

浅析国内网络虚假信息研究方法

张恒瑞

(郑州航空工业管理学院信息管理学院, 河南 郑州 450046)

【摘要】 [目的/意义] 对国内虚假信息的相关研究进行分类, 归纳分析相关研究方法, 有助于丰富虚假信息相关研究, 推动虚假信息治理。[方法/过程] 本文从中国知网数据库中, 收集整理了107篇虚假信息相关文献。将这些文献分为虚假信息检测、虚假信息传播、虚假信息治理三部分, 探析了虚假信息这三部分在研究方法的特征。[结果/结论] 研究发现, 模型展示法和理性思辨法得到了较多的使用, 图书情报学和传播学领域的基础理论的出现频次最多, 各个部分的研究方法与基础理论占比具有差异。

【关键词】 虚假信息; 研究方法; 基础理论; 图书情报; 国内

【DOI】 10.12294/j.issn.1673-0992.2022.06.058

一、引言

在线社交网络中, “虚假信息泛滥”现象日益严重, 并逐渐引起了人们的广泛关注。2020年新冠疫情期间, 由于人们对网络信息交流的依赖, 大量疫情相关的虚假信息充斥在网络中, 加剧了人们的焦虑与恐慌, 给疫情防控和社会治理带来巨大危害。我国虚假信息相关研究显著增多, 其研究内容主要涉及图书情报学、心理学、传播学、计算机科学、社会学等多个学科领域, 体现出显著的交叉性特点。

陈燕方等^[1]从在线社会网络谣言的基本概念和特征研究出发, 分别基于目标、对象和时间三个属性, 分析了在线社会网络谣言检测研究基本问题的形式化定义。王晰巍等^[2]通过搜集和分析国内外社交媒体网络谣言的相关文献, 对国内外网络谣言研究的起源和发展、国内外研究热点及未来研究趋势进行对比和可视化分析。李丹丹等^[3]以复杂社会网络中谣言传播模型为对象, 系统论述了谣言传播模型的发展历程。闵庆飞等^[4]对谣言相关研究进行了分析综述, 提出了谣言研究的三个发展阶段, 以及计算机媒介与社会化媒体情境下谣言研究的重大意义与未来方向。本文将重点对虚假信息研究方法进行总结分析, 以期发现学科交叉特点, 为后续研究提供思路与参考。

二、我国虚假信息研究现状

以“虚假信息”或“谣言”或“伪信息”或“虚假新闻”为关键词, 在中国知网中进行检索, 时间范围设定为2015年3月至2021年3月, 文献来源设定为北大核心期刊和CSSCI期刊, 共获得相关核心论文235篇, 并筛选掉有关经济学造假和会计学造假的相关论文, 最终得到文献107篇。

纵观虚假信息相关文献, 其研究主要集中在虚假信息检测、虚假信息传播和虚假信息治理三个方面。①在虚假信息检测研究方面, 学者多以微博为主要研究对象,

重点关注虚假信息源头和虚假信息早期形成时的特点, 探索快速有效的虚假信息检测和甄别方法。②在虚假信息传播的相关研究中, 大量的传播学理论被引用, 许多学者从虚假信息的传播特征出发, 通过构建研究框架, 对虚假信息的传播和治理进行了深入探究。③在虚假信息治理的相关研究中, 公共事件辟谣和用户行为得到了较多关注, 通过辟谣事件的研究探讨了虚假信息的治理方法、体系及政府治理政策。

值得注意的是, 虚假信息的检测、传播和治理三方面的研究并不是完全孤立的, 其内容是有交叉的, 如虚假信息治理的相关研究会涉及虚假信息传播, 但具体研究中会有所侧重, 具有较高的辨识度。从学科领域来看, 虚假信息研究的相关文献主要分布在图书情报学、传播学、计算机科学、心理学和社会学等多个学科领域, 其中基础理论大多是以图书情报学和传播学为主, 计算机科学领域主要提供有力的技术支持。

综上, 根据研究内容, 本文将所收集的107篇文献按照虚假信息检测、虚假信息传播和虚假信息治理分为三组, 在文献分组时, 存在研究内容交叉的现象, 做重复归类处理。

三、我国虚假信息研究的方法论分析

(一) 分类统计

为方便对虚假信息研究的相关方法进行统计分析, 本文选择以武汉大学司莉等撰写的《近五年我国图书情报学研究方法应用与演化的实证研究——基于5种期刊的统计分析》文章^[5]为参考, 选取文章里图书情报研究方法分类统计表(如表1所示)为分类依据。

本文以此为基础对三组虚假信息相关文献的研究方法进行归类与统计, 将虚假信息研究涉及到的方法进行归纳。依据方法的固有特性, 可将这12种方法分为定量研究方法和定性研究方法。定量研究方法主要包括实验、调查、仿真、回归统计分析等, 重视的是假设与理

表 1 研究方法分类统计表

研究方法	释义	举例
模型展示法	构建模型、平台、框架、本体等	本体构建、演化分析、模型分析等
仿真研究法	利用模型复现实际系统中发生的本质过程，并通过对系统模型的实验来进行研究	物理仿真、计算机仿真（数学仿真）、半实物仿真
数学方法	运用数学的算法、公式进行分析	因果链算法、马尔科夫过程、权重计算等
理性思辨法	在已有知识基础上采用概念、判断和推理等形式进行理性思考而形成思想观点	概念辨析、特征分析、对策研究、功能分析、VIKOR（多属性决策法）、分类方法等
统计分析法	运用了统计学的方法进行分析	SPSS 软件、层次分析法、回归分析、PageRank 算法等
关联分析法	运用工具分析事物间的潜在关联	可视化分析、知识图谱、社会网络分析法、鱼骨图等
案例分析法	结合实际例子对相关理论进行研究	
数据挖掘法	对数据或文本进行挖掘	专利挖掘法、日志挖掘法、知识抽取法
计量分析法	以计量为基础进行分析	文献计量、引文分析、信息计量等
实验观察法	通过进行实验观察获得结论	用户实验法、有声思维法
调查法	按照拟定的题目进行相关调查	访谈、田野调查、德尔菲法、文献调研、网络调研
实证分析法	通过样本观测数据的分析验证得出结论	可行性验证等

论验证，强调变量之间的因果关系，因此其分析过程、结论多以统计分析或图表为主。定性研究方法则重视意义解释与理论建构，因此其推理过程、结论多以文字描述为主，更为广义的定性研究还将思辨性研究、经验描述性研究纳入其中。

（二）数据分析

在虚假信息的检测、传播和治理三方面的研究中，每部分涉及到的研究方法并不相同，且各个部分研究方法占比也有较大差异，通过统计每部分研究方法的种类和数量，绘制了图 1 虚假信息研究方法的分布统计表和图 2 虚假信息研究方法频次表，用于研究方法的纵向分布分析。

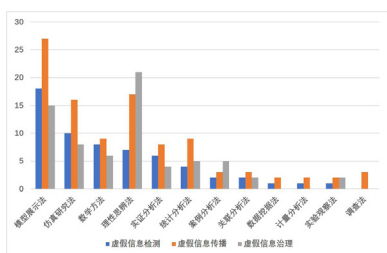


图 1 虚假信息研究方法分布统计表

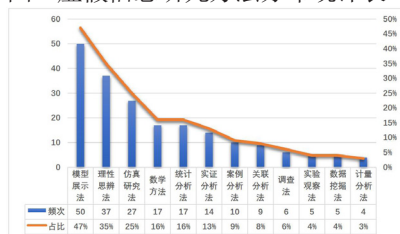


图 2 虚假信息研究方法频次表

1. 共性分析

①模型展示法是使用次数最多的研究方法，在虚假信息全部文献中使用频次和占比位列第一。如刁海伦^[6]采用无尺度网络的算法构建微博网络，对虚假信息和检测同时进行建模，建立微博虚假信息的集中甄别仿真模型 MCFDM，以此研究最佳的虚假信息检测甄别方案。张

卫东^[7]在已有的传统信息博弈模型基础上，提出了“完全信息博弈模型”的概念，在社交媒体用户之间建立博弈模型，分析用户的选择策略。这表明，相关文章在研究中比较侧重于虚假信息这一抽象事物运动规律的探究，希望通过搭建系统化的模型来模拟和仿真虚假信息的“形成”或“传播”机制。

②在虚假信息相关研究中，模型展示法很少单独使用，多与仿真法和数学方法结合使用，因为在构建模型时往往需要运用到大量的数理计算和仿真实验来进行结果验证。如倪培昆^[8]提出了两阶段贪婪算法 (TSGA)，通过实际在线社交网络数据对模型和算法的有效性进行了分析，实验表明本文提出的算法比现有算法更加有效。胡欢^[9]构建了政府、网络运营商和造谣者三方演化博弈模型，然后通过系统动力学软件进行仿真模拟，最后依据模拟结果提出网络谣言管控的相关建议。故仿真法和数学方法也有着很高的使用频次和占比。

③理性思辨法单独使用频次最多，在虚假信息全部文献中使用频次和占比位列第二，且在三个部分中均多次被采用。与模型展示法不同的是，采用理性思辨法的研究往往只是对相关问题的理论和概念进行研究和讨论，大多单独的采用理性思辨法。如韩娜^[10]以行动者网络理论作为一种社会建构和政府治理的研究理论，为网络虚假信息治理提供了理论框架，通过分析各个主体之间的连接点和行为困境，建构克服行动障碍的路径，以形成稳定的虚假信息治理行动者网络联盟，以期推进网络社会的治理。何勇^[11]结合对突发公共卫生事件网络谣言的理论分析，提出了应该从供给侧体系建设、德治与法治相结合、网络社区治理多方面建构符合社会治理现代化要求的谣言治理体系理论。

2. 差异性分析

虚假信息检测、传播和治理三个方向在研究方法使用上各有侧重，不尽相同。

①在虚假信息检测研究中，模型展示法和仿真研究法使用最多，原因在于“虚假信息检测”研究中引入了

较多的人工智能和神经网络算法,且减少了对单纯理论的研究和讨论。如米源^[12]提出在目前网络谣言鉴别方法中,通过采用半监督学习的图卷积神经网络,可有效利用无标注数据,本方法可有效减少人工标注代价,鉴别社交媒体和网络新闻中的谣言。尹鹏博^[13]使用卷积—长短期记忆网络模型进行谣言早期检测,减少了等待产生评论转发信息的时间,实现了微博谣言的实时早期检测。谣言模型实验检测结果表明,对于虚假信息检测而言,更多的定量研究有助于验证和发现新的相关检测方法,完善其研究成果。

②“虚假信息传播”相关研究数量最多,涉及的研究方法种类也是最多,说明这一研究方向受关注度更高。除模型展示法运用频次最多以外,“虚假信息传播”的相关文章还使用了较多的理性思辨法,这是主要由于涉及到了大量的传播学和其他学科理论。如王芳^[14]以百度新闻帖子数量的变化为依据,采用模型展示法,构建了网络谣言传播规律的宏观模型,并进行仿真与实证分析。林华^[15]结合网络谣言传播的理论与实践,以类型化为视角分析不同网络谣言传播的逻辑异同,包括普通网络谣言的传播逻辑理论和焦点网络谣言的传播逻辑理论。统计表明这一部分研究更为繁多复杂,既有较多的定量研究,也有大量的定性研究,是虚假信息研究的核心部分。

③“虚假信息治理”的相关文章中则是理性思辨法使用的最多,且使用频次和占比均高于其他两部分。在虚假信息的治理方面,大多是有关理论概念的研究和讨论,内容中也的确有较多的文章研究和讨论了相关信息治理政策的制定和实施。如王景侠^[16]在分析公共图书馆开展公众媒介信息素养教育的重要性和主要优势的基础上,重点探讨了构建科学高效的媒介信息素养教育体系及其主要途径,以提升我国图书馆界在应对信息治理和重大突发公共卫生事件中的影响力。王灿发^[17]提出民众传播“谣言”是在真相信息传播缺失下的话语表达,与控制信息的一方进行博弈,以激发真相信息的释放与传播,辟谣的有效手段应是秉持开放的心态,利用可以利用的一切传播渠道及时公开真相信息。结果表明,“虚假信息治理”的相关研究主要以定性研究为主,但定量研究方法的使用仍保持一定占比。

四、讨论与总结

在国内虚假信息相关研究中,2016年之后文献数量快速增长,尤其在新冠疫情爆发后,相关文献数量迎来了指数爆发型增长。从研究方向上来看,主要分为“虚假信息的检测”、“虚假信息的传播”、“虚假信息的治理”三个方面,且内容上互有交叉。

在互联网迅速发展的环境下,网络虚假信息的泛滥愈发严重,增加了人们的信息获取成本,扰乱了新时代信息生活的稳定秩序。为有效的对虚假信息进行研究,我们应不断深化和创新虚假信息的相关研究方法,跨学

科、多领域的发展和更新虚假信息的相关基础理论,结合国内外虚假信息的研究进展及未来的研究趋势,推动虚假信息研究的创新发展与优化。

参考文献:

- [1] 陈燕方,李志宇,梁循,齐金山.在线社会网络谣言检测综述[J].计算机学报,2018,41(07):1648-1677.
- [2] 王晰巍,李文乔,韦雅楠,张柳.社交媒体环境下网络谣言国内外研究动态及趋势[J].情报资料工作,2020,41(02):39-46.
- [3] 李丹丹,马静.复杂社会网络上的谣言传播模型研究综述[J].情报理论与实践,2016,39(12):130-134.
- [4] 闵庆飞,刘晓丹.谣言研究综述:基于媒介演变的视角[J].情报杂志,2015,34(04):104-109.
- [5] 司莉,陈金铭,马天怡,等.近五年我国图书情报学研究方法应用与演化的实证研究——基于5种期刊的统计分析[J].图书馆,2019(6):15-21.
- [6] 刁海伦,王树义,王楠.基于多主体的微博网络虚假信息的集中甄别方法研究[J].情报科学,2016,34(02):37-44.
- [7] 张卫东,李松涛,梁恩平.基于完全信息博弈模型的社交媒体用户跟随行为研究[J].情报科学,2019,37(08):114-119.
- [8] 倪培昆,朱建明,王国庆.在线社交网络虚假信息交互量最小化的边阻断策略研究[J/OL].中国管理科学:1-10[2021-09-23].
- [9] 胡欢,郭晓剑,梁雁茹.基于前景理论的重大疫情网络谣言管控三方演化博弈分析[J].情报科学,2021,39(07):45-53.
- [10] 韩娜.行动者网络视角下虚假信息治理的困境与路径[J].领导科学,2020(22):31-34.
- [11] 何勇,杨映瑜.突发公共卫生事件网络谣言的特征及治理[J].现代传播(中国传媒大学学报),2020,42(06):130-135.
- [12] 米源,唐恒亮.基于图卷积网络的谣言鉴别研究[J].计算机工程与应用,2021,57(13):161-167.
- [13] 尹鹏博,潘伟民,彭成,张海军.基于用户特征分析的微博谣言早期检测研究[J].情报杂志,2020,39(07):81-86.
- [14] 王芳,连芷莹.公共危机中谣言真实度计算及其与正面信息的交锋研究[J].图书与情报,2020(01):34-50.
- [15] 林华.普通网络谣言与焦点网络谣言的传播逻辑异同[J].情报杂志,2020,39(09):116-120.
- [16] 王景侠.公共图书馆参与信息治理提升公众媒介信息素养的思考[J].数字图书馆论坛,2020(05):30-35.
- [17] 王灿发,邵全红.重大突发卫生疫情谣言的多维认知与治理——以新冠肺炎疫情期间的谣言为例[J].新闻爱好者,2020(05):14-16.