/j.csdg.2097-1788.2022.01.007

引用格式: 田蕾,洪晟. “大数据杀熟”的法律规制研究[J]. 网络安全与数据治理, 2022, 41(1): 41-46.

Research on legal regulation of the Big Data affinity

Tian Lei, Hong Sheng

(School of Cyber Science and Technology, Beihang University, Beijing 100083, China)

Abstract: With the development of Big Data technology, the phenomenon of “Big Data Affinity” is getting more serious, infringing on legitimate rights and interests of consumers. Therefore, this paper puts forward suggestions for improving relevant legal regulations on the basis of fully studying its causes and current regulatory dilemmas, so that Big Data can better play its due value. This is also helpful for finding a balance between the benefits of Internet platforms and the protection of rights and interests of consumers.

Key words: Big Data affinity; Big Data; legal system; consumers' rights and interests

0 引言

近年来,随着电商平台规模的不断增长,互联网经济模式愈发多样,我国居民已经逐渐形成了全新的网络消费习惯,互联网消费已成为大势所趋^[1]。然而,互联网消费数据由于具有体量大、隐私性强等特点,其在流动过程中保护难度较大。而一旦这些数据被泄露或滥用,其后果将不堪设想。因此,对互联网消费数据空间进行治理已成为当务之急。

目前,在消费互联网中广泛使用了大数据技术,该技术已有近 20 年的发展历史。2005 年,雅虎公司开发了 Hadoop 项目,使对结构化数据的快速、可靠分析变为现实^[2]。2008 年,计算社区联盟发表白皮书^[3],最早提出大数据概念。2011 年,麦肯锡全球研究院在报告^[4]中对大数据进行了全方面介绍和展

望,标志大数据时代的正式到来。2017 年,《大数据安全标准化白皮书》正式发布^[5],加速了大数据落地深耕,促进了行业生态发展。

大数据技术完美符合互联网时代的数据特点,近年来备受重视。然而随着该技术的广泛应用,“杀熟”现象逐渐产生。本文在对“大数据杀熟”现象进行深入分析的基础上,指出了“大数据杀熟”所存在的法律问题并对其法律规制提出建议。

1 “大数据杀熟”概述

1.1 “大数据杀熟”现象

随着大数据技术的日渐成熟,其与生活联系越发紧密,但同时也引发了新的问题,其中较为突出的便是对公民个人信息的侵犯^[6],危害很大。“大数据杀熟”现象是指:许多互联网平台利用大数据技术收集并分析消费者的消费信息,并以此对不同消费者进行区别定价^[7]。这样的行为严重损害了消费

* 基金项目:北京航空航天大学研究生精品课程建设项目(403918)

者的合法权益,后果十分严峻。

亚马逊早在 2000 年就被爆出存在“大数据杀熟”现象。在亚马逊网站进行购物时,一些用户发现某些商品在删除 cookie 记录前后存在价格不一致的情况^[8]。2014 年,国外研究者^[9]调查了多家电商平台后发现,在一些平台上,用户使用不同的电子设备搜索出的结果不尽相同,同一酒店的住宿价格也有所不同。2017 年,加拿大新闻媒体 CBC NEWS^[10]的调查人员以浏览器、设备、国籍等为变量对同一酒店进行搜索,结果显示搜索出的价格不尽相同。一般而言,匿名浏览器、手机搜索出的价格更便宜,而美国籍账号所搜索出的价格则更贵。也就是说,这些客观因素的不同都会影响商品的价格。这在国外引起了广泛的社会关注。

2018 年,国内首次爆出“大数据杀熟”现象。一位微博网友爆料称其在某一网站预订酒店的价格会比其不经常使用该网站的朋友的账号更贵。此后,陆续有网友反映:使用打车、外卖等软件时老用户的价格比不经常使用该软件的用户和新用户高。例如,滴滴熟客张女士试验发现,自己打车时所显示的价格确实比不经常打车的丈夫高出很多。

北京市消费者协会所发布的关于“大数据杀熟”的调查问卷结果显示:在所有调查对象中,认为该现象非常普遍的占 88.3%,而仅有 11.7%认为该现象不普遍或一般。值得注意的是,所有被调查者均承认了“大数据杀熟”现象的存在。被调查者在各领域遭受“大数据杀熟”的情况如图 1 所示。

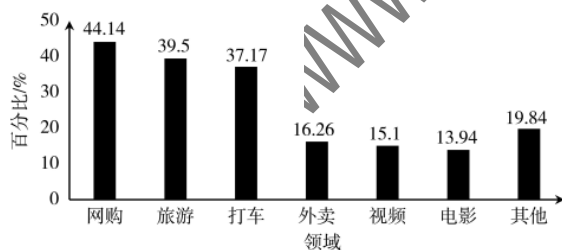


图 1 各领域“大数据杀熟”现象统计

现阶段,“大数据杀熟”现象颇有持续发酵的趋势,这主要有两个原因:(1)大数据技术的逐渐成熟使得互联网平台获取用户数据的成本和门槛越来越低;(2)由于互联网领域是近年来的新兴领域,因此我国尚未形成一套完善的法律体系。这会导致用户对互联网平台的不信任感与日俱增。因此,从长远发展来看,国家非常有必要采取适当措施对消费

者个人权益进行保护。

1.2 “大数据杀熟”的定义

相较传统的“杀熟”而言,“大数据杀熟”广泛利用了大数据技术,互联网平台通过收集和分析消费者个人信息和消费信息,可以描绘出客户的消费习惯等特征,在此基础上形成用户画像,进而以此为依据调整商品的价格,进行差异定价,从而尽可能达到利益最大化^[11]。

“大数据杀熟”的核心基础是对个人信息的不当收集和使用,其本质是一种数据滥用,这严重侵犯了消费者的合法权益。目前,学界对“大数据杀熟”的定义主要有两种^[12],一种强调杀熟目的^[13],即行为主体具有明显的获利倾向;另一种则不强调杀熟目的。前者认为“大数据杀熟”现象的核心是经营者为了追求私利,基于大数据技术所形成的用户画像,从而对不同用户采用不同的定价方式。正是因为经营者的获利倾向明显,暗中损害消费者利益,因此称之为“杀熟”。后者则并未事先进行目的预设,认为“大数据杀熟”仅仅是一种价格差异现象,而该现象只是由于所使用的定价算法所导致的,而并不需要考虑这样的差异是否是由经营者的获利倾向所导致的,仅仅着眼于差别定价现象本身。值得注意的是,两种观点均将“大数据杀熟”作为价格歧视行为来看待。詹好^[14]等指出非价格歧视也可能导致价格歧视行为,但由于区别定价现象属于价格歧视的表现形式,因此目前学界普遍认为将其定性为价格歧视更具有合理性。

通过以上分析,“大数据杀熟”的主要特点^[15]为:(1)互联网平台利用大数据收集、分析用户信息;(2)不同用户购买同一商品的价格有所不同;(3)熟客遭受“杀熟”的概率更高。

1.3 “大数据杀熟现象”的成因

目前的研究调查显示,“大数据杀熟”现象的成因主要有:交易双方信息不对称、用户数据获取成本较低、定价标准难以解释、行业相关法律法规尚未完善等。

(1)交易双方信息不对称

交易双方信息不对称主要指的是在交易过程中,互联网平台和用户获取信息的渠道有很大差别,这导致双方最终获取的信息差别也很大。总的来说,互联网平台相对于消费者而言掌握更多的市场信息,并且占有更多技术资源,从而在交易过程中

具备明显优势。

互联网平台获取数据的渠道非常丰富,且由于技术的发展,其获取的数据规模也非常庞大,用户在注册和使用平台的过程中都会为其提供丰富的数据。利用大数据技术,平台可以迅速收集和分析信息,从而在很短的时间内就能掌握一定的规律。而消费者由于时间、工具等客观因素限制,只能以自己手机、电脑页面上所展示的极为有限的信息为基础决定购买哪些商品,无法对商品有全面了解。同时,消费者在交易过程中只能看见自己购买该商品所需支付的价钱,而无法与其他用户进行比较。

此外,互联网平台还会采取其他手段例如设定优惠券等方式隐藏价格,这进一步增加了消费者比对价格的难度^[11],从而加剧交易双方信息的不对称程度。

(2) 用户数据获取成本较低

数据只有在被收集利用后才会具备一定的经济价值。消费者在互联网平台交易过程中,除涉及个人隐私的信息外,产生的其他数据对个人而言并无价值。因此消费者往往不够重视对自身数据的保密,这使得互联网平台很容易进行数据收集。而这些数据在形成一定规模后则具有重要参考意义。另外,随着大数据技术的发展,互联网平台收集数据的技术也更加高效,进一步降低了收集用户数据的成本。

(3) 定价标准难以解释

目前,互联网平台主要的定价方式为:采用深度学习、数据挖掘等技术,形成一套通用的定价算法,而这些算法是不透明的,对用户而言是不可见的。平台可以完全控制和决定商品的价格,而用户只能被动地接受价格,无法影响价格的制定。受各种因素影响,商品的定价标准对用户而言并不透明,用户无法分析其价格和成本之间的关系^[11]。

(4) 行业相关行规、法规尚未完善

互联网是新兴领域,电商平台等更是近年来才飞速发展,因此行业并未形成一套成熟的体系。一些平台在前期通常会采用各种优惠方式来吸引用户,而在后期,为了弥补企业亏损,则会利用一系列方式如差异化定价等来追求利益最大化。互联网行业尚未形成良好的竞争氛围,而关于电商平台定价的相关法律法规也尚未完善,因此“大数据杀熟”现象治理颇具难度。

1.4 “大数据杀熟”的运作模式

目前“大数据杀熟”的一般运作模式^[8]为:信息收集、数据分析、动态差异化定价。

(1) 信息收集

互联网平台收集用户信息的主要方式有:(1)用户在注册时填写个人信息;(2)用户在使用平台时产生的全部交易信息;(3)利用技术手段收集用户在互联网上的公开信息。这些信息均为之后的数据分析提供了海量素材。

(2) 数据分析

通过大数据挖掘等技术,互联网平台能够尽可能多地分析整理用户的消费习惯、偏好等数据,形成所谓的用户画像,在此基础上针对性地向某一用户推荐一些特定的商品,对其进行诱导,提高该用户的消费可能性,从而获得更高利润。

(3) 动态差异化定价

根据用户画像,互联网平台会通过各种手段进行价格调整以达到“杀熟”目的。例如,通过对用户搜索某个关键词的次数和频率等来推断出用户的迫切程度,以此为依据进行价格调整。而用户一旦进入了互联网平台的某些杀熟路径,便很容易被多轮杀熟。

2 “大数据杀熟”的法律问题

2.1 “大数据杀熟”违法性分析

“大数据杀熟”现象中的差异化定价是建立在价格歧视原理上的,这虽然在一定程度上能够使得社会总产出有所增加,但是却是对消费者合法权益的严重侵害^[16]。

(1) 侵害消费者的知情权

《消费者权益保护法》第八条规定,消费者享有对其所购买的商品的真实情况的知悉权。

然而现实中却存在着严重的信息不对称问题。用户对商品的全部所知信息均来源于互联网平台,且其只能看见自己最终支付的价格,很难获知他人的支付价格,也无法判断价格上的出入是由何种原因引起。互联网平台为了逐利,利用用户的价格不敏感心态,只提供商品部分信息,对消费者全面了解商品造成阻碍,这实质上是对消费者知情权的侵害。

(2) 侵害消费者的公平交易权

《消费者权益保护法》第十条规定,消费者享有公平交易的权利^[17]。尽管这里的公平并不是指所有用户购买同一商品的价格必须完全相同,但是在现

代法中对差别定价进行了明确规定,其必须遵守三原则^[18]。但从目前的“大数据杀熟”现象来看,差别定价往往仅仅由消费者的消费习惯所引起,不是正当理由,这是对消费者公平交易权的严重侵害。

(3)侵害消费者的自主选择权

《消费者权益保护法》第九条规定,消费者享有自主选择的权利。马爽^[19]指出,该权力最突出的特征为自愿性和自由性。前者体现在:消费者在进行消费决策时应是完全自愿的,任何外在因素都不能对其进行干扰;后者则体现在:消费者在消费时自己拥有充分的决定权,其意愿不会因为外在因素而被迫改变。

而在大数据技术下,互联网平台所展现给用户的信息均是经过处理的信息,其可以完全控制展示哪些信息和怎么展示信息。平台所展示出的信息往往具有明显导向性,这实际上是对消费者进行正常、合理选择的严重干扰,是对消费者自主选择权的侵害。

(4)侵害消费者的隐私权

我国《民法总则》中明确了对个人信息的保护。这要求对个人信息的收集和使用必须严格遵循特定的原则,且应有意识地对其进行严格保护^[20],不得以任何理由在未经用户同意的情况下泄露或出售这些信息^[21]。当今时代,用户在使用互联网平台进行交易时,需要提供相关个人信息且会形成交易记录。这些信息都可能会被平台收集,而平台怎样收集信息、收集哪些信息对消费者而言是隐蔽的,严重侵害了消费者的隐私权。

2.2 “大数据杀熟”法律规制困境

“大数据杀熟”行为是违法行为,对其采取法律手段进行规制有法可依,但是在实际规制中却存在诸多问题,如:“大数据杀熟现象”难以界定、现行法律使用困难、缺少对算法权力的制约、举证难度大、维权成本高等。

(1)“大数据杀熟”行为难以界定

要用现行法律体系对“大数据杀熟”予以规范就要首先证明其存在。然而,目前对于大数据“杀熟”行为的揭露大多是消费者的自发行为,很难确保消费者在自行测试是否存在大数据“杀熟”行为时,有没有对实验变量进行严格控制,也无从考证其实验结果是否可信。另一方面,基于市场供求关系所引起的价格正常涨落和互联网平台的“大数据

杀熟”在结果上均表现为导致了价格变动,如何对这两者进行区分尚无明确的判断标准^[16]。

(2)现行法律使用困难

虽然我国近年来出台了一系列法律法规以尽可能在大数据时代保证消费者的合法权益,但在具体应用过程中却存在较大难度。

朱建海^[22]指出,“大数据杀熟”的实现的两个基本要素为对消费者的信息进行大量收集和以此为基准所实现的差别定价。前者可能会涉及对消费者隐私权的侵犯,但是由于个人隐私的界定尚且模糊,因此究竟哪些信息可以纳入个人隐私的范畴还存在较大争议。不能因为经营者收集了消费者的消费信息等就将其认定为侵犯了消费者的隐私权。后者则可能会涉及对消费者的公平交易权和知情权的侵犯。同样,目前对“公平”和“知情”的定义也较为模糊,究竟怎样的交易才算是公平,究竟哪些信息才是消费者应该知道的,目前尚无统一标准。因此,在经营活动中,若仅因为经营者选择性地向消费者展示了其想要展示的信息(即使这些信息具有一定的导向性)就将其认定为侵权行为,这在法律上并不可行。由此可见,现行法律在实际使用时面临诸多问题,难以对“大数据杀熟”其进行有效规制。

(3)缺少对算法权力的制约

互联网平台基于大数据技术所进行的“用户画像”是其进行“大数据杀熟”的关键,然而其算法的不透明性和隐蔽性使得监管机构和社会公众很难进行有效监督。

(4)举证难度大,维权成本高

根据我国《民事诉讼法》规定,在提起诉讼时所采用的举证原则为:“谁主张,谁举证”。因此,若消费者要对其遭受“大数据杀熟”的不公平经历提起诉讼,其必须要提供大量有力证据。然而由于信息的不对称,消费者仅持有自己的消费记录等,而“大数据杀熟”算法则为平台经营者所持有,故消费者很难提供充足证据。

另外,在实际生活中,即使消费者遭受“大数据杀熟”,若其为了维护自身合法权益而向法院提起诉讼,往往需为此付出比自己损失高得多的时间和精力,得不偿失。

3 “大数据杀熟”的法律规制建议

近日,国家网信办等四部门联合发布《互联网信息服务算法推荐管理规定》,于3月1日起正式

施行。在该规定中明确指出,为逐步规范算法推荐服务,有关部门应开展三方面工作:(1)加强监管工作,建立健全算法机制机理审核;(2)推动实现算法公开化、透明化;(3)加大用户权益保护力度等。

值得注意的是,该规定中只对算法权力进行了定性约束,在实际应用中仍存在一定的模糊性,有待进一步细化。

从我国现阶段来看,较为可行的解决方法有:进一步完善对算法权力的约束、加强社会监督与行政监管、提高消费者自身维权意识、细化现有法律法规等。

3.1 完善对算法权力的约束

(1)加强对算法权力的监管

对于算法权力的监管,域外模式已经取得了较大成效,具有代表性的有:美国模式和欧盟模式。前者核心为“算法问责”,后者核心为“数据保护”,两者均具有启示意义。在今后的法律规制过程中,我国可以充分借鉴这些成功经验,进一步明确算法和数据使用的权力,定期对算法中相关部分的参数等进行检查,以规范算法的使用。

(2)促进新兴技术的发展

随着科技不断发展,“数据可用不可见”的概念也受到了广泛关注。近年来,安全多方计算、隐私计算等技术对于实现数字资产化具有重要意义,这些技术在有效保护用户隐私信息等的同时也使得大数据能得到充分利用,这有助于在互联网平台利益与消费者合法权益保护之间找到平衡点。

3.2 加强社会监督与行政监管

朱喆峰^[23]指出,在对“大数据杀熟”行为进行规制时,既需要完善现有的法律体系,还需要加强监督与监管。可对互联网平台实施自律与监督相结合的方式,推动行业自身形成良好的竞争氛围,必要时可通过监督手段督促其进行整改,以使个人信息保护与互联网平台发展形成双赢。

3.3 提高消费者自身维权意识

胥雅楠^[15]指出,在“大数据杀熟”现象频出的现状下,消费者应提高自身维权意识,同时应注重货比三家,提升对价格变化的敏感度,降低对单一平台的依赖。另外还要注重保护个人隐私信息。在自身权益收到侵害时,要坚定维权,否则只会让“杀熟”现象更加严重。

3.4 细化现有法律法规

现有法律法规的模糊性使得其在实际应用中存在困难,立法机关应当及时完善现行法律法规。阎文聪^[24]对不同法律法规提出了不同的完善方式:(1)对于《反垄断法》,应完善构成要件认定标准,降低认证难度;(2)对于《消费者权益保护法》,应拓宽消费者知情权的外延,降低信息的不对称程度;(3)对于《价格法》,应完善不正当价格行为的构成要件,规范对商品的定价;(4)对于《电子商务法》,应详细解释相关规定,以避免“大数据杀熟”所带来的不可预知的风险。

4 结论

“大数据杀熟现象”的愈演愈烈严重恶化了消费者与互联网平台之间的关系,影响互联网经济发展,当务之急是采用有效法律规制手段以遏制该现象的进一步发展。《互联网信息服务算法推荐管理规定》^[27]已从监管层面和算法推荐服务提供者层面提出了规范建议,除了进一步细化相关法律法规外,促进新兴技术发展、提高消费者自身维权意识等也具有重要意义。相信在全社会的共同努力之下,大数据一定能更好地发挥其应有的价值。

参考文献

- [1] 赵文莉.互联网经济对推动我国居民消费的作用探究[J].经济师,2020(11):55-56.
- [2] 大数据技术概述[EB/OL].[2020-12-21].https://blog.csdn.net/m0_52879270/article/details/111466007.
- [3] BRYANT R, KATZ R H, LAZOWSKA E D. Big-data computing: creating revolutionary breakthroughs in commerce, science and society[J]. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'13, 2013: 1-15.
- [4] MANYIKA J, CHUI M, BROWN B, et al. Big data: the next frontier for innovation, competition, and productivity[M]. McKinsey Global Institute, 2011.
- [5] 《大数据安全标准化白皮书》正式发布[J]. 信息网络安全, 2017(4): 98.
- [6] 邓华. 大数据背景下侵犯公民个人信息犯罪研究[J]. 信息网络安全, 2020(S1): 63-66.
- [7] 金幼芳, 王凯莉, 张汀菡. 《个人信息保护法》视角下“大数据杀熟”的法律规制[J]. 浙江理工大学学报(社会科学版), 2021, 46(6): 693-701.
- [8] 张海玥. “大数据杀熟”的法律规制研究[D]. 兰州: 兰州大学, 2020.

- [9] HANNAK. Measuring price discrimination and steering on E-commerce web sites[C]. Conference on Internet Measurement Conference, 2014: 305-318.
- [10] How companies use personal data to charge different people different prices for the same product[EB/OL]. [2020-02-28]. <https://www.cbc.ca/news/business/marketplace-online-prices-profiles-1.4414240>.
- [11] 朱昱潼. “大数据杀熟”现象的成因与对策[J]. 中国经贸导刊, 2021(19): 62-63.
- [12] 贺桂华, 尚玉萌. “大数据杀熟”视野下消费者权益保护研究[J]. 经济研究导刊, 2021(28): 157-160.
- [13] 邹开亮, 刘佳明. 大数据“杀熟”的法律规制困境与出路——仅从《消费者权益保护法》的角度考量[J]. 价格理论与实践, 2018(8): 47-50.
- [14] 詹好, 邵靳天, 黄智威. “大数据杀熟”: 概念澄清及解决方案[J]. 软件, 2019, 40(8): 62-65.
- [15] 胥雅楠, 王倩倩, 董润, 等. “大数据杀熟”的现状、问题与对策分析[J]. 改革与开放, 2019(1): 15-20.
- [16] 毕思琪. 大数据“杀熟”的法律规制路径研究[D]. 上海: 华东理工大学, 2020.
- [17] 李昌麒, 许明月. 消费者权益保护法[M]. 北京: 法律出版社, 2014.
- [18] 范进学. 论宪法比例原则[J]. 比较法研究, 2018(5): 108-112.
- [19] 马爽. 大数据“杀熟”中的消费者权益保护[D]. 石家庄: 河北经贸大学, 2021.
- [20] 陈明茗. 我国个人信息权的权利内容研究[D]. 重庆: 西南政法大学, 2014.
- [21] 吴宏伟. 《消费者权益保护法》[M]. 北京: 中国人民大学出版社, 2014, 28.
- [22] 朱建海. “大数据杀熟”的法律规制困境及其破解路径[J]. 重庆邮电大学学报(社会科学版), 2021, 33(1): 64-72.
- [23] 朱喆峰. 大数据“杀熟”行为的法律规制[D]. 武汉: 华中师范大学, 2020.
- [24] 阎文聪. 大数据“杀熟”行为法律规制研究[D]. 兰州: 甘肃政法大学, 2021.

(收稿日期: 2022-02-23)

作者简介:

田蕾(2001-), 女, 本科, 主要研究方向: 网络信息安全。

洪晟(1981-), 通信作者, 男, 博士, 副教授, 博导, 主要研究方向: 网络信息安全、复杂系统安全运行状态监测与健康管理, 复杂系统通用质量“六性”技术。
E-mail: shenghong@buaa.edu.cn。



版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.cn