

DOI:10.12154/j.qbzlgz.2025.01.010

突发事件情境下网络舆情信息摩擦评估与分析*

安 璐^{1,2,3} 成皓飞² (¹武汉大学信息资源研究中心 湖北 430072;
²武汉大学信息管理学院 湖北 430072; ³武汉大学数据智能研究院 湖北 430072)

摘 要: [目的/意义]突发事件情境下,信息搜寻阻碍和信息不对称导致应急信息传播受阻。对突发事件情境下网络舆情信息摩擦评估与分析,进而为突发事件应急管理提供可参考的理论工具,对政府舆情治理能力提升具有重要意义。[方法/过程]基于信息经济学中的信息摩擦概念,从信息搜寻阻碍和信息不对称两个维度构建突发事件情境下网络舆情信息摩擦评估指标体系,利用熵权法对不同主题时段的信息摩擦进行评估和排序。基于Kruskal-Wallis非参数检验对不同生命周期和主题类型的网络舆情信息摩擦进行比较分析。基于公众角度对应急处置效果进行测度,通过相关性分析、构建应急处置效果预测模型和SHAP模型探究信息摩擦对应急处置效果的影响机制。[结果/结论]研究建立了网络舆情信息摩擦的量化标准,分析发现不同生命周期和主题类型信息摩擦存在显著差异,起始期和爆发期信息摩擦较高,调查问责类和情绪表达类主题信息摩擦较高。信息摩擦与应急处置效果之间存在显著负相关关系,信息搜寻阻碍相较于信息不对称对应急处置效果的影响更大。

关键词: 信息摩擦 社交媒体 网络舆情 突发事件 应急处置

Evaluation and Analysis of Online Public Opinion Information Friction in Emergency Situations

An Lu^{1,2,3} Cheng Hao-fei²

(¹Center for Studies of Information Resources, Wuhan University, Hubei, 430072;

²School of Information Management, Wuhan University, Hubei, 430072;

³Institute of Data Intelligence, Wuhan University, Hubei, 430072)

Abstract: [Purpose/significance] In emergency situations, information search obstacles and information asymmetry lead to the obstruction of event information dissemination. Evaluating and analyzing the friction of online public opinion information in emergency situations, and providing reference theoretical tools for emergency management are of great significance for improving the government's public opinion governance capabilities. [Method/process] Based on the concept of information friction in information economics, an evaluation indicator system for online public opinion information friction in the context of public events is constructed from two dimensions: information search obstacles and information asymmetry. Using the entropy weight method to evaluate and rank information friction in different thematic time periods, we compare online public opinion information friction based on the Kruskal Wallis non parametric test for different lifecycles and topic types. The effectiveness of emergency response is measured from a public perspective. The impact mechanism of information friction on emergency response effectiveness is revealed through correlation analysis. [Result/conclusion] This study establishes a quantitative standard for the friction of online public opinion information. Results find significant differences in information friction among different lifecycles and theme types, with higher friction in the initiation and outbreak periods, and higher friction in the topics of investigation accountability and emotional expression. There is a significant negative correlation between information friction and the effectiveness of emergency response. Information search obstacles have a greater impact on the effectiveness of emergency response compared with information asymmetry. The findings provide suggestions for government departments to govern the dissemination of online public opinion information during public events.

Keywords: information friction social media online public opinion emergency emergency response

*本文系国家社会科学基金重大项目“不确定环境下韧性社会智能情报支持与决策研究”(批准号:23&ZD230)的研究成果之一。

1 引言

近年来我国突发事件频发,不仅对社会安全造成危害,也给应急管理部门带来了挑战。《突发事件应急预案管理办法》中提到要充分运用信息化手段推动应急管理创新^[1]。互联网是应急管理信息化的重要阵地。第53次《中国互联网络发展状况统计报告》显示,截至2023年12月,我国网民规模达10.92亿人,互联网普及率达77.5%^[2]。面对日益增长的网民规模与互联网普及率,突发事件发生时的信息传播成本极速下降,政府部门与新闻媒体可以通过社交媒体及时发布应急信息,社会公众可以实时关注突发事件相关进展并发表观点与建议。然而,社交媒体平台上的信息传播同样存在各种阻力。

信息摩擦是信息经济学中的重要概念。在经济学领域,研究者将交易双方在完成预期目标的过程中,因为搜寻摩擦和信息不对称导致的正常支出成本之外消耗的成本称为信息摩擦^[3]。社交媒体网络舆情中同样广泛存在信息摩擦,主要表现在由于信息搜寻阻碍和信息不对称造成的信息传播的额外成本增加。在突发事件的发展过程中,社交媒体将呈现信息井喷态势,海量信息和观点充斥网络空间,其中不可避免的包含低质量信息,增加了用户的信息搜寻、筛选和辨别成本。同时,突发事件发生时网络舆情中也存在着信息不对称现象,即不同用户主体掌握的信息程度不同。这不仅阻碍了信息的正常流动,增加了公众对于信息需求的表达成本,还将影响政府与公众之间良好互信沟通机制的建立,降低公众满意度,甚至造成社会混乱。

总之,突发事件情境下网络舆情信息摩擦不仅会导致应急信息传播受阻,还会导致政府应急处置效率降低,大大提高舆情发展态势不确定性,造成应急秩序混乱。因此,构建突发事件情境下网络舆情信息摩擦评估指标体系,对网络舆情信息摩擦进行评估与分析,对政府舆情治理能力提升具有重要意义。基于此,本文旨在解决以下问题:(1)如何对突发事件情境下网络舆情信息摩擦进行测度与评估?(2)突发事件不同生命周期和不同主题类型的信息摩擦是否存在差异?(3)网络舆情信息摩擦如何影响突发事件的应急处置效果?

2 理论基础与文献回顾

2.1 理论基础

信息摩擦的概念主要分为两方面:一是搜寻摩擦,

指的是交易双方需求不同且市场上存在大量供给与需求,使得信息搜寻和匹配难度增加,供给双方需要付出额外成本才能达成交易;二是信息不对称,指的是交易双方拥有的信息量不同^[4],交易双方易出现“道德风险”与“逆向选择”,从而导致成本增加。当前,有关社交媒体中信息摩擦的研究较少,主要是在理论层面对信息摩擦的类型、伦理影响以及两面性进行探讨^[5],少有研究者将信息摩擦与突发事件情境下的网络舆情研究相结合。因此,本文将信息摩擦与突发事件情境下的网络舆情结合,为突发事件情境下应急信息传播效率提高和政府舆情治理提供新的视角。

基于信息摩擦概念,本文结合信息熵理论、风险信息搜寻与利用理论以及危机沟通理论,构建和完善突发事件情境下网络舆情信息摩擦评估指标体系。信息熵理论由Shannon^[6]于1948年根据统计热力学提出,指信息环境中各种信息量的平均,常表示信息环境中信息的无序和混乱程度^[7],用于信息分析和计算权重。风险信息搜寻与利用理论由Yang等^[8]于2014年提出,被广泛运用于信息行为学领域。这一理论强调用户在风险环境下为获得有效信息会不断进行信息搜寻行为,用户对于信息质量的评估作为用户搜寻行为的关键动机,将直接影响信息搜寻行为是否继续。危机沟通理论由Coombs^[9]于2007年提出,被广泛用于危机情境下的公共管理领域。该理论认为在危机情境下政府等主体应针对危机回应准确、及时和有用的信息,为利益相关者提供援助并恢复秩序。危机情境下的高质量回应和沟通对于应急管理十分重要,如果沟通不当,将产生负面影响^[10]。上述理论分别从信息量化、信息搜寻行为和危机情境回应与沟通出发,为网络舆情信息摩擦评估指标体系构建过程中的信息无序程度量化、信息搜寻摩擦和信息不对称指标构建问题提供了坚实的理论基础。

本文基于新公共管理理论对应急处置效果进行测度。新公共管理理论是公共管理领域的重要理论,于20世纪80年代在西方国家兴起。该理论重新定义了政府与公众的关系,提出政府应为以工作为本的服务提供者,应遵循“顾客至上”这一管理方法^[11],即政府的各项工作应以公众为导向,以公众满意度为核心,对公众负责。这一理论对于本文的借鉴意义在于该理论对公众的高度重视,政府在工作过程中,应重视和聆听公众需求,并提高工作的效率、质量和有效性。基于上述

理论,本文将基于公众的角度,对应急处置效果进行评估。

2.2 信息摩擦评估的相关研究

经济学领域中信息摩擦的评估方法主要有两种,一是直接计算真实成本的增加,二是量化信息不对称或搜寻摩擦的程度^[12]。现有的网络舆情相关研究中,鲜有学者对信息摩擦进行评估。网络舆情中的信息摩擦评估也可以通过计算用户信息传播成本增加,或通过量化网络舆情中信息不对称和信息搜寻阻碍程度进行评估。目前,网络舆情中用户信息传播成本的研究主要集中在用户感知成本对信息行为的影响^[13]。由于用户在信息传播过程中的成本增加主观感受性强、涉及因素多,难以直接计算,其度量方法大多基于访谈、问卷等主观数据。

现有对网络舆情中信息搜寻阻碍与信息不对称的评估研究,为构建突发事件情境下的网络舆情信息摩擦评估体系提供了重要的理论参考。信息搜寻阻碍的相关研究认为信息搜寻阻碍主要表现为用户自身搜寻能力的缺乏和信息搜寻环境中的信息阻碍两方面^[14-16]。现有信息搜寻阻碍的评估主要是通过问卷、访谈等主观方法进行,尚未有研究通过真实客观数据评估。信息不对称的相关研究主要集中于政务舆情,且少有学者对其进行评估。李晓龙和徐鲲^[17]提出了政务微博信息对称程度模型,用政府作为程度、公众反响状况、政务微博平台建设情况和政治法律环境四个指标评估政务微博的信息对称程度。其他学者主要是在理论层面上探讨网络舆情中信息不对称的表现,部分学者认为信息不对称是造成网络舆情谣言兴起的主要原因^[18-20]。李靖和李慧龙^[21]认为政务舆情中的信息不对称表现为政府发布信息质量以及信息能否全面真实的被传送以及被理解。张勤和李静^[22]认为政府的信息公开和信息回应能力决定着网络舆情中的信息不对称。

综上所述,过去学者对于网络舆情中信息摩擦的评估大多只是在理论层面进行探讨,缺乏系统性的评估体系和实证研究;早期研究在指标构建、评估方法和数据获取过程中依赖主观数据和主观评判,缺乏客观量化。本文将基于微博客观指标的构建及客观权重设定,构建突发事件情境下网络舆情信息摩擦评估指标体系,对网络舆情中信息摩擦程度进行评估。

2.3 应急处置效果评估的相关研究

现有关于应急处置效果评估的研究较少,一部分

学者从应急处置对社会的综合影响出发进行评估^[23-24],一部分学者则侧重社交媒体在应急处置中发挥的效果^[25-26],以网络舆情为出发点进行评估。例如,王宇^[27]从信息公开程度、信息发布质量、次生危机控制、初次反应速度、持续回应情况、议题引导效果、媒体运用情况等方面评估重大突发事件社会舆情应急处置效果。徐选华等^[28]基于公众风险感知的大小和感知风险的可控性对应急决策质量进行评估,进而对应急处置效果进行评估。

以上研究结合定性和定量的方法对突发事件应急处置效果进行分析与评估,评估指标多以应急处置效果的影响因素为主,鲜有研究将应急处置后的结果表现作为评估指标。此外,从现有研究可以发现公众与网络舆情中的应急处置效果密切相关,因此本研究将以公众为出发点研究网络舆情中的应急处置效果。

3 研究设计

3.1 文本主题挖掘

突发事件情境下,人们往往会对事件的不同主题进行探讨,通过对社交媒体上突发事件相关博文进行主题挖掘,可以细化评估事件不同主题在信息传播过程中的信息摩擦。本文通过 Word2Vec 和 K-means 对突发事件相关博文进行主题聚类。聚类完成后,选取与簇中心距离最近的前十个词作为该聚类簇的特征词,然后总结主题,并将主题划分为事实类、情绪表达类、救援善后类、调查问责类和应对指南类五个类别^[29]。

3.2 信息摩擦定义与指标构建

网络舆情中的信息摩擦是指由于信息搜寻阻碍和信息不对称造成信息传播的额外成本增加。其中,信息搜寻阻碍指的是用户在海量信息中为获取有效信息需克服的阻碍;信息不对称则是指不同主体对于信息的掌握程度不同。相较于经济学领域的信息摩擦概念,网络舆情信息摩擦适用情境由商品交易情境转变为网络舆情情境;参与主体由商品交易双方转变为社交媒体用户;额外成本表现形式由交易成本转变为信息传播成本。然而,两者的本质都是由于信息搜寻阻碍和信息不对称导致的额外成本增加。本文根据定义,从信息搜寻阻碍和信息不对称两方面构建网络舆情信息摩擦评估体系,并遵循独立性、科学性、客观性和可操作性等原则,对突发事件不同主题时段的信息摩擦进行评估。

信息搜寻阻碍是信息摩擦的重要来源,用户面对网络舆情中的搜寻阻碍往往需要花费更多的成本搜寻可靠信息。本文从信息过载和信息质量两个方面评估网络舆情中的信息搜寻阻碍。其中,信息过载指的是用户所处信息环境存在大量信息,用户的信息处理能力不能满足其信息需求,导致用户无法获得有效的信息^[30]。基于信息熵理论,分别通过主题信息量和观点信息熵评估信息环境中的信息过载。主题信息量是指某一主题时段的信息数量,观点信息熵则为不同信源发出的各种观点的平均,体现了观点的无序程度。观点的多样性和无序性会造成用户信息搜寻过程中的信息过载,进而造成搜寻困难。根据风险信息搜寻与利用理论,信息质量同样会影响用户搜寻行为。信息质量决定着用户能否从信息发布内容中获取客观世界中事物真实和完整的表现、特征和规律^[31]。结合相关文献,本文认为信息的真实性、明确性、可读性、丰富性和相关性可以减少用户在搜寻过程中的阻碍。信息真实性是指非谣言信息的占比,信息的可信度或真实性越高,信息质量越高^[32];信息明确性是指信息表达是否明晰,信息明确可以提高信息发布的有效性^[29];信息可读性是指信息是否简单易懂,信息用语越简短往往越容易引发用户共鸣^[33];信息丰富性则是指信息的发布是否带有图片或者视频,图片和视频往往可以传达更多的信息^[34];信息相关性衡量了信息内容与主题的相关程度,相关程度越高,信息质量越高^[35]。

信息不对称是信息摩擦的另一重要来源,信息优势方的信息公开程度决定着网络舆情中信息优势方的信息能否全面并准确地传播并被公众接收和理解。突发事件网络舆情中,政务新媒体是最具权威性的主体和权威信息的发布者。除此以外,主流媒体作为政府的“喉舌”,也是突发事件发生时权威信息的重要来源。因此,政务新媒体和主流媒体作为权威主体处于信息优势地位。本文借鉴相关研究^[22]对于网络舆情中信息不对称的阐述,以政务新媒体和主流媒体作为信息优势方的信息公开程度评估网络舆情中的信息不对称。危机沟通理论强调了政府等主体在危机情境下回应和

沟通对于应急管理的重要性,本文从信息公开广度、信息公开强度和信息公开深度三个方面衡量权威主体的信息公开程度。信息公开广度是指突发事件情境下累计参与信息公开的权威主体数量,数量增多意味着信息公开更加广泛。信息公开强度包括信息公开数量和信息公开间断时间,信息公开数量是指权威主体累计公开的信息数量^[36],信息公开间断时间则体现了信息公开是否持续进行^[37]。信息公开深度是指信息公开内容的质量,包括信息公开真实性、信息公开明确性、信息公开可读性、信息公开丰富性、信息公开相关性,同时,本文增加信息公开滞后性这一指标,信息公开越滞后,信息公开质量越低^[38]。

基于此,本文从网络舆情中的信息搜寻阻碍和信息不对称两个维度的指标出发,构建了网络舆情信息摩擦评估指标体系,如表1所示。

表1 网络舆情信息摩擦评估指标体系

目标层	准则层	一级指标层	二级指标层	基础指标层	属性
信息摩擦	信息搜寻阻碍	信息过载	主题信息量	主题信息量	+
			观点信息熵	观点信息熵	+
		信息质量	信息真实性	非谣言微博占比	-
			信息明确性	确切词使用情况	-
				#和【】使用情况	-
			信息可读性	短微博占比	-
			信息丰富性	文字+视频微博占比	-
				文字+图片微博占比	-
			信息相关性	主题相关性	-
	信息不对称	信息公开广度	参与公开权威主体数量	累计参与公开权威主体数量	-
		信息公开强度	信息公开数量	累计信息公开数量	-
			信息公开间断时间	权威主体最新博文时间间隔	+
		信息公开深度	信息公开真实性	权威主体非谣言微博占比	-
			信息公开明确性	权威主体博文确切词使用情况	-
				权威主体博文#和【】使用情况	-
			信息公开可读性	权威主体短微博占比	-
			信息公开丰富性	权威主体文字+视频微博占比	-
				权威主体文字+图片微博占比	-
			信息公开相关性	权威主体博文主题相关性	-
			信息公开滞后性	权威主体博文首发时间间隔	+

3.3 信息摩擦的测度与评估

3.3.1 信息搜寻阻碍维度

信息过载通过主题信息量和观点信息熵进行评估。主题信息量即某一主题时段内的博文数量。观点

信息熵指某一主题时段内观点的信息熵,本文通过对博文进行SPO(Subject-Predicate-Object)三元组抽取实现网络舆情中特征级别的观点识别。使用哈尔滨工业大学社会计算与信息检索研究中心研制的语言技术平台(LTP)进行语义角色标注和依存句法分析,抽取原始SPO三元组,并对主谓宾缺失和过长三元组进行清洗,将相同的三元组归为一类观点,计算观点信息熵,计算方法如公式(1)所示。

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (1)$$

其中, $p(x_i)$ 表示达成同一类观点的三元组数量占所有三元组数量的比例, n 表示观点数量。

信息质量从信息真实性、明确性、可读性、丰富性和相关性5个方面进行评估。信息真实性通过非谣言微博占比评估,基于微博官方平台推出的辟谣账号“微博辟谣”和“微博社区管理中心”发布的辟谣微博识别谣言微博。信息明确性分别通过确切词使用情况和#和【】使用情况进行评估。确切词使用情况指某一主题时段内博文确切词数量与微博数量之比,如一定、一直等词语,确切词来自LIWC2015中文词典^[39],此词典用于中文文本词语的量化分析;#和【】使用情况指某一主题时段内博文中#和【】数量与微博数量之比。

信息可读性通过短微博占比进行评估,短微博占比越大,信息可读性越高,信息摩擦越低。微博平台将不超过140字的微博定义为短微博。短微博占比指某一主题时段内发布的短微博数量与所有微博数量的比值。信息丰富性通过文字+视频微博占比和文字+图片微博占比进行评估。文字+视频微博占比指某一主题时段内发布的带有视频的微博数量与所有微博数量之比;文字+图片微博占比指某一主题时段内发布的带有图片的微博数量与所有微博数量之比。信息相关性通过主题相关性进行评估。主题相关性指某一主题时段内发布的微博文本词向量与该主题关键词向量余弦相似度的均值。本文通过TF-IDF算法计算该主题微博文本词频并提取词频前15的关键词,生成主题关键词向量,通过平均法计算每条微博文本词向量与该主题关键词向量的余弦相似度。计算方法如公式(2)所示。

$$\text{Top}_{\text{relevance}} = \frac{\sum_{i=1}^m \text{sim}(P, U_i)}{m} \quad (2)$$

其中, P 表示该主题关键词向量, U_i 表示第*i*条微博文本的词向量, m 表示该主题时段所有微博数量。

3.3.2 信息不对称维度

由于权威主体的信息公开对于突发事件应急处置的影响是持续且深远的,仅对评估时间段的信息公开程度进行评估是片面和不完整的。因此,本文对权威主体对该主题的累计信息公开程度进行评估。信息公开广度指截止到待评估主题时段结束,累计参与该主题信息公开的政务微博和主流媒体数量。信息公开强度通过信息公开数量和信息公开间断时间进行评估。信息公开数量指截止到待评估主题时段结束,政务微博和主流媒体累计发布主题相关微博数量。信息公开间断时间指截止到待评估主题时段结束,政务微博和主流媒体发布的最新主题相关微博与评估时间的时间间隔。

信息公开深度通过信息公开真实性、明确性、可读性、丰富性、相关性和滞后性进行评估。信息公开真实性指截止到待评估主题时段结束,权威主体累计发布主题相关微博中非谣言微博的占比;信息公开明确性指截止到待评估主题时段结束,权威主体累计发布主题相关微博确切词数量以及#和【】数量与权威主体累计发布的相关主题微博数量之比;信息公开可读性指截止到待评估主题时段结束,权威主体累计发布主题相关短微博数量与权威主体累计发布主题相关微博数量的比值;信息公开丰富性指权威主体累计发布主题相关微博中带有视频的微博和带有图片的微博占比;信息公开相关性指截止到待评估主题时段结束,权威主体累计发布主题相关微博文本向量与该主题关键词向量余弦相似度的均值,计算方法同上文公式(2)所示;信息公开滞后性指截止到待评估主题时段结束,权威主体累计发布主题相关微博发布时间与该主题首发微博发布时间间隔的均值,时间间隔均值越大,信息公开越滞后。

本文将突发事件每个主题按天分组作为一个主题时段,针对不同主题时段的网络舆情抽取表1中的基础指标进行Max-Min标准化,并生成指标矩阵。其中,矩阵中的每一行代表一组数据,每一列代表一个基础指标,第*i*行第*j*列数据用 x_{ij} 表示。熵权法是一种旨在通过客观分析每个指标提供的信息量来确定各指标权重的方法,可以减少评估过程中主观判断对结果的影响,具有操作性强、客观性高等优点。本文采用熵权法计算信息摩擦评估指标体系各个指标的信息熵值、差异系数,进而确定指标权重。根据指标矩阵和指标权重

计算得出不同主题时段的信息摩擦,如公式(3)所示。

$$F_i = \sum_{j=1}^m X_{ij} \omega_j \quad (3)$$

其中, F_i 指的是第*i*个主题时段的信息摩擦取值, X_{ij} 指的是第*i*个主题时段标准化后的第*j*个指标值, ω_j 指的是第*j*个指标的权重。

3.4 基于生命周期和主题类型的信息摩擦比较分析

依据生命周期理论,将研究选取的三个突发事件的生命周期划分为起始期、爆发期、蔓延期和衰退期,将主题划分为调查问责类、情绪表达类、救援善后类、应对指南类和事实类共5个主题类型。本文采用Kruskal-Wallis非参数检验,判断不同生命周期和主题类型的信息摩擦是否存在显著性差异。Kruskal-Wallis非参数检验用于检验多个独立样本是否来自同一概率分布,即用于检验多个独立样本之间的差异,适用于数据量较小且数据不服从正态分布的情况。当P值小于0.05时,认为两个独立样本之间存在显著性差异。

3.5 突发事件应急处置效果测度

新公共管理理论强调满足公众需求是政府工作的基本导向。社交媒体作为政府应急处置工作的重要阵地之一,更应注重公众诉求,以防止突发事件舆情态势恶化。因此,本文基于新公共管理理论,从公众的角度测度应急处置效果。

一方面,公众的情绪倾向是舆情的核心内容,反映了舆情态势的变化以及公众对于应急处置的满意度^[40]。本文用公众博文情感倾向评估公众的情感倾向,公众博文情感倾向(*Sentiment*)指该主题在某一主题时段内公众发布微博情感分数的均值。利用百度智能云情感分析系统分析博文情感倾向,得到博文的消极、中性和积极情感分数分别为0、1、2。

另一方面,应急处置效果与公众的表达密不可分。首先,公众博文中的疑问词使用情况可以反映公众对于突发事件应急处置过程的信息需求以及公众对失职方的问责。公众博文疑问词使用情况(*Interrog*)指该主题在某一主题时段内公众发布微博包含的疑问词数量与公众微博数量之比,包括什么、怎么等词语,疑问词来自LIWC2015中文词典^[39]。其次,应急处置效果也体现在权威主体公开的信息是否完整、准确的传达到社会公众。信息在权威主体和公众之间传播的过程中容易遭遇信息畸变与失真^[21],导致应急处置效果不佳。因此,本文通过公众观点与政务微博、主流媒体观点的相似度衡量应急处置效果。公众观点与政务微博

观点相似度(*Government_{sim}*)指待评估主题距离本时段最新发布的10条政务微博观点与该时间段公众博文观点相似度。本文通过抽取微博的SPO三元组,计算政务微博三元组与公众微博三元组的余弦相似度,计算方法如公式(4)所示。

$$Government_{sim} = \frac{\sum_i \sum_j sim(SPO_i, SPO_j)}{10m} \quad (4)$$

其中, SPO_i 表示第*i*条政务微博文本中抽取的三元组, SPO_j 表示第*j*条公众微博文本中抽取的三元组, m 表示当前时间段公众发布微博数量。

公众观点与主流媒体观点相似度(*News_{sim}*)指待评估主题距离本时段最新发布的10条主流媒体微博观点与该时间段公众博文观点相似度,计算方法同上。

综上,分别对上述指标进行Max-Min标准化,生成指标矩阵,采用熵权法计算各个指标的信息熵值、差异系数,进而确定指标权重。根据指标矩阵和指标权重计算得出突发事件某一主题时段的应急处置效果,计算方法如公式(5)所示。

$$E_i = \alpha Sentiment_i - \beta Interrog_i + \gamma Government_{sim(i)} + \delta News_{sim(i)} \quad (5)$$

其中, E_i 表示第*i*个主题时段的应急处置效果值, $Sentiment_i$ 表示第*i*个主题时段的公众博文情感倾向, $Interrog_i$ 表示第*i*个主题时段的公众博文疑问词使用情况, $Government_{sim(i)}$ 表示第*i*个主题时段的公众观点与政务微博观点相似度, $News_{sim(i)}$ 表示第*i*个主题时段的公众观点与主流媒体观点相似度, α 、 β 、 γ 和 δ 为各指标权重。

3.6 网络舆情信息摩擦对应急处置效果的影响机制分析

网络舆情中的信息摩擦作为突发事件情境下有效信息传播的阻碍,会对突发事件应急处置效果造成影响。为探究突发事件情境下网络舆情信息摩擦与应急处置效果之间的相关关系,本文对信息摩擦与应急处置效果两个变量进行斯皮尔曼(Spearman)等级相关性分析,并进行显著性检验。

具体的,以表1中的信息摩擦评估基础指标作为自变量,应急处置效果作为因变量,构建回归预测模型。本文构建的模型主要有支持向量机回归模型(SVR)、Light Gradient Boosting Machine 回归模型(LGBMRegressor)、随机森林回归模型(RandomForestRegressor)、eXtreme Gradient Boosting 回归模型(XGBRegressor)和极限树回归模型(ExtraTreeRegressor)。通过 MAE

(Mean Absolute Error, 平均绝对误差)、MSE(Mean Square Error, 均方误差)和 R^2 对模型性能进行评估和比较,选择最优模型并将结果输入SHAP中。SHAP模型基于博弈论,可以对机器学习模型的输出结果进行解释^[41]。通过SHAP模型的可视化分析,不仅可以对不同特征的重要性进行排序,还可以展现不同特征对于因变量的正负影响。

4 实验与结果分析

4.1 案例选取与数据采集

本文以2023年齐齐哈尔体育馆坍塌、2024年湖北冻雨以及2024年梅大高速坍塌三个事件作为研究案例,通过Python爬虫爬取相关事件的原创与转发微博,以“齐齐哈尔体育馆”为关键词,获取2023年7月23日至2024年2月22日的相关微博共22472条;以“武汉冻雨”“湖北冻雨”为关键词,获取2024年1月30日至2024年3月1日的相关微博共25544条;以“梅大高速”“梅龙高速”为关键词,获取2024年5月1日至2024年7月1日的相关微博共38061条。同时,爬取微博发布者的用户相关信息,根据微博提供的微博认证字段筛选政务微博和媒体,当字段取值为“1”时,该微博账号为政务微博;当字段取值为“3”时,该微博账号为媒体。本文通过筛选主流媒体相关关键词识别媒体用户中的主流媒体^[42]。当媒体微博账户认证信息包含报纸、日报、时报、都市报等关键词时,该媒体用户为主流媒体。

4.2 主题识别与提取

为确保模型和算法训练效果,对微博文本进行数据清洗和分词。将分词后的微博文本通过Word2Vec模型进行训练,使用K-means算法对微博文本进行聚类,结合肘部法则确定不同事件的最佳K值,齐齐哈尔体育馆坍塌、湖北冻雨、梅大高速坍塌三个事件的最佳K值分别为10、10和8。通过对不同聚类簇的特征词进行人工总结,提取主题,将主题划分为调查问责类、事实类、情绪表达类、救援善后类和应对指南类五个类别。根据上文研究方法对于数据的要求,将三个事件不同主题的数据按天进行分组,得到940组主题时段数据。

4.3 网络舆情信息摩擦评估与排序

本文基于熵权法对三个突发事件不同主题时段的信息摩擦进行评估。首先,对数据进行Max-Min标准化,并生成指标矩阵。其次,对属性为负的指标进行转

换,确保 F_i 越大,信息摩擦越大。最后,通过熵权法计算各指标的信息熵值和差异系数,进而得出指标权重。将各指标权重与标准化后的数值代入公式(3),计算得到信息摩擦。为便于观察以及避免得到负值,将信息摩擦值进行数据平移,并扩大100倍,最终得到940组数据的信息摩擦值并排序,结果如表2所示。

表2 信息摩擦评估与排序

排名序号	事件	主题概括	日期	信息摩擦值
1	齐齐哈尔体育馆坍塌	事故伤亡情况	2023.07.24	32.20
2	齐齐哈尔体育馆坍塌	事故伤亡情况	2023.07.25	30.84
3	梅大高速坍塌	悼念遇难者	2024.05.02	29.28
4	梅大高速坍塌	悼念遇难者	2024.05.01	27.98
5	齐齐哈尔体育馆坍塌	事故起因调查	2024.02.21	27.84
.....				
936	梅大高速坍塌	认真学习灾害重要提示	2024.06.06	1.87
937	湖北冻雨	天气预报与预防	2024.02.26	1.54
938	湖北冻雨	天气预报与预防	2024.02.29	1.01
939	湖北冻雨	天气预报与预防	2024.02.25	0.06
940	湖北冻雨	天气预报与预防	2024.02.24	0.01

由表2可知,排名前5的主题为齐齐哈尔体育馆坍塌事件“事故伤亡情况”“事故起因调查”和梅大高速坍塌事件“悼念遇难者”。对照对应时段的原始数据可以发现,社交媒体用户就以上主题产生了大量讨论,发帖的观点无序程度较高,这增加了用户搜寻有效信息的成本。由于前四个时段处于事件前期,权威主体对于事故原因和伤亡的信息公开程度较低,这使得权威主体与公众之间的信息不对称程度较高,从而增加了信息摩擦。体育馆坍塌事件中的权威主体对于“事故起因调查”的信息发布时间距离事发时间已过七个月,信息公开滞后,信息摩擦较大。

排名最后的主题为高速坍塌事件“认真学习灾害重要提示”和冻雨事件“天气预报与预防”。对照对应时段的原始数据可以发现,梅大高速坍塌事件“认真学习灾害重要提示”这一主题时段,权威主体在社交媒体应急处置期间累计发布了众多相关主题的高质量博文,大大减少了这一主题的信息不对称程度,从而减少了信息摩擦。同样的,在对应时段对于“天气预报与预防”这一主题的博文信息质量高;参与信息公开的权威主体众多,权威主体对于信息公开强度和深度都较高,大大减少了信息不对称,使得这一主题的信息摩擦处

于较低水平。

4.4 基于生命周期和主题类型的网络舆情信息摩擦比较分析

4.4.1 基于生命周期的网络舆情信息摩擦比较分析

通过计算不同生命周期的信息摩擦平均值和中位数可以发现,起始期和爆发期的信息摩擦平均值和中位数均高于蔓延期和衰退期。为探究不同生命周期信息摩擦的差异,本文进行了 Kruskal-Wallis 非参数检验,检验结果如表3所示。可以发现,除了起始期和爆发期这对数据的显著性大于0.05(即不存在显著性差异)外,其他生命周期的信息摩擦之间均存在显著差异。

表3 不同生命周期信息摩擦的Kruskal-Wallis非参数检验

	检验统计	标准误差	标准检验统计	显著性	调整后显著性 ^a
起始期-爆发期	47.897	50.574	0.947	0.344	1.000
起始期-蔓延期	196.361	45.825	4.284	<.001	0.000
起始期-衰退期	127.355	43.931	2.899	0.004	0.022
爆发期-蔓延期	148.419	32.588	4.554	<.001	0.000
爆发期-衰退期	79.456	29.867	2.660	0.008	0.047
蔓延期-衰退期	-68.963	20.837	-3.310	<.001	0.006

注:已针对多项检验通过 Bonferroni 校正法调整显著性值。

4.4.2 基于主题类型的网络舆情信息摩擦比较分析

通过计算不同主题类型的网络舆情信息摩擦平均值和中位数可以发现,调查问责类和情绪表达类主题的信息摩擦平均值和中位数较高,事实类主题次之,应对指南类和救援善后类主题的信息摩擦平均值和中位数较低。为探究不同主题类型信息摩擦的差异,本文进行了 Kruskal-Wallis 非参数检验,检验结果如表4所示。可以发现,除了救援善后类和应对指南类、调查问责类和情绪表达类两对数据的显著性大于0.05(即不存在显著性差异)外,其他主题类型的信息摩擦之间均存在显著差异。

总体来看,不同生命周期、不同主题类型的信息摩擦表现出显著的差异性。起始期和爆发期信息摩擦较高,调查问责类和情绪表达类主题的信息摩擦较高。相关部门在事件起始期和爆发期应重点关注调查问责类和情绪表达类主题的信息传播疏导,减少网络舆情环境的不确定性;在事件后期则需均衡各类主题信息的传播,保持网络舆情环境的稳定性。

表4 不同主题类型信息摩擦的Kruskal-Wallis非参数检验

	检验统计	标准误差	标准检验统计	显著性	调整后显著性 ^a
救援善后-应对指南	-32.807	34.426	0.953	0.341	1.000
救援善后-事实	141.056	26.033	5.418	<.001	0.000
救援善后-调查问责	290.738	33.493	8.680	<.001	0.000
救援善后-情绪表达	336.571	28.558	11.786	<.001	0.000
应对指南-事实	183.862	30.395	5.702	<.001	0.000
应对指南-调查问责	323.544	36.986	8.748	<.001	0.000
应对指南-情绪表达	369.377	32.584	11.336	<.001	0.000
事实-调查问责	-149.682	29.335	-5.103	<.001	0.000
事实-情绪表达	-195.515	23.543	-8.304	<.001	0.000
调查问责-情绪表达	45.833	31.597	1.451	0.147	1.000

注:已针对多项检验通过 Bonferroni 校正法调整显著性值。

4.5 网络舆情信息摩擦对应急处置效果的影响机制分析

4.5.1 网络舆情信息摩擦与应急处置效果的相关性分析

本文基于公众角度,运用熵权法对突发事件的应急处置效果进行测度。为保证测度的有效性,筛选公众发量大于5的主题时段,计算各指标的权重,得到应急处置效果值。为探究信息摩擦与应急处置效果的关系,对信息摩擦值与应急处置效果值进行斯皮尔曼(Spearman)等级相关性分析,并进行显著性检验。如表5所示,信息摩擦与应急处置效果的 Spearman 相关系数为-0.3697,P值小于0.01。这表示突发事件情境下信息摩擦与应急处置效果两个变量之间存在显著的负相关关系。突发事件情境下,对信息摩擦进行控制将有利于应急处置效果的提升。

表5 信息摩擦与应急处置效果的相关性分析

信息摩擦	Similarity	
	Spearman 相关系数	P 值
应急处置效果	-0.3697	0.000

4.5.2 应急处置效果回归预测模型构建

为探究信息摩擦不同指标对应急处置效果的影响,本文基于5个机器学习模型构建应急处置效果回归预测模型。为避免变量之间存在多重共线性,计算每个变量的方差扩大因子(VIF),当变量的 VIF 大于等于10时,认为变量之间存在多重共线性。通过计算发现,累计参与公开权威主体数量和累计信息公开数量两个变量的 VIF 大于10,因此本文删除 VIF 较大的变量,即

累计参与公开权威主体数量。然后,将数据按照8:2分为训练集和测试集,通过MAE、MSE和 R^2 对模型表现进行评估,如表6所示。可以看出,极限树回归模型(ExtraTreeRegressor)具有最好的效果。

表6 各回归模型表现

回归模型	训练集			测试集		
	MAE	MSE	R^2	MAE	MSE	R^2
SVR	0.07347	0.00897	0.54415	0.06812	0.00778	0.60727
LGBMRegressor	0.02679	0.00150	0.92392	0.07612	0.01052	0.46874
RandomForestRegressor	0.03296	0.00248	0.87388	0.07622	0.01107	0.44112
XGBRegressor	0.00201	0.00001	0.99963	0.07158	0.00845	0.57364
ExtraTreeRegressor	0.01028	0.00153	0.98791	0.06359	0.00723	0.63484

4.5.3 影响机制分析

将极限树回归模型(ExtraTreeRegressor)的预测结果输入SHAP模型中,计算不同特征的SHAP值,SHAP值可以体现不同特征对于模型输出结果的贡献。SHAP模型的结果如图1所示,图1(a)为SHAP全局特征解释图,x轴为SHAP值,y轴为不同的特征,散点的颜色深浅代表特征值的大小,解释了不同特征对于模型预测结果的正负影响;图1(b)为SHAP全局特征重要性排序图,反映了不同特征对于因变量的影响程度。

结合图1,综合信息摩擦各级指标来看,信息搜寻阻碍相较于信息不对称对应急处置效果的影响更大。信息搜寻阻碍相关指标中,信息质量比信息过载对应急处置效果的影响更大,其中信息搜寻环境中#和【】使用情况、文字+视频微博占比、文字+图片微博占比、主题相关性在一定程度上与应急处置效果呈正相关,公众对于搜寻环境中信息质量的感知将提高应急处置效果。值得注意的是,信息搜寻环境中的确切词使用情况与应急处置效果呈负相关,这可能是因为确切词的使用虽然能够在一定程度上提高信息明确性,但突发事件具有不确定性和复杂性,强调确定性的语言可能与公众的实际需求和感知不符、与事件的实际复杂性不匹配,进而影响公众对应急处置的满意度,降低应急处置效果。信息过载相关指标对应急处置效果的影响较小,其中观点信息熵与应急处置效果呈负相关。

信息不对称相关指标中,信息公开深度和信息公开强度相关指标对应急处置效果的影响较大。信息公开深度相关指标中,权威主体累计公开博文的文字+图片微博占比、文字+视频微博占比在一定程度上与应急处置效果呈正相关,这表明权威主体信息公开形式越丰富,应急处置效果会越好。信息公开强度相关指标

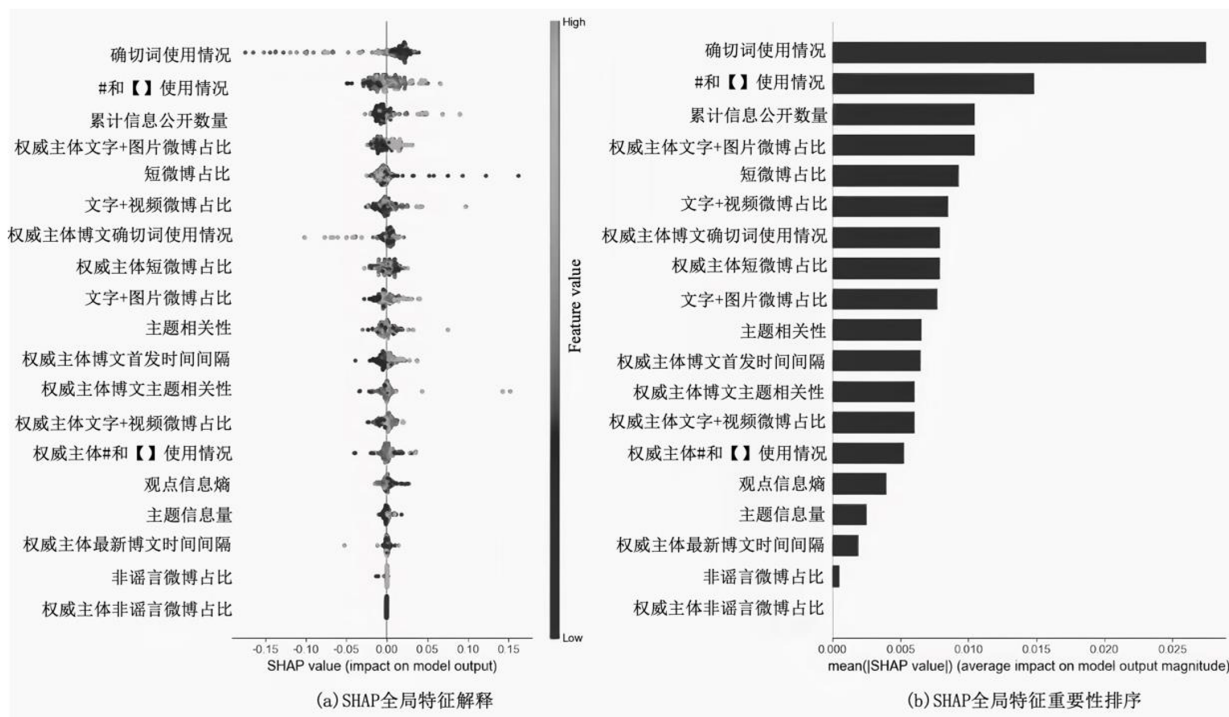


图1 SHAP模型结果图

中,累计信息公开数量与应急处置效果呈正相关,这表明权威主体应保证信息公开的数量,公开数量越多,应急处置效果越好。

5 结论与展望

本文基于信息经济学中信息摩擦的定义,结合国内外相关文献构建了突发事件情境下网络舆情信息摩擦评估指标体系,实现了对不同主题时段的信息摩擦评估,并基于生命周期和主题类型对信息摩擦进行比较分析。此外,本文通过构建机器学习模型和SHAP模型探究了信息摩擦对应急处置效果的影响机制。本文的研究结论在于:第一,构建了突发事件情境下网络舆情信息摩擦评估体系,基于客观赋值的熵权法计算各指标权重,对3个突发事件各主题时段的信息摩擦进行评估与排序,建立了突发事件情境下网络舆情信息摩擦的量化方法,帮助相关部门客观研判突发事件网络舆情信息传播情况,及时干预以减少信息环境的不确定性;第二,基于生命周期和主题类型对信息摩擦进行比较分析,可以发现不同生命周期和主题类型的信息摩擦存在显著差异,其中,事件起始期和爆发期的信息摩擦较高,调查问卷类和情绪表达类主题的信息摩擦较高;第三,基于公众角度测度突发事件各主题时段的应急处置效果,并探究信息摩擦对应急处置效果的影响机制,结果显示信息摩擦和应急处置效果之间存在显著负相关关系,信息搜寻阻碍相较于信息不对称对应急处置效果的影响更大。

本文的创新点在于:第一,本文将信息摩擦概念引入突发事件情境下网络舆情的研究中,从全新的视角探讨了突发事件相关信息传播可能遇到的阻碍;第二,本文构建了突发事件情境下网络舆情信息摩擦评估指标体系,为突发事件应急管理提供可参考的理论工具,对政府应急舆情治理能力提升具有重要意义;第三,本文探究了信息摩擦对应急处置效果的影响机制,揭示了信息摩擦与应急处置效果的负相关关系,相关部门可以重点关注高摩擦主题和高摩擦生命周期,以便及时干预,具有较高的实践意义。研究的不足之处在于,本文只基于微博数据展开研究,未来将考虑结合其他多模态网络平台数据,并进一步完善指标体系,以探讨网络舆情信息摩擦评估指标体系对于其他网络平台数据的普适性。

感谢图书情报国家级实验教学示范中心为本研究提供实验支持!

参考文献

- [1] 国务院办公厅关于印发《突发事件应急预案管理办法》的通知[EB/OL].[2024-02-17].https://www.gov.cn/zhengce/zhengceku/202402/content_6930817.htm.
- [2] 中国互联网信息中心(CNNIC).第53次《中国互联网发展状况统计报告》[EB/OL].[2024-03-22].<https://www.cnnic.net.cn/n4/2024/0322/c88-10964.html>.
- [3] 李三希,王泰茗,武珂璠.数字经济的信息摩擦:信息经济学视角的分析[J].北京交通大学学报(社会科学版),2021,20(4):12-22.
- [4] 作志忠.信息不对称理论及其经济学意义[J].经济学动态,1997(1):66-69.
- [5] TOMALIN M. Rethinking online friction in the information society[J]. Journal of Information Technology, 2023, 38(1): 2-15.
- [6] SHANNON C E. A mathematical theory of communication[J]. The Bell System Technical Journal, 1948, 27(3): 379-423.
- [7] 李根强,刘莎,张亚楠,等.信息熵理论视角下网络集群行为主体的观点演化研究[J].情报科学,2020,38(1):42-47,86.
- [8] YANG Z J, ALOE A, FEELEY T H. Risk information seeking and processing model: a meta-analysis[J]. Journal of Communication, 2014, 64(1): 20-41.
- [9] COOMBS W T. Protecting organization reputations during a crisis: the development and application of situational crisis communication theory [J]. Corporate Reputation Review, 2007, 10(3): 163-176.
- [10] VEIL S R, BUEHNER T, PALENVCHAR M J. A work in process literature review: incorporating social media in risk and crisis communication[J]. Journal of Contingencies and Crisis Management, 2011, 19(2): 110-122.
- [11] 何颖,李思然.新公共管理理论方法论评析[J].中国行政管理,2014(11):66-72.
- [12] ZHANG H G, CAO T T, LI H, et al. Dynamic measurement of news-driven information friction in China's carbon market: theory and evidence[J]. Energy Economics, 2021(95): 104994.
- [13] 刘宇桐,王晰巍,王楠阿雪,等.重大突发公共卫生事件下社交媒体信息传播的算法抵抗行为研究[J].图书情报工作,2024,68(9):98-109.
- [14] 赵栋祥,马费成,张奇萍.老年人健康信息搜寻行为的现象学研究[J].情报学报,2019,38(12):1320-1328.
- [15] JIA X, PANG Y, LIU L S. Online health information seeking behavior: a systematic review[J]. Healthcare, 2021, 9(12): 1740.

- [16] 刘静,安璐.突发公共卫生事件中社交媒体用户应急信息搜寻行为画像研究[J].情报理论与实践,2020,43(11):8-15.
- [17] 李晓龙,徐鲲.信息不对称视角下的政务微博管理策略研究[J].电子政务,2014(12):28-36.
- [18] 杨正,胡象明,刘浩然.从“缺场”到“在场”:公众在网络谣言协同治理中的新“人设”及特征[J].电子政务,2021(11):69-79.
- [19] JU C, JIANG Y, BAO F, et al. Online rumor diffusion model based on variation and silence phenomenon in the context of COVID-19[J]. Frontiers in Public Health, 2022(9): 788475.
- [20] 董伟建.网络传播中的信息不对称及其防范[J].中南民族大学学报(人文社会科学版),2013,33(5):161-164.
- [21] 李靖,李慧龙.政务舆情回应中的信息不对称探微[J].上海行政学院学报,2018,19(6):21-30.
- [22] 张勤,李静.论互联网背景下的政府公信力建设[J].中国行政管理,2015(8):19-22.
- [23] 陈雪龙,张钟,李悦.临机决策视角下的非常规突发事件应急处置方案生成[J].中国管理科学,2024,32(10):109-122.
- [24] 潘静静,王若琳,尤爱国,等.传染病暴发事件应急处置效果评价体系研究[J].现代预防医学,2020,47(2):330-335.
- [25] MUNIZ RODRIGUEZ K, OFORI S K, BAYLISS L C, et al. Social media use in emergency response to natural disasters: a systematic review with a public health perspective[J]. Disaster Medicine and Public Health Preparedness, 2020, 14(1): 139-149.
- [26] 梁新华,赵丽.基于DEMATEL方法的突发公共卫生事件信息公开关键影响因素识别研究[J].科技情报研究,2021,3(4):80-94.
- [27] 王宇.重大突发事件社会舆情应对效能评价及考核制度的探索[J].当代传播,2017(2):10-12,34.
- [28] 徐选华,马志鹏,陈晓红.基于公众偏好大数据分析的大群体应急决策质量动态演化研究[J].中国管理科学,2022,30(7):140-149.
- [29] 安璐,陈苗苗.突发事件情境下政务微博信息发布有效性评估[J].情报学报,2022,41(7):692-706.
- [30] KARR-WISNIEWSKI P, LU Y. When more is too much: operationalizing technology overload and exploring its impact on knowledge worker productivity[J]. Computers in Human Behavior, 2010, 26(5): 1061-1072.
- [31] 冯纓,张瑞云.基于用户体验的微博信息质量评估研究[J].图书馆学研究,2014(9):62-67,101.
- [32] 李全喜,徐嘉徽,魏骏巍,等.基于BP神经网络的共享服务平台资源信息质量评价研究——以短租类共享服务平台为例[J].图书情报工作,2019,63(10):125-133.
- [33] KAPLIAN A M, HAENLEIN M. Users of the world, unite! The challenges and opportunities of Social Media[J]. Business Horizons, 2010, 53(1): 59-68.
- [34] 孙帅,周毅.政务微博对突发事件的响应研究——以“7·21”北京特大暴雨灾害事件中的“北京发布”响应表现为个案[J].电子政务,2013(5):30-40.
- [35] KULKARNI A, WONG M, BELASRE T, et al. Quantifying the quality of web-based health information on student health center websites using a software tool: design and development study[J]. JMIR Formative Research, 2022, 6(2): e32360.
- [36] 王玥,郑磊.中国政务微信研究:特性、内容与互动[J].电子政务,2014(1):66-77.
- [37] 高萍,李爱生.制度、传播与技术逻辑:新冠疫情背景下的政府信息公开[J].现代传播(中国传媒大学学报),2021,43(8):125-130.
- [38] 肖诗依,文庭孝.后疫情时代重大突发公共卫生事件信息公开质量影响因素分析——立足公众视角[J].图书情报工作, 2021,65(23):41-49.
- [39] PENNEBAKER J W, BOYD R L, JORDAN K, et al. The development and psychometric properties of LIWC2015[R/OL]. [2024-11-27]. <http://hdl.handle.net/2152/31333>.
- [40] 邓卫华,吕佩.反转或缓解?突发事件政府回应有效性研究——基于在线文本情感分析[J].中国行政管理,2021(2):123-130.
- [41] MENG Y, YANG N H, QIAN Z L, et al. What makes an online review more helpful: an interpretation framework using XGBoost and SHAP values[J]. Journal of Theoretical and Applied Electronic Commerce Research, 2021, 16(3): 466-490.
- [42] 安璐,惠秋悦.热点事件情境下微博舆情反转预测[J].信息资源管理学报,2022,12(3):21-34.

【作者简介】安璐,女,1979年生,武汉大学信息管理学院教授。
成皓飞,女,2000年生,武汉大学信息管理学院硕士研究生。
收稿日期:2024-08-16

欢迎关注人大复印报刊资料
信息资源管理系列刊官方公众号
——图情档学界

