

融媒体环境下的受众计算：途径与挑战

刘肖凡¹, 吴晔^{2,3}, 许小可⁴

(1 香港城市大学媒体与传播系互联网挖掘实验室, 香港

2 北京师范大学计算传播学研究中心, 珠海

3 北京师范大学新闻传播学院, 北京

4 大连民族大学信息与通信工程学院, 大连)

摘要: 融媒体改革是传统媒体应对传媒经济平台化和传媒平台智能化浪潮的一项重要措施。传统机构在融媒体环境中利用数据驱动的计算传播学方法“重新”认识受众、量化传播效果, 既是时代的必然, 又是时代的需要。本文论述融媒体环境中受众计算的“为什么”和“怎么做”问题, 将受众分析分解为三个纵向数据维度(文本、网络和时空数据)和两个横向方法维度(统计分析和机器学习)。最后, 我们着重讨论融媒体受众计算所面临的困难和挑战, 并提出可能的解决方案。

关键词: 融媒体; 平台经济; 受众分析; 传播效果; 计算传播

中图分类号: G206 文献标识码: A

Computational Audience Analysis in Media Integration: Research Directions and Challenges

LIU Xiao Fan¹, WU Ye^{2,3}, Xu Xiao-Ke⁴

(1 Web Mining Laboratory, Department of Media and Communication, City University of Hong Kong, Hong Kong SAR

2 Computational Communication Research Center, Beijing Normal University, Zhuhai

3 School of Journalism and Communication, Beijing Normal University, Beijing

4 College of Information and Communication Engineering, Dalian Minzu University, Dalian)

Abstract: The media integration reform is an essential measure for traditional media to respond to the media's economic platformization and media platform intelligentization. "Re-understanding" the audience and "re-quantifying" communication effects in the new media environment using data-driven computational communication methods are the necessity of the times and the needs of the times. This article discusses the "why" and "how" issues of audience calculation in the media integration reform and decomposes audience analysis into three vertical data dimensions (text, network, and spatiotemporal data) and two horizontal method dimensions (statistical analysis and machine learning). Finally, we focus on discussing the difficulties and challenges in computational audience analysis and propose possible solutions.

Key words : media integration; platform economy; audience analysis; media effect; computational communication

1 平台经济视角下的融媒体改革¹

尽管融媒体改革近年来才被业界和学界频频提起,但实际上传媒行业与互联网的融合(convergence)早在1978年就由麻省理工大学Nicholas Negroponte(Media Lab的共同创始人)提出。在互联网传媒平台尚未兴起前,传统媒体还提出了泛媒体、全媒体等概念,认为数字媒介会像过往所有新媒介一样与自己渐变共融^[1-5]。然而,在过去的几十年间传统媒体并没有有效将互联网融入自己,反而是互联网平台出人意料地冲击了整个传统媒体行业^[6,7]。如果从经济学的角度来审视这一过程,这个结果虽然在意料之外,但是也在情理之中。

二十一世纪以来移动互联网基础设施的完善催生了一种新的经济形式——互联网平台经济。平台经济指的是由互联网平台主导的、需求侧和供给侧自由进入、平台提供信息撮合机制的一种经济业态^[8]。互联网平台降低了信息流通的成本,间接使得商品流通更加快速、各种经济活动更加便利。尤其近年来人工智能迅速发展后,深度学习等算法赋能互联网平台,进一步缩短了用户寻找商品的时间,促进了商品流通。平台经济已经颠覆了多个行业的传统商业模式,例如电子商务。

无论是国外还是国内,传媒平台都是最早出现的互联网平台之一。从较为简单的新闻聚合器(news aggregator)、包含社交功能的社交媒体(博客、微博、微信),到近年来迅速崛起的智能推荐平台(今日头条、抖音),互联网传媒平台的形式几经变革。从行业功能角度来看,平台在媒体采、编、发的业务流程中担当了发行功能,负责撮合媒体内容和受众。互联网传媒平台的发展壮大对传统媒体行业的冲击有两个维度。首先,由于受众消费媒体的时间是有限的、受众注意力是一个零和游戏^[9],互联网传媒平台对受众的消费时间的挤占就意味着传统渠道的市场份额下降。其次,内容采编与发行功能的独立催生了内容采编的草根化和职业化,这导致传统机构必须与用户生成内容(UGC)和由自身分离出去的优质内容在第三方渠道中直接竞争。更甚者,智能化算法让互联网平台变相拥有了对信息触达的控制权。失去了赖以生存的垄断渠道、面对大量对

手的高强度竞争,传统媒体的受众触达能力必然断崖式下降。

在此背景下,传统媒体进行融媒体改革是必要和紧迫的。习总书记在2018年8月和2019年2月发表两次重要讲话,旨在推动传统媒体以互联网思维为基础的媒体改革。在这次自下而上的改革中,传统机构采取了一系列举措来应对渠道冲击和内容竞争。改革的一个重要的措施是渠道融合。随着传统受众迁徙到互联网平台,传统媒体也必须向互联网阵地转移,否则只能面对受众越来越少这一残酷的现实。“三微一端”,即微信、微博、微视频和客户端,是传统机构在大屏、纸张之外的小屏幕上拓展传播渠道的尝试。但是从渠道性质上来看,三微一端又分为两个类别。微信、微博和微视频账号代表的是传统机构进驻互联网平台参与和其他内容生产者的同场竞技。而客户端,也包括多个传统机构联合打造的新闻聚合平台,则是从更低的层面与其他互联网平台展开正面交战。

融媒体改革的另一项重要措施是内容采编流程的改革。媒体机构的组织形式多年来一直在变革,例如地方报业的整合和集团化过程^[10,11]。在融媒体改革中,中央和地方的广电和报业部门进一步以集团形式整合记者资源,在内部打造“中央厨房”式的大编辑室,将内容进行多角度编辑发放在多个平台上,最终形成“一次采集、多种生成、多元发布”的内部流程。这种改革的目的在于提高内容生产的质量和效率,以应对在第三方平台上的受众争夺战。2018年3月,原中央电视台、中央人民广播电台和中国国际广播电台整合成立中央广播电视总台,是传统机构集团化的历史里程碑。传统渠道强强联合既是融媒体背景下有效管理的需要,也是应对传统媒体渠道受众触达能力下降的重要举措。

2 融媒体环境下受众分析的重要性

在融媒体的环境下,面对互联网平台对原有渠道的冲击,传统机构除了迎合互联网受众的口味而改变风格外,还要继续担负起传播正能量的责任。想要科学地评估多种改革举措是否有效、投入的成本是否能够形成收益^[12],还得回到传播学理论的根本:受众分析上来。我们认为融媒体改革中,传统媒体需要加强融媒体中、尤其是互联网渠道中受众的原因有以下三点。

(1) 新渠道中的传播效果量化需要受众分析。

基金项目:国家社科基金重大项目“大数据时代计算传播学的理论、方法与应用研究”(项目号:19ZDA324)

作者简介:刘肖凡(1985-),香港城市大学媒体与传播系助理教授,xf.liu@cityu.edu.hk。

衡量三微一端的建设成效需要明确的量化标准，科学的衡量方法是受众的触达量和他们对媒体内容的反馈。

传统媒体衡量受众触达的方法是订阅量或收视率。在互联网平台中则通常以日/月活跃用户、点击量，甚至更简单的粉丝量、关注量作为指标。相比传统传播渠道，互联网数据的可靠性不高，流量造假、评论造假的情况非常常见^[13]。此外，受众的触达也并不等于正面的传播效果。要评估传统机构在互联网平台的融入程度和新媒体平台的建设成效，仅仅以点击量等简单指标来衡量是远远不够的。更真实的传播效果分析需要基于用户的实际行为，即通过文字、符号等媒介表达出来的真实态度，作为标准来全方位地把握。

(2) *内容生产需要基于对受众的理解*。传统机构的内容生产曾经以多样性为目标。地方广电和报业集团由于市场垄断，只需通过有限几个频道和几种报纸类型即可瓜分受众群体。然而互联网平台打破了空间的分割，导致如果各地媒体依旧生产全覆盖式的内容，就会形成同质化的竞争，并且被受众群体抛弃。为避免这种情况，媒体在生产内容前就需要明确自己的定位、了解自己的受众群体。

对受众的理解除了有助于内容生产外，也有助于指引业务拓展的方向。如今，不少传统机构从传播渠道中拓展出满足受众需求的多种服务，提供如政务、生活、教育培训等。应该拓展哪些业务、向哪些受众提供服务，这些问题都需要坚实的受众理解作为基础。

(3) *网络生态治理的需要*。国家互联网信息办公室《网络信息内容生态治理规定》要求建立健全的网络综合治理体系、营造风清气正的网络空间、建设良好的网络生态^[14]。传统媒体在融入网络生态的同时，也需要起到引导生态的作用。但是在能够引导生态之前，认清网络生态尤为重要。中宣部和广电总局编制的《县级融媒体中心建设规范》要求省级技术平台采集和汇聚受众数据、对受众进行实时画像，并对传播方式、范围、效果展开分析。这些举措都体现了受众分析的紧迫性。

3 计算传播学中的受众计算方法

受众分析的主要算法主要来自于以数据挖掘技术为核心建立起来的计算传播学领域。计算传播学研究通过数据的搜集、清洗和分析，以量化的方式考察受众的行为特征、兴趣爱好、情绪表达，以及媒

体的传播效果^[15]。使用计算方法进行融媒体环境下的受众分析，是互联网时代的必然。

互联网数据的获取主要有两个渠道：网站爬虫和数据管道（data hose）。网站爬虫是以用户身份访问互联网平台并存储数据技术。数据管道则基于分析者与互联网平台间签订的访问协议，从后台快速下载数据。互联网平台上可以获得的用户数据包括用户简介（profile）、社交关系、活跃时间、浏览记录、发布内容、点赞、评论和转发等行为。

我们把基于互联网数据的计算思维分为三个纵向数据维度和两个横向方法维度。数据维度包括文本数据、网络数据和时空数据^[16]；方法维度包括统计分析和机器学习算法。需要强调的是，我们并不认为计算方法与人工智能是等同的概念，而是一种行为数据驱动、量化指标测量的分析思维。接下来将从三个纵向维度梳理目前已有的分析方法和典型的任务目标。

(1) *文本类型数据*。文本数据主要包括用户在简介、发布和评论中留下的文本、图片、符号等多维度信息。这类数据的特点是非结构化、分析流程复杂，但是包含的信息非常丰富。

文本分析可以完成大部分受众分析和传播效果分析的需求，包括用户画像、舆论议程分析、情感和态度测量等。用户使用的文字、符号、表情等信息可以直接反映他的教育水平、文化背景和兴趣爱好等人口特征。通过对媒体内容进行主题分类，可以把握当前舆论场中的议程设置。受众发布的评论内容则可以用来把握受众对议程的具体态度，了解受众真实情感。

文本分析的方法主要有字典法和机器学习方法两种类别^[17]。字典法由人工将字词按照属性做分类或打标签，例如“太好了”表达高兴的感情，然后对文本中出现的各种情感的关键词计数，以判断文本的类别。机器学习方法的不同之处在于人工标记的对象并不是字词，而是长文本。机器学习算法则从每类文本中找出相似之处和类别之间的不同之处，利用找到的规律（又称为模型）对其他未标记的文本进行分类。

(2) *网络类型数据*。网络类型数据，又称为关系型数据，包含一个群体内个体间的关联信息。在融媒体环境中，个体间的关联包括社交、关注、二级传播等。由于互联网上用户群体数量巨大，网络类型数据的特点是结构复杂、信息隐藏较深、处理复杂度较高。

社交网络分析和二级传播网络分析的目的略有不同。用户间的社交网络反应了用户在该平台上表现出的社会属性，例如意见领袖、信息中介等社会角色。社交网络是受众分析的重要组成部分。消息传播网络则反映了二级传播行为^[18]。对消息传播网络的分析目的主要是检验媒体内容或形式在受众中的传播深度、广度等传播能力指标。这些信息均有

助于量化媒体内容的传播效果。

网络分析的方法也分为统计方法和机器学习方法。统计方法采用传统社交网络分析（SNA）和网络科学分析方法抽取网络中的个体（节点）和关系（边），以及自中心网络的统计量，例如节点的度、边的中心性等。然后通过与随机网络模型（零模型或 ERGM）进行对比，得出统计意义上的显著性结论^[19]。这些统计量会作为实际任务中对个体或关系的衡量指标，例如符合某种结构规律的个体是意见领袖。针对网络数据的机器学习算法则通过人工标签信息找出不同类别网络在结构特征上的规律，再通过这些规律为未标注的网络进行分类，完成预测任务。

*(3) 时空类型数据。*除了文本和社交关系数据之外，同样具有重要价值的还有日志（log）数据。日志数据以条目的形式顺序记录用户行为的时间和地点。因此，时空数据的本质都是高度结构化和序列化的。

时空分析的目标是从序列中的频率、顺序等信息中找到用户行为的规律^[20]，例如稳定性、周期性、关联性等。这些规律可以进一步反映受众的人口信息和行为可预测性，为下游任务提供重要的参考。

时空分析的主要方法同样分为统计方法和机器学习方法两类。统计方法依赖传统指标寻找序列的规律，例如用熵来反映可预测性、用动态时间规整度（Dynamic Time Warping）来判断数据周期性等。机器学习方法则用算法来找到时空数据的内部规律和与外部事件间的联系，预测在特定时间地点可能发生的事件类型。

4 融媒体环境下受众计算的挑战

虽然计算传播方法已经发展出了一套完整的理论和工具可用于受众分析，但是融媒体环境中的受众分析也有其特殊之处。这一节中我们着重讨论融媒体受众分析的几个显著挑战和可能的解决方案。有的解决方案并不是简单依靠能力提升就能实现，而需要全社会的共同努力来推动。

*(1) 互联网平台受众的根本性不同。*传统媒体所面对的受众是较为明确的。报刊、电视都有明确的订阅人，他们的身份和地址等信息为用户画像提供了丰富的素材。然而，传统渠道里的客户流失意味着过往用户画像的失效。虽然传统机构重新在互联网传媒平台上获得了一定受众群体，但是新受众与旧受众仍有根本性的不同。这种不同首先体现在受众身份上。互联网平台是扁平的，一条媒体内容可以顺时触达所有受众。这种扁平性打破了传统条块化媒体受众的人口同质性。与传统媒体受众不同，互联网平台上的受众从经济条件、教育水平到文化

差异，都体现了极强的异质性。

其次，受众的内容消费方式也产生了变化。传统受众倾向于直接从媒体消费信息或进行短距离的人际传播。但是互联网平台的社交性和智能性改变了这种消费习惯。由于社交网络的存在，受众的信息消费受第三人影响非常大。考虑到不同的平台的特性，受众有可能受到强关系（微信）、弱关系（微博），甚至多层以外社会关系的影响消费内容的选择。智能化算法更是可以根据受众兴趣强化内容推荐，造成了第一人效应的强化。当然，各大平台也都在调整算法，以尽力避免“过滤气泡”或“信息茧房”的出现^[21]。

总之，互联网传媒平台中的受众不再是身份明确的单一个体或家庭，而变成了一个高维的兴趣向量及其自我网络（ego network）的组合。受众的信息来源渠道也已经从单一媒体转变为了社交推荐和算法推荐。理解这些变化对传统机构的采编人员而言是进行下一步工作的前提。

*(2) 跨平台受众分析的困难。*融媒体环境中传统机构的内容会在多个渠道同时发布。这也就意味着想要分析内容的传播效果，必须综合多渠道的受众反馈。然而，由于同一个互联网用户也会在多个平台消费内容，所以多个平台的传播效果不能简单累加。严谨的传播效果分析需要首先识别出不同平台上的重复用户并进行受众整合。

跨平台的用户身份匹配是一个具有挑战性的科学问题。互联网用户具有匿名性，在互联网上通常不会主动披露自己的身份。虽然互联网用户实名化是一个趋势，但是互联网平台也不会轻易披露用户真实身份供媒体机构分析。可行的解决方案包括匹配用户惯用 id、简介中披露的个人信息以及用户间的社交关系。例如，如果能够通过用户 id 或简介定位几个明确的跨平台用户，并且获得这些用户在不同平台上的社交关系，那么就有可能通过对比社交网络结构的办法识别出其他有可能的跨平台用户^[22]。

跨平台受众识别对技术能力要求很高，给融媒体改革的受众分析带来了巨大的挑战。传统媒体机构不一定能够在短期内掌握所需的数据和技术资源。所以我们建议互联网平台能够一定程度地开放数据，协助传统媒体和其他入驻媒体找准受众定位。

*(3) 互联网受众样本缺乏代表性。*当传统媒体通过网络空间与受众互动时，通常需要想象自己的受众，并决定说什么和做什么^[23]。虽然通过计算方法能够使假想的受众浮出水面，但是也要强调，能够被计算的用户只是互联网平台受众的一个带有偏差的样本子集。

如果将互联网上的用户根据内容生产量和消费

时长大致分为活跃/不活跃的传播者和受众四个类别，那么能够通过数据搜集到的用户行为大多仅仅描述了活跃的受众^[23]。而沉默的受众所具有的行为特征以及他们的态度和情绪则很难从零星的数据中总结出来。问卷调查是一种可能的解决方案。但是相较于传统的电话和上门调查方法，由于互联网环境的匿名性和用户接受调查时的随意性，使用网上问卷往往也存在系统性的采样偏差^[24]。

受众分析的样本偏差导致的后果可能是严重的。许多传统机构认为融媒体的内容生产需要迎合受众口味、顺应受众兴趣。通过反向工程平台的推荐算法，一些新闻机构会在媒介形式、新闻标题、内容情感上做出调整，以活泼抓眼的形式吸引互联网平台受众^[25, 26]。这些方式从表面上看也许受到部分用户的肯定，但是对沉默的受众而言效果并不可考。例如，根据企鹅智酷2017年调查，超过50%的网民认为调查、科普、观察性才是被认可的新闻资讯。传统机构在全民媒体时代维护“公知”形象，也符合受众对新闻正义和公众教育的期待。这样的用户需求就很难从活跃用户的行为中分析出来。到底是选择迎合用户口味，还是坚持生产严肃内容，抑或是在两个方向之间做出妥协，是摆在传统机构面前的一个开放问题。需要全面搜集并科学评估受众的反馈，才能找到内容生产的平衡点。

(4) 互联网数据获取困难。全世界对数据产权认知变革的里程碑是2018年5月开始实施的欧洲《通用数据保护条例》(General Data Protection Regulation, 简称GDPR)。它也被称为“史上最严格的隐私保护法”。它针对过去二十年来互联网企业对用户数据的垄断性掌控和滥用问题，将用户的数字数据当做隐私来保护，规定互联网平台保存和分析用户数据前必须征得用户的明确同意。

互联网平台深知用户数据的商业价值。各平台对数据爬取的政策从大约十年前开始逐步收紧，并在GDPR法案实施后变得更加困难。如今，一个普通科研团队或传统媒体的研究机构已经很难利用爬虫的方式从成熟的互联网平台获得大量的用户数据。

数据获取问题的可能解决方式是业界面向学界开放数据。虽然学术研究不一定能够为业界带来直接的经济利益，但是基础研究的进步还是能够为整个产业带来长远的收益。

(5) 互联网数据质量问题。互联网数据具有量大和信息稀疏的特点，数据采集的过程也难免出现疏漏错误。因此，采集到的数据通常存在缺失值、重复值和异常值(过高、过低或人为造假)的情况。这些问题的处理费时费力，其重要性常常被忽视。但是计算机行业有俗语“垃圾进，垃圾出”，即没有

经过精细清洗的数据往往不能产出任何有效的分析结果。

数据质量问题除了技术原因之外，还有一个更加基础的问题，即网络用户身份的不可知性。除了前面提到的样本偏差问题之外，数据中还掺杂了社交机器人产生的假数据。“数据公关”是互联网平台上的一个重要产业链。需求方通过雇佣水军或使用个社交机器人的方式，制造虚假点赞数据和评论内容，引导舆论方向。

水军和社交机器人识别与受众计算类似，也使用文本数据、社交网络和时空数据。利用异常检测、特征识别等方法，水军识别已经形成一套成熟的技术堆栈^[27]。但是需要注意的是，在水军识别技术提升的同时，虚假内容的生成也在“道高一尺，魔高一丈”地进步。这两者之间军备竞赛式地竞争将是互联网受众分析的长期主题。

(6) 计算方法还有很大的提升空间。文本、网络和时空数据的分析技术都还有巨大的发展空间。文本分析中字典的构建、机器学习中的数据标记都是成本极高的工作。由于对字词或文本的标记工作需要耗费大量成本，公开可获得的文本数据标记尚且有限。目前的标签数据集大多是公开比赛或学术研究领域通用的数据集，在其他应用场景中并没有很强的普适性。因此，在其基础上建立起来的上层建筑缺乏泛化(generalization)能力。

当今机器学习算法常常被应用领域所诟病的问题是其生成的模型缺乏可解释性。也就是说，用来描述用户行为规律的复杂模型并不能够被人类所完全理解。由此造成的后果是应用领域从业人员对算法的不信任，甚至抛出所谓“算法歧视”的问题。模型可解释性是计算方法发展的难点和重点，但是避免出现“算法歧视”实际上有法可依。在训练算法之前，只需要对样本数据中包含的信息进行事先分析，保证不存在“数据歧视”问题，即可最大程度避免“算法歧视”。

最后，在计算复杂度上，无论是网络数据分析算法还是深度学习算法在面对百万至上亿级受众分析任务时都需要的大量计算资源，使得普通机构无法完成大量受众分析。降低算法复杂性也是计算方法发展的当务之急。

5 结论

本文从平台经济学的角度分析了融媒体产生的背景，提出融媒体面对的冲击来源于传统渠道被侵占，所以不得不在不受控的第三方平台上面对自媒体的直接竞争。融媒体改革的成效，即受众触达量和传播效果评估都需要更准确的受众分析作为前提。传统的调查方法在互联网平台中面临很大困难，而互联网世界中原生的计算方法以及计算传播学领域

的定量分析就成为自然的选择。计算方法的流程包括数据的获取、清洗和分析。针对文本、关系和时空类型数据,受众计算可以采取统计分析和机器学习的办法获得用户行为的规律,以完成用户画像和传播效果量化的任务。计算方法虽然已经表现出了很强的实用性,但是也存在明显的问题。出现问题的一部分原因源于互联网受众的特殊性,另一部分源于从互联网平台获取数据的困难以及数据的不准确性。这些问题的解决任重而道远,需要今后业界和学界以开放的态度共同攻关解决。

传统媒体应当在融媒体改革中融入计算的元素。按照实现方式从易到难,但是预期效果逐渐提高的顺序,具体方法可以分为三种。(1)与拥有计算能力的调查机构合作,或与互联网平台合作,购买受众分析结果;(2)提高行业从业者的计算思维,增加计算能力的训练,在日常工作中将受众反馈作为考核形式;(3)进行全面的数字化转型,以证据作为决策的基础,将信息技术提升至机构中最重要的地位,围绕信息化、数字化重构行业。融媒体的改革方兴未艾,传统媒体与未来世界的对话也才刚刚开始。我们认为,拥抱未来的最好方式,就是成为未来,受众计算将会起到非常关键的作用。

致谢

作者感谢祝建华教授、王丹博士和刘薇薇女士在文章撰写过程中提出的宝贵意见和建议。

参考文献

- [1] Fidler, R. F. (1997). *Mediamorphosis: Understanding new media*. Pine Forge Press.
- [2] 蔡雯,王学文.角度·视野·轨迹——试析有关“媒介融合”的研究[J].国际新闻界,2009(11):87-91.
- [3] 张大伟.数字出版即全媒体出版论——对“数字出版”概念生成语境的一种分析[J].新闻大学,2010(01):113-120.
- [4] 李玮.跨媒体·全媒体·融媒体——媒体融合相关概念变迁与实践演进[J].新闻与写作,2017(06):38-40.
- [5] 黄旦.试说“融媒体”:历史的视角[J].新闻记者,2019(03):20-26.
- [6] 胡泳,陈秋心.中国网络传播研究:萌芽、勃兴与再出发[J].新闻战线,2019(03):53-59.
- [7] Barker, A. Digital ad market set to eclipse traditional media for first time. *Financial Times*, June 23 2020.
- [8] Rochet, J. C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European economic association*, 1(4), 990-1029.
- [9] Zhu, J. H. (1992). Issue competition and attention distraction: A zero-sum theory of agenda-setting. *Journalism Quarterly*, 69(4), 825-836.
- [10] 李良荣,林晖.垄断·自由竞争·垄断竞争——当代中国新闻媒介集团化趋向透析[J].新闻大学,1999(02):3-5.
- [11] Klinenberg, E. (2005). Convergence: News production in a digital age. *The Annals of the American Academy of Political and Social Science*, 597(1), 48-64.
- [12] 黄楚新,彭韵佳.2017年中国媒体融合发展报告[J].现代传播(中国传媒大学学报),2018,40(04):9-15.
- [13] Arthur, C (August 2, 2013). "How low-paid workers at 'click farms' create appearance of online popularity". *the Guardian*.
- [14] 国家互联网信息办公室.《网络信息内容生态治理规定》,2019年12月15日.
- [15] 张伦,王成军,许小可. (2018). *计算传播学导论*. 北京师范大学出版社.
- [16] 祝建华, 黄煜, & 张昕之. (2018). 对谈计算传播学: 起源, 理论, 方法与研究问题. *传播与社会学刊*, (44), 1-24.
- [17] Jurafsky, D., & Martin, J. H. (2018). Speech and language processing (draft). *Chapter A: Hidden Markov Models (Draft of September 11, 2018)*. Retrieved March, 19, 2019.
- [18] Katz, E., & Lazarsfeld, P. F. (1955). Personal influence: the part played by people in the flow of mass communications.
- [19] 汪小帆, 李翔, & 陈关荣. (2012). *网络科学导论*. 高等教育出版社.
- [20] Zhu, J. J., Chen, H., Peng, T. Q., Liu, X. F., & Dai, H. (2018). How to measure sessions of mobile phone use? Quantification, evaluation, and applications. *Mobile Media & Communication*, 6(2), 215-232.
- [21] Nechushtai, E. & Lewis, S.C. (2019). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, 90, 298-307.

- [22] Narayanan, A., & Shmatikov, V. (2009). *De-anonymizing Social Networks*. The 30th IEEE Symposium on Security and Privacy, 17-20 May 2009.
- [23] 李晓静,付思琪.智能时代传播学受众与效果研究:理论、方法与展望——与香港城市大学祝建华教授、斯坦福大学杰佛瑞·汉考克教授对话[J].国际新闻界,2020,42(03):108-128.
- [24] Ward, M. K., Meade, A. W., Allred, C. M., Pappalardo, G., & Stoughton, J. W. (2017). Careless response and attrition as sources of bias in online survey assessments of personality traits and performance. *Computers in Human Behavior*, 76, 417-430.
- [25] Tandoc Jr, E. C., & Maitra, J. (2018). News organizations' use of Native Videos on Facebook: Tweaking the journalistic field one algorithm change at a time. *New Media & Society*, 20(5), 1679-1696.
- [26] Lischka, J. A. (2018). Logics in social media news making: How social media editors marry the Facebook logic with journalistic standards. *Journalism*, 1464884918788472.
- [27] 莫倩, 杨珂. 网络水军识别研究[J].软件学报,2014, 25(7): 1505-1526.