

企业危机舆情主题与情感演化研究 ——以“双汇生产车间乱象”为例

龚 睿,王筱莉,朱 株,连卓毅

(上海工程技术大学 管理学院,上海 201620)

摘 要: 针对网民讨论主题与情感对企业危机舆情传播的影响,基于生命周期理论,提出融合知识图谱和情感图谱的企业危机舆情主题与情感演化模型。将“双汇生产车间乱象”事件作为研究案例,首先收集与舆情相关的微博文本并进行舆情阶段划分;其次采用知识图谱、LDA 主题模型、朴素贝叶斯情感分类器和情感图谱对舆情各阶段的网民关注热点、讨论主题和情感演化规律进行实证分析。研究表明:网民讨论内容随舆情演化略有不同,但部分话题会被持续讨论;网民情感变化与主题演化规律相吻合,负向情感在爆发期达到峰值;媒体用户影响力高于微博红人,但普通用户易与微博红人产生情感共鸣。

关键词: 企业危机舆情;主题演化;情感演化;情感图谱;知识图谱

中图分类号: F279.23

文献标识码: A

DOI: 10.19358/j.issn.2097-1788.2023.01.011

引用格式: 龚睿,王筱莉,朱株,等. 企业危机舆情主题与情感演化研究——以“双汇生产车间乱象”为例[J]. 网络安全与数据治理, 2023, 42(1): 79-85.

Research on the theme and emotional evolution of enterprise crisis public opinion——Take “chaos in Shuanghui workshop” as an example

Gong Rui, Wang Xiaoli, Zhu Zhu, Lian Zhuoyi

(School of Management Studies, Shanghai University of Engineering Science, Shanghai 201620, China)

Abstract: Aiming at the influence of netizens' discussion topics and emotions on the spread of corporate crisis public opinion, based on the life cycle theory, this paper proposes an evolution model of corporate crisis public opinion topics and emotions that integrates knowledge maps and emotion maps. Taking the event of “chaos in Shuanghui production workshop” as a research case, firstly microblog texts related to public opinion and divide the public opinion stages are collected; Secondly, the knowledge map, LDA topic model, naive Bayesian emotion classifier and emotion map are used to make an empirical analysis of the Internet users' focus, discussion topics and emotion evolution laws at each stage of public opinion. The research shows that the discussion content of netizens varies slightly with the evolution of public opinion, but some topics will be discussed continuously; The change of netizens' emotion coincides with the evolution of the theme, and the negative emotion reaches its peak in the outbreak period; The influence of media users is higher than that of microblog celebrities, but ordinary users tend to have emotional resonance with microblog celebrities.

Key words: corporate crisis public opinion; theme evolution; emotional evolution; emotion map; knowledge graph

0 引言

随着互联网的发展,以微博为代表的新媒体平台层出不穷并成为舆情发生和发酵的主要场地^[1]。近年来,企业危机事件频繁发生并通过新媒体平台快速进入公众视野,网民针对舆情事件发表自己的观点和情感。网民情感对企业危机舆情的传播和发

酵具有刺激作用,若不积极引导,将影响企业形象甚至社会稳定。在企业危机舆情传播过程中,公众情感与讨论主题均在动态演变,各阶段往往具有不同的主题,不同主题下所表达的情感直接影响舆论的走向及舆情传播速度。主题和情感是舆情评价的重要指标^[2],同时意见领袖的行为能够影响网民的

观点形成和情感演变。因此,探究企业危机舆情主题演化特征及网民情感演变规律,了解意见领袖对网民情感的影响机制,对应对企业危机舆情有重要的现实意义。

目前关于企业危机舆情生成方面的研究,少数学者利用 fsQCA 探究企业危机舆情的生成机理^[3-4]。在企业危机舆情传播影响因素方面,有些学者基于主体层面研究不同主体对企业危机舆情传播的驱动作用^[5],部分学者基于行为层面分析不同干预措施对企业危机舆情传播的影响^[6-7]。在企业危机舆情应对方面,学者们主要从应对效果、应对策略等角度对企业危机舆情进行研究。在舆情应对效果研究中,大多学者采用实证分析法^[8]和构建指标体系^[9]探究舆情应对效果。在舆情应对策略的研究中,部分学者基于系统动力学理论对舆情传播机制进行分析,并在此基础上提出舆情应对策略^[10-11]。随着研究的不断深入,部分学者将微博文本与企业危机舆情应对策略相结合,探究舆情主题与网民情感变化规律,并在此基础上提出舆情应对策略^[12]。

综上所述,学者们从不同角度针对企业危机舆情进行了大量研究并取得丰硕成果。但针对企业危

机舆情探究主题与情感演变规律的研究还相对较少,且在现有的研究中学者们较少通过知识图谱整体展示舆情各阶段网民讨论热点,也较少使用情感图谱深入探究舆情情感演变规律揭示意见领袖对网民情感的影响机制。基于此,在现有研究基础上,本文考虑生命周期理论,融合知识图谱与情感图谱构建企业危机舆情主题与情感演化模型,并以“双汇生产车间乱象”事件为例进行实证分析,以此来揭示企业危机舆情主题及情感演化规律并了解意见领袖对网民情感的影响,为企业危机舆情治理提供理论依据。

1 企业危机舆情主题与情感演化模型构建

本文以企业危机舆情为研究对象,基于生命周期理论,融合知识图谱与情感图谱构建企业危机舆情主题与情感演化模型,如图 1 所示。

该模型主要由数据获取及预处理、舆情阶段划分、演化分析三个环节组成。

1.1 数据获取及预处理

通过分布式爬虫框架 Scrapy 爬取企业危机舆情相关微博及评论文本数据并对文本数据进行预处理,包括删除重复项使用 Jieba 分词脚本和自定

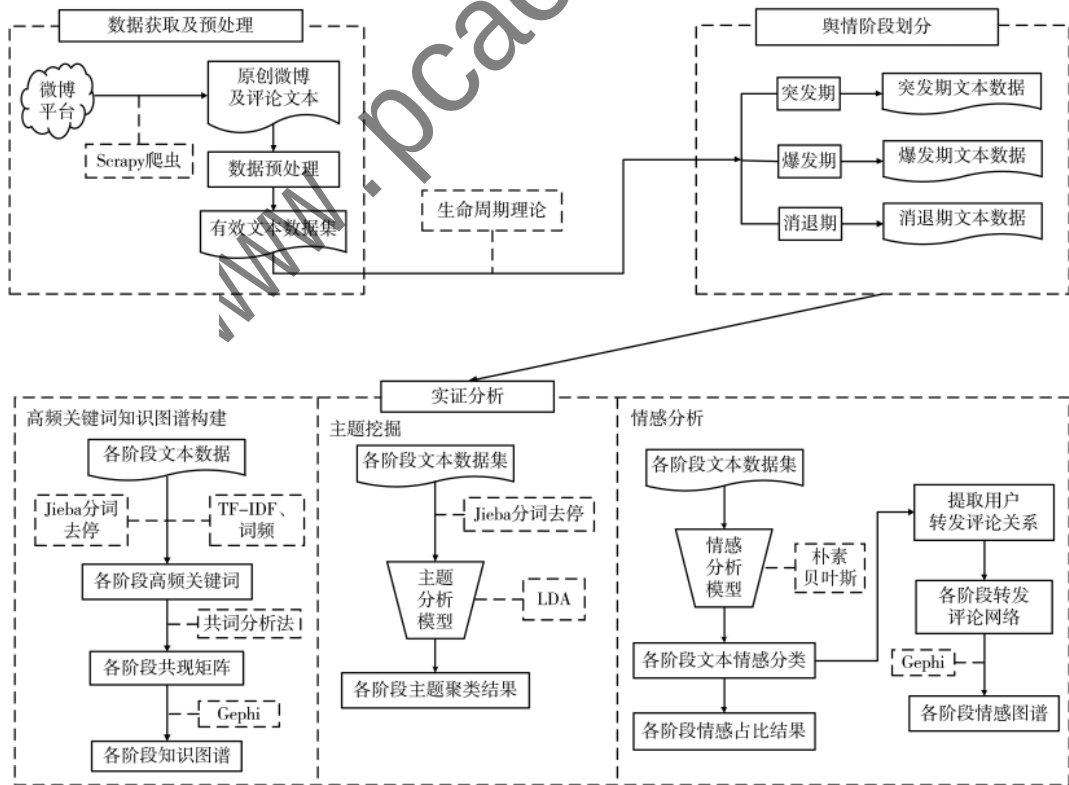


图 1 企业危机舆情主题与情感演化分析模型

义词典进行分词、去停等,得到实验语料库。

1.2 舆情阶段划分

基于生命周期理论,结合舆情发生的时间和传播趋势,将舆情划分为突发期、爆发期和消退期三个阶段。同时将实验语料库按照舆情阶段划分为三个阶段实验语料库。

1.3 演化分析

该环节基于各阶段实验语料库进行各阶段高频关键词知识图谱构建、主题挖掘和情感分析,内容如下:

(1)高频关键词知识图谱构建。首先基于实验语料库统计词语词频并结合 TF-IDF 算法计算词语权重,提取各阶段高频关键词。

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (1)$$

$$idf_i = \lg \frac{|D|}{|\{j:t_i \in d_j\}|} \quad (2)$$

$$tfidf_{i,j} = tf_{i,j} \times idf_i \quad (3)$$

其中,tfidf 是评估词语重要性的加权方法,tf 指词词频,idf 指逆词频。

其次,用共词分析法构建各阶段共现矩阵,利用 Gephi 软件绘制各阶段高频关键词知识图谱。整体把握网民的关注热点且综合认识网民讨论内容,并了解高频关键词在舆情事件中的地位及相互间的关系。其中,共现矩阵计算公式见式(4),当词 a 和词 b 同时出现在一句话中,则表示词 a 和词 b 之间有共现关系并形成一条连边,其边的权重加 1,反之两个词之不存在共现关系,即边的权重为 0^[13]。

$$y = \begin{cases} +1, & \text{词 } a、b \text{ 存在共现} \\ 0, & \text{词 } a、b \text{ 不存在共现} \end{cases} \quad (4)$$

(2)LDA 主题挖掘。基于实验语料库利用 LDA 主题模型进行各阶段主题挖掘。LDA 模型结构如图 2 所示。生成文档步骤如下:①计算文档-文档话题分布,针对所有文档进行采样,选择参数为 α 的狄利克雷分布,分布结果 $Q_m \sim \text{Dir}(\alpha)$, $m \in [1, M]$;②计算单词-话题分布,针对所有话题进行采样选择参数为 β 的狄利克雷分布,分布结果 $Q_m \sim \text{Dir}(\beta)$, $m \in [1, M]$;③针对第 m 篇文档的第 n 个词,从 Q_m 中抽取一个话题 $Z_{m,n}$,使 $Z_{m,n} \sim \text{Mult}i(Q_m)$,从 φ_k 中采用词 $w_{m,n}$,使 $w_{m,n} \sim \text{Mult}i(\varphi_k)$ 。其中, $\text{Mult}i(Q_m)$ 是 Q_m 的多项分布, $\text{Mult}i(\varphi_k)$ 是 φ_k 的多项分布。

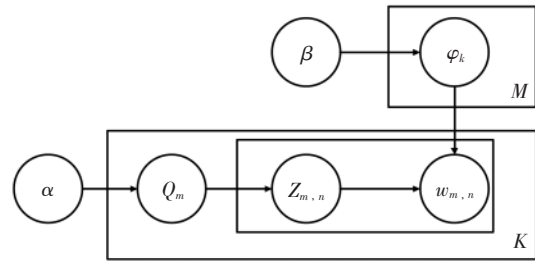


图 2 LDA 模型结构

(3)情感分析。该部分主要包括各阶段情感倾向占比分析和各阶段情感图谱绘制。首先,利用朴素贝叶斯分类器对各阶段实验语料库进行情感分类。其描述过程如下:①设 $x = \{a_1, a_2, \dots, a_m\}$, 其中 a 为 x 的某个特征属性;②有类别集合 $C = \{y_1, y_2, \dots, y_n\}$;③计算 $P(y_1|x), P(y_2|x), \dots, P(y_n|x)$;④若存在 $P = (y_k|x) \max\{P(y_1|x), P(y_2|x), \dots, P(y_n|x)\}$, 则 $x \in y_k$ 。

其中步骤③为整个计算过程中的关键步骤,该步骤无法直接计算得出条件概率,因此需要对训练集中各属性各特征属于各类别的条件概率进行统计。即:

$$P(y_i|x) = \frac{P(x|y_i)P(y_i)}{P(x)} \quad (5)$$

$$P(x|y_i)P(y_i) = P(y_i) \prod_{j=1}^m P(a_j|y_i) \quad (6)$$

其次,根据各阶段文本情感分类结果确定发表该文本的用户的情感倾向,通过 Python 提取各阶段用户间的转发评论关系,以微博用户为节点,用户之间的转发以及评论关系为边^[14],利用 Gephi 软件绘制各阶段情感演化图谱进行可视化情感演化分析。

2 实证分析

2.1 数据获取与舆情阶段划分

考虑到微博平台的高影响力和高关注度,本文以“双汇生产车间乱象”事件作为研究案例,基于微博平台借助分布式爬虫框架 Scrapy 爬取与该事件相关的热门微博及评论信息共获取相关数据 11 622 条并进行预处理,最终得到 7 874 条有效文本数据。依据生命周期理论,结合事件发生的时间以及传播量等特征,将此次舆情事件划分为三个阶段:3 月 14 日突发期阶段,共有 343 条有效文本数据;3 月 15 日~3 月 16 日爆发期阶段,共有 6 495 条有效文本数据;3 月 17 日~3 月 31 日消退期阶段,共有 1 009 条有效文本数据。

本节通过 TD-IDF 算法及词频统计提取各阶段权重排名前 300 且词频排名前 100 的高频关键词构建知识图谱,结果如图 3 所示。其中,词语出现频



由图3可知,在舆情整个时期,均是“双汇”一词为中心节点,其他关键词向边缘扩散。“双汇”“火腿肠”“食品”“安全”“生产”“车间”“猪排”等词联线较粗共现明显,这些关键词围绕此次舆情事件以及由双汇事件所引发的食品安全问题,表明双汇事件本身和食品安全问题是整个舆情期间网民的关注热点。

本节采用 LDA 主题模型进行各阶段主题挖掘, 具体结果如表 1 所示。

通过对表 1 中内容分析可知,部分话题会在相邻的两个阶段连续出现受到网民持续性地讨论,表明这些话题不仅是舆论热点还是网民关切的问题。网民对双汇事件本身持有不满意的态度,并从监管力度、处罚力度、事件发生根源等方面对监管部门和国家法律提出质疑。双汇集团对于此次舆情的应对效果并不好,究其缘由,从过去报道以及突发期挖掘结果可以发现双汇集团曾发生食品安全问题导致网民对该集团失去信任。因此,企业以及相关

表 1 各阶段主题挖掘结果

舆情阶段	主题编号	主题特征词(部分)	主题概括
突发期	TopicI-1	双汇、外包、曝光、瘦肉精、之前、举报、老子	双汇之前存在的问题
	TopicI-2	生产车间、双汇、南昌、发臭、落地、猪排、装袋	双汇事件本身
	TopicI-3	不吃、干净、能吃、问题、火腿肠、牌子、双汇、卫生	批评与抵制双汇
爆发期	TopicII-1	食品安全、问题、企业、品牌、相信、消费者	食品安全问题
	TopicII-2	双汇、消毒、员工、南昌、记者、生产、猪肉、外包、劳动法	双汇事件本身
	TopicII-3	双汇、平时、监管、监管部门、王中王、问题、爆出来、处罚	食品安全监管问题
	TopicII-4	目前、调查结果、双汇发展、安全隐患、联合、调查组	双汇发展进行调查
	TopicII-5	火腿肠、不吃、喜欢、恶心、老坛酸菜、干净、泡面	关于火腿肠和酸菜的讨论
	TopicII-6	停职、反省、声明、责任人、封存、深刻、处置	双汇致歉
消退期	TopicIII-1	罚罚、严查、责任人、判刑、愤怒、违法	重罚问题企业及责任人
	TopicIII-2	食品安全、良心、过轻、负责、太轻、根源、屡禁不止	食品安全事件频发的根源
	TopicIII-3	道歉、平时、干净、整改、犯罪、警察	不接受道歉和整改措施
	TopicIII-4	食品安全、老坛酸菜、瘦肉精、奶粉、三鹿、金锣、食品卫生	食品安全问题

部门可以实时监测舆情演变动态并根据舆情主题演化特征,针对各阶段不同的热点舆论采取合理及差异化的应对措施。

2.4 情感分析

2.4.1 情感倾向性分析

本节利用朴素贝叶斯分类器对微博文本进行情感分类,分类结果如表 2 所示。

表 2 各阶段网民情感倾向占比

舆情演化阶段	负向情感占比/%	中性情感占比/%	正向情感占比/%
突发期	63	27	10
爆发期	75	17	8
消退期	62	29	9
总和	73	18	9

由表 2 可知,在整个舆情期间,负向情感占主导地位占比为 73%,而正向情感占比在整个舆情期间波动不明显,表明在此次舆情中网民对食品安全问题的厌烦且在整个舆情传播过程中有小部分网友能坚持自己的立场。结合主题挖掘结果,舆情突发期阶段中性情感占比 27%相对较高,是由于部分网民对微博报道进行一个转载使得有些评论信息和微博没有明显的情感倾向。舆情爆发期阶段,舆情事件得到了网友的广泛参与,多数网友从双汇事件本身上升到企业食品安全问题、监管问题,并对监管部门和双汇集团发表了负面评论。在此期间酸菜也被爆出负面新闻,部分网友将两者联系在一

起,更激发了网民对于企业食品安全问题的负面情绪,因此爆发期阶段负向情感占比高达 75%,在整个舆情演化周期中占比最高。在舆情消退期阶段,负面情感由爆发期的 75%下降到 62%,中性情感由爆发期的 17%上升到 29%。此阶段部分网友开始理智地探讨食品安全问题频发的根源,形成了负向情感向中性情感的转变。

2.4.2 情感图谱构建

为了更加深入地了解企业危机舆情情感演化规律并了解意见领袖对普通用户情感的影响机制,本节进行情感图谱绘制,如图 4 所示。其中小、中、大节点分布代表拥有负面、中性、正面情感的网民。舆情突发期,共有 316 个节点、313 条边,普通节点形成了一定的负面情感聚集。舆情爆发期,共有 3 968 个节点、3 975 条边,普通节点数量迅速增加,负面情感聚集现象更加集中,以媒体用户为核心的节点其影响力变大。舆情消退期,共有 285 个节点、282 条边。表明企业危机舆情的突发期和爆发期是舆情管控的关键时期。

从节点影响力来看,在舆情突发期,与微博红人和媒体用户相连的普通用户节点数量差距不大;在舆情爆发期,与微博红人相比,普通用户与媒体用户连接更多;在舆情消退期,与普通用户节点相连的多为媒体用户节点。表明在整个舆情传播过程中,媒体用户影响力最为显著,微博红人影响力略低于媒体用户。从节点情感倾向来看,媒体用户多为中性情感。微博红人情感倾向多为负向。但多数

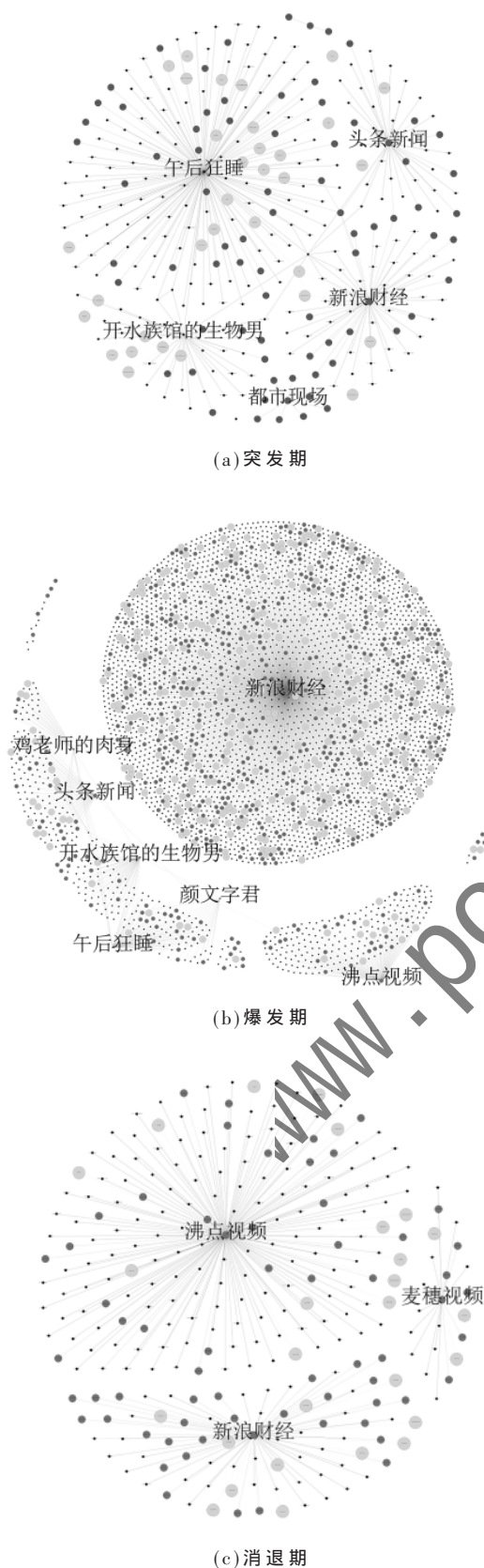


图4 各阶段情感图谱

普通节点与微博红人情感类别一致,表明其情绪对社群内其他普通用户的情感有较为直接的影响,容易引起普通用户的情感共鸣。因此,在舆情管理中,应加强对微博红人的舆论引导,以防其可能的不当言论,扩大事件影响力。

3 结论

本文基于生命周期理论,融合知识图谱和情感图谱构建企业危机舆情主题与情感演化模型。以“双汇生产车间乱象”事件为例,采用知识图谱、LDA主题模型、朴素贝叶斯情感分类器和情感图谱对舆情不同阶段的网民关注热点、讨论主题和情感演变规律进行实证分析。通过实证分析得出结论:(1)随着舆情的不断演化,网民讨论内容略有不同,但部分话题受到持续性讨论。(2)网民情感变化与舆情主题演化规律相吻合。突发期阶段开始产生负面情绪;随着舆情的扩散负面情绪在爆发期阶段达到峰值;舆情消退期部分网民回归理智负面情绪较爆发期有所下降。(3)媒体用户和微博红人多为舆情传播过程中的核心节点,媒体用户的影响力略高于微博红人,但普通用户的情感容易受到微博红人的影响。

本文仅选择“双汇生产车间乱象”事件进行分析,而不同企业危机舆情事件可能会存在差异,将多个案例进行对比研究,进一步深入探讨企业危机舆情主题与情感演化规律将是接下来的研究重点。

参考文献

- [1] 安璐,吴林.融合主题与情感特征的突发事件微博舆情演化分析[J].图书情报工作,2017,61(15):120-129.
- [2] 刘勇,韩清云.基于主题与多元情感融合的网络舆情动态分析方法研究[J].竞争情报,2021,17(5):10-18.
- [3] 杨波,谢乐.企业危机事件网络舆情传播态势生成机理研究——基于信息生态的多阶段fsQCA分析[J].管理评论,2022,34(7):339-352.
- [4] WANG M, SUN J. Generation mechanism of corporate online public opinion hotness based on multicase qualitative comparative analysis[J]. Discrete Dynamics in Nature and Society, 2021: 1-11.
- [5] 姚晶晶,姜靓,姚洪兴.基于SIR模型的情绪信息传播研究[J].情报科学,2018,36(10):25-29.
- [6] 王康,李含伟.自媒体时代的企业网络舆情应对策略研究——基于上市公司百度指数的研究[J].情

- 报科学, 2018, 36(1): 113-117.
- [7] KIM Y, LI H, LI S. Corporate social responsibility and stock price crash risk[J]. Journal of Banking & Finance, 2014(43): 1-13.
- [8] 刘春林, 张宁. 上市公司传闻的澄清效果研究——来自中国证券市场的证据[J]. 管理科学学报, 2012, 15(5): 42-54.
- [9] 张军玲. 基于层次分析法的企业网络舆情危机应对评价研究[J]. 经济数学, 2015, 32(3): 60-63.
- [10] 阎海燕, 詹凌云, 陈明明, 等. 基于系统动力学的企业危机事件网络舆情传播与应对研究[J]. 系统科学学报, 2021, 29(1): 92-97.
- [11] 齐丽云, 李晓鸿, 曹硕. 企业社会责任负面事件网络舆情演化与政企合作研究[J]. 系统工程理论与实践, 2020, 40(7): 1792-1805.
- [12] 张尧政, 邓少灵. 基于文本情感分析的企业网络舆情应对策略比较研究[J]. 电子商务, 2019(5): 32-35.
- [13] 令狐秋萍, 何世群, 齐梦珂, 等. 基于知识图谱和层次聚类的水族文化主题演化研究[J]. 现代计算机, 2022, 28(1): 1-9.
- [14] 张柳, 王晰巍, 王铎, 等. 微博环境下高校舆情情感演化图谱研究——以新浪微博“高校学术不端”话题为例[J]. 现代情报, 2019, 39(10): 119-126, 135.
- (收稿日期: 2022-11-16)
- 作者简介:
- 龚睿(1998-), 女, 硕士研究生, 主要研究方向: 企业舆情管理。
- 王筱莉(1981-), 通信作者, 女, 博士, 副教授, 硕士生导师, 主要研究方向: 应急管理、舆情管理。E-mail: 03140021@sues.edu.cn。

(上接第 71 页)

- 07-26[2022-11-03]. https://www.qianxin.com/news/detail?news_id=5770.
- [7] ASRC. 报告漏洞[EB/OL]. (2021-06-02)[2022-11-03]. <https://security.alibaba.com/report.htm>.
- [8] 绿盟科技新安盟. 软件供应链安全技术白皮书[EB/OL]. (2021-06-02)[2022-11-03]. <http://blog.nsfocus.net/wp-content/uploads/2022/07/NSFOCUS.pdf>.
- [9] 哈力木拉提·巴图尔. 浅谈软件开发中开源组件的应用及产生的安全问题[J]. 中国新通信, 2019(6): 105-105.
- [10] 苏仟, 赵娆. 开源软件供应链安全风险分析与发展建议[J]. 信息安全研究, 2022, 8(8): 831-835.
- [11] 王晓冬. 我国开源软件产业面临的突出风险及对策研究[J]. 信息安全研究, 2021, 10(10): 973-976.
- (收稿日期: 2022-11-23)
- 作者简介:
- 余建利(1973-), 男, 硕士研究生, 高级工程师, 主要研究方向: 电信网络的安全防护。
- 姜荣霞(1994-), 女, 硕士研究生, 主要研究方向: 高级威胁检测与响应、软件供应链安全。
- 卢蓉(1985-), 女, 本科, 主要研究方向: 高级威胁检测与响应、软件供应链安全、零信任安全防护体系。



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