

# 不同社交媒体的舆情比较研究： 以“胖猫事件”为例

郭军巧, 钱颖

上海理工大学管理学院, 上海

收稿日期: 2024年9月16日; 录用日期: 2024年9月24日; 发布日期: 2024年10月22日

## 摘要

研究目的: 互联网加速了舆情事件在社交媒体平台上的传播, 除了在以文本内容为主的微博上快速传播外, 视频网站也是舆情发酵的重要平台。既有舆情研究多聚焦微博, 对视频网站的研究较少, 更缺乏对两者的对比。以“胖猫事件”为案例, 对比研究微博和Bilibili用户的话题和态度。研究方法: 利用TF-IDF提取文本特征词及权重, 结合LDA主题聚类识别焦点词和核心议题; 使用百度情感分析API, 比较不同平台和主题下的用户情感。研究结果: 微博和Bilibili在主题和情感表现上既有相似性, 也存在差异: 1) 主题内容: 微博和Bilibili前期的讨论主要集中在情感与性别话题; 但警方通报后, 微博侧重对真相的反应和网络舆论的讨论, Bilibili则延续了前期对情感和家庭影响的关注。微博讨论主题相对分散, 而B站较为集中。2) 情感表现: 负面情感在两平台均占主导, 但微博更为强烈。这与微博的快速传播和部分蹭热度行为密切相关, 微博上“流量与热度”和“网络舆论”两个主题的负面情绪尤其高。综上, 两平台舆情主题和情绪有一定相似性, 但微博蹭热度现象更明显, 这与文本为主的社交平台信息发布门槛低有关, 而视频网站的讨论相对更集中更深入。这一发现为社交媒体管理和舆情监控提供了有价值的参考。

## 关键词

网络舆情, 平台差异, LDA模型, 情感分析

# A Comparative Study of Public Opinion on Different Social Media Platforms: The Case of the “Fat Cat Incident”

Junqiao Guo, Ying Qian

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Sep. 16<sup>th</sup>, 2024; accepted: Sep. 24<sup>th</sup>, 2024; published: Oct. 22<sup>nd</sup>, 2024

## Abstract

**Research purpose:** The Internet has accelerated the dissemination of public opinion events on social media platforms. In addition to the rapid dissemination on text-based platforms like Weibo, video websites are also important platforms for the escalation of such events. Existing research on public opinion predominantly focuses on Weibo, with comparatively little attention given to video platforms, and even fewer studies offering a comparative analysis between the two. Using the “Fat Cat Incident” as a case study, this research conducts a comparative analysis of user topics and attitudes on Weibo and Bilibili. **Research method:** Using TF-IDF to extract textual feature words and weights, combined with LDA topic clustering to identify focus words and core topics; using Baidu Sentiment Analysis API to compare user sentiment under different platforms and topics. **Research result:** Weibo and Bilibili exhibit both similarities and differences in thematic focus and emotional expression: 1) In the early stages, discussions on both platforms centered primarily around emotions and gender-related issues. However, after the police report was released, Weibo shifted its focus to reactions to the truth and broader public opinion, whereas Bilibili continued to emphasize the emotional and familial impacts. Discussions on Weibo were more dispersed, while Bilibili’s were more concentrated. 2) Emotional Expression: Negative emotions dominated on both platforms, but they were more intense on Weibo. This heightened negativity is closely linked to Weibo’s rapid dissemination and instances of opportunistic content for attention-seeking. In particular, negative sentiment was significantly higher in topics related to “traffic and trending” and “online public opinion” on Weibo. In summary, while there are notable similarities in the themes and emotions across both platforms, the phenomenon of attention-seeking is more pronounced on Weibo, likely due to the lower barriers to content dissemination on text-based social media. In contrast, discussions on video platforms like Bilibili tend to be more focused and in-depth. These findings offer valuable insights for social media management and public opinion monitoring.

## Keywords

Online Public Opinion, Media Platform Differences, LDA Model, Sentiment Analysis

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

互联网发展和社交媒体应用普及为公众创造了社会参与及获取教育、娱乐、法律等服务的平台。各类平台吸引众多用户交流和分享日常生活中的想法和感受,成为公众讨论社会问题、表达观点的重要渠道。随着不同类型社交媒体的蓬勃发展,跨平台使用成为了普遍现象。

网络舆情通常伴随着社交媒体的公众舆论。社交媒体不仅为公众情感的传播提供了媒介,蕴含的内容更是缓解网络舆情的宝贵数据来源[1]。由于文本数据易于获取,信息、情感丰富,可以有效识别主体的评价和情感[2]。因此,主题挖掘和情感分析广泛应用于网络舆情领域。Bai and Yu (2016)通过分析含有灾害信息的微博帖子,提出了一个三阶段结构化框架:识别灾害相关信息-筛选负面情感信息-预测灾民负面情感变化趋势,并利用雅安地震的数据验证了该框架在潜在公共危机事件监测中的有效性,对灾后救援和应急管理具有重要意义[3]。王冲等(2024)通过挖掘 Bilibili 上关于“ChatGPT 教育应用”的舆论,发现社会大众对“ChatGPT 教育应用”总体持积极态度,关注点涵盖“学生学习、教育公平”等方向,

帮助从师生政企的层面提出优化建议[4]。此外, YouTube [5] [6]、知乎[7]、小红书[8]等也是研究网络舆情现象的重要数据来源, 但以往既有研究多采用单一平台或整合多平台数据。

跨平台使用的日益增加, 使不同类型社媒平台的用户观点行为进入研究者视野。刘威等(2019)分析微博和微信用户自我呈现的差异, 认为微信与用户间联系更强, 更偏向“前台”, 而微博更偏向“后台”[9]。Wang and Huang (2022)将微博与小红书进行对比, 发现受访者在日常生活中使用微博的频率更高, 但小红书更有可能影响他们旅游目的地选择[10]。胡媛等(2024)对比文字(知乎)、视频(抖音、Bilibili)、混合(微博)三个类型的社媒平台, 得出视频、混合型平台内容更情绪化, 文字类型平台的内容更专业与严谨等结论[11]。李萌认为微博、抖音和 Bilibili 在社交属性、短视频代表性和用户画像三个角度存在异同, 故选取这三个平台的评论分析食品安全这类负面新闻, 发现在用户情绪差异上, 微博用户理性回归最明显, 这是因为微博上除了热衷娱乐的用户, 还有专业性强的专家和组织等特定用户群体[12]。通过文献研究发现, 对比不同类型社媒平台网络舆情差异的研究较少, 多聚焦微博、微信和知乎等偏文本型的平台, 对近年来广泛流行的视频型社交媒体——Bilibili、抖音的对比研究更加缺乏。同一网络舆情事件在不同类型平台上发酵的情况往往不同, 用户的态度和情绪也有所差异。微博和 Bilibili 作为中国两大知名社交媒体平台, 具有不同的社交媒体特征。微博以其简洁的文本信息和快速传播力著称, 使用门槛低, 用户可以通过短文本快速分享和讨论热点话题, 是影响力日益增强的网络舆论传播中心[13]。近年来, Bilibili 在年轻人中影响力逐渐提升, 吸引了共青团等主流媒体入驻。作为典型的亚文化网络社区, Bilibili 以其丰富的视频内容和强大的社区互动功能见长, 拥有稳定的用户群体, 活跃度高且及时关注热点事件[4], 通过弹幕评论和兴趣内容的分享形成了独特的社交生态。

本研究以热点事件为案例, 对微博帖子 and Bilibili 评论从内容和情感两个角度分析, 研究不同类型平台的舆情差异, 探索舆情热点和情感倾向, 从而为制定更有效的舆情管理策略提供理论支持和实践指导, 具有重要的现实意义和学术价值。

## 2. 研究设计

### 2.1. 案例选择

案例选取最近引起网络暴力的社会热点——“胖猫事件”。2024年4月11日21岁游戏代练“胖猫”在重庆长江大桥跳江后离世。5月2日, “胖猫”姐姐在微博上晒出“胖猫”和女友双方的聊天记录的细节信息, 引发舆论热议, 舆情热度逐渐上涨。3日, “胖猫”姐姐发布了弟弟的火化证照。网民自发悼念, 通过外卖平台购买奶茶、汉堡等食物送至涉事大桥, 随后茶百道、华莱士等品牌被曝外卖出现空杯、空包、白水等情况, 引发舆论不满。随后相关涉事品牌相继发布处理结果并致歉。7日, “胖猫”姐姐回应已经正式交给公安机关处理, 律师团队也正对相关证据进行取证收集。由于舆情热度, 微博上还出现冒充男女双方的账号来骗取流量的行为。19日, 重庆市公安局南岸区分局官方“平安南岸”发布警情通报。

以5月19日警方通报为时间节点, 将“胖猫事件”分为前期和后期。以此为案例考虑到话题与公众的社会距离小, 且舆论在网络上持续的时间也比较长(张晓君, 2017) [14]。

### 2.2. 方法框架

图1展示了本研究的方法框架。以“胖猫事件”作为爬虫关键词, 收集指定时间内的微博帖子 and Bilibili 视频评论。其次, 对收集的文本数据进行预处理, 去除重复数据、无关数据和异常数据, 使用 Python 中的结巴分词(jieba)包进行中文分词和去停用词处理, 输出可以使用的有效文本。最后, 通过 TF-IDF 方法提取文本数据中的特征词及其权重, 生成词云图; 运用 LDA 主题建模生成焦点主题; 利用百度情感分析 API 进行情感分析。

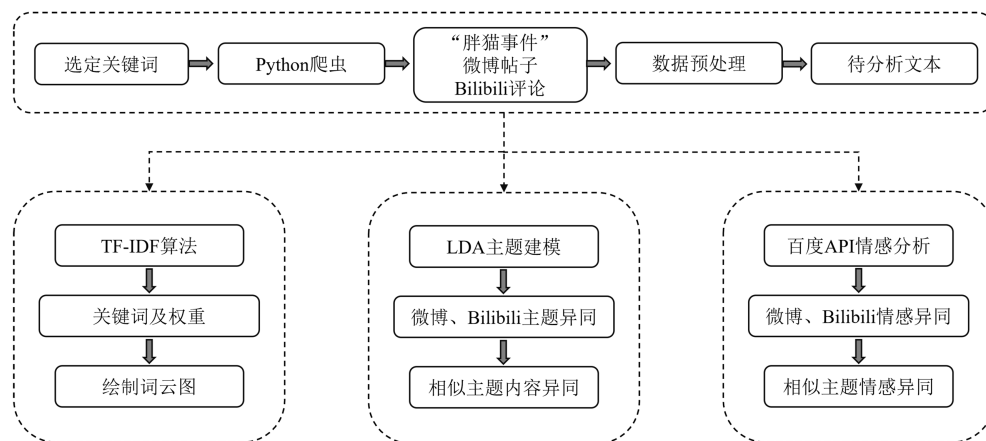


Figure 1. Research methodological framework

图 1. 研究方法框架

## 2.3. 数据收集与预处理

选定“胖猫事件”为关键词，采集时间为2024年5月1日至28日。在微博上初步采集到16,122条帖子，预处理后，得到15,616条有效帖子，共计432,707评论，平均每条帖子有27.7条评论回复。在Bilibili上初步抓取946个视频，总计622,759条评论。每天选择1个回复量多且讨论“胖猫事件”整体内容而不是某个侧面的视频，最终筛选出28个视频，共计43,225条评论，平均每个视频1544条评论。微博帖子数量庞大，但单条帖子的平均评论量仅有27.7条，讨论分散，Bilibili则显示出高度集中的讨论趋势。

## 2.4. 研究方法

### 2.4.1. TF-IDF

TF-IDF是自然语言处理中一种算法，广泛应用于信息检索与文本挖掘，提取并计算文档中术语的重要性[15][16]。TF-IDF涉及两个关键部分：术语频率( $tf_{i,d}$ )和逆文档概率( $idf_i$ )。术语频率通过将术语的总出现次数除以特定文档中的术语总数来衡量术语在文档中出现的频率。逆文档概率通过计算文档总数与包含该术语的文档数量之间的比率对数来评估术语在整个文档集合中的重要性。通过TF-IDF分数排名，可以提取最能代表文档内容的关键术语。

假如术语*i*在文本*D*中出现的词频为 $tf_{i,d}$ ， $idf_i$ 为含有术语*i*的文本数，则TF-IDF函数为：

$$tf-idf_{i,d} = tf_{i,d} \times idf_i = tf_{i,d} \times \log \frac{D}{1 + |j: t_i \in d_j|} \quad (1)$$

### 2.4.2. LDA 主题建模

潜在狄利克雷分配(LDA)主题建模是Blei(2003)等提出的一个“文本-主题-词汇”三层贝叶斯概率模型[17]，通过判断单词分布来识别文档集合中潜在的主题[18]。它假设每个文档都包含这些潜在主题的混合物，每个主题都以单词分布为特征。文本间的区别主要在于它们的主题混合规则不同。通过LDA可以获取每个主题下的词语分布概率，以及文档对应的主题概率，根据所得的概率进行主题分类。LDA因其对用户友好的配置成为目前使用最广泛的算法之一，研究将利用LDA算法挖掘网络舆情的主题。

LDA公式如下：

$$p(\vec{w}, \vec{z} | \vec{\alpha}, \vec{\beta}) = p(\vec{w} | \vec{z}, \vec{\beta}) p(\vec{z} | \vec{\alpha}) \quad (2)$$

使用困惑性指数作为指标来确定 LDA 结果中最优的主题数量, 公式如下:

$$\text{perplexity}(D_{\text{test}}) = \exp \left[ \frac{\sum_{d=1}^M \log p(w_d)}{\sum_{d=1}^M N_d} \right] \quad (3)$$

### 2.4.3. 情感分析

百度情感分析 API 是一种基于自然语言处理技术的工具, 自动分析文本情感倾向。使用 Python 调用百度情感分析的 SDK 模块, 应用百度 AI 平台的用户账号(API\_ID\_API\_Key)和密钥(Secret\_Key), 以 post 方式调用百度 AipNlp 情感分析接口, 输入文本后, 返回相应的情感倾向(正面-2、中性-1、负面-0)及置信度。

## 3. 文本情感与舆情主题分析

### 3.1. 高频特征词及权重

高频词反映用户对舆情事件的关注角度和讨论焦点, 分析高频词可以有效把握事件的核心议题。“胖猫事件”前、后期的前 15 个特征词及其频率如表 1 所示。

**Table 1.** Top 15 characteristic words of Weibo and Bilibili  
**表 1.** 微博、Bilibili 特征词(Top 15)

前期 - 微博		前期 - Bilibili		后期 - 微博		后期 - Bilibili	
特征词	权重	特征词	权重	特征词	权重	特征词	权重
姐姐	0.00758	女人	0.01492	网络	0.01079	反转	0.01539
女人	0.00740	男人	0.01283	姐姐	0.00921	姐姐	0.01225
男人	0.00688	女性	0.01027	公布	0.00876	恋爱	0.01204
证明	0.00486	社会	0.00809	警方	0.00802	转账	0.00948
死亡	0.00470	男性	0.00676	重庆	0.00794	女方	0.00870
恋爱	0.00452	恋爱	0.00607	细节	0.00792	官方	0.00781
网民	0.00436	爱情	0.00560	谣言	0.00725	网民	0.00742
社会	0.00431	结婚	0.00539	恋爱	0.00709	聊天记录	0.00732
爱情	0.00425	纯爱	0.00501	账号	0.00698	通报	0.00705
樱子	0.00424	感情	0.00490	舆论	0.00672	男方	0.00656
外卖	0.00414	世界	0.00459	通报	0.00641	评论	0.00650
女性	0.00411	男女	0.00457	流量	0.00609	警方	0.00624
感情	0.00402	姐姐	0.00456	央视	0.00510	通告	0.00599
发胖	0.00393	喜欢	0.00434	暴力	0.00509	网络	0.00559
发声	0.00383	对方	0.00384	网民	0.00506	差额	0.00558

表 1 显示, 前期, “姐姐”、“女人”、“男人”、“女性”、“感情”、“恋爱”、“社会”和“恋爱”这 8 个特征词在微博和 Bilibili 均有出现。涉及男女话题讨论量庞大, 两性舆情作为比较敏感的话题

迅速引发网络热议。在前 15 个特征词中, 微博用户讨论内容涵盖多方面, 质疑“姐姐”是否在吸引流量、借机敛财(证明、死亡, 13.07%), 关注性别话题(女人、男人、女性, 25.15%), 讨论“胖猫”感情(恋爱、爱情、感情, 17.49%), 明星参与(樱子、发声, 11.04%)外卖空包现象(外卖, 5.66%)。Bilibili 用户则相对集中在两性话题(女人、男人、女性、男性, 44.01%)和情感方面(恋爱、爱情、结婚、纯爱、感情、喜欢, 30.77%)。

后期, “网络”、“姐姐”、“警方”、“恋爱”、“通报”和“网民”是微博和 Bilibili 上相同的 6 个特征词。两个平台都关注了警方对事件的通报(公布、细节、官方、通报、警方、通告), 此外, 微博上出现蹭热度和吸引流量的行为(谣言、账号、舆论、流量)。相比之下, Bilibili 持续了对情感问题的关注(恋爱、女方、男方); 特征词更具体, 对事件中经济细节的关注(转账、聊天记录、差额), 以及对警方对事件的通报(官方、通报、警方、通告)。

特征词显示, 微博用户讨论更分散, Bilibili 用户相对集中。通过分阶段词云图, 更直观揭示两个平台用户在不同时期关注主题, 如图 2 所示。词越大, 表示其出现频率越高, 用户对该话题的关注度越高。



Figure 2. Word cloud  
图 2. 词云图

### 3.2. 热点话题分析

参考格里菲斯(Griffiths)实验的参数设置方法, 首先将先验参数  $\alpha$  和  $\beta$  设置为 0.01。在图 3 中, 随着

K 的增大, 困惑度逐渐减小, 当  $K = 5$  和  $6$  时, 困惑度下降趋势逐渐开始平缓。通过实验和参数分析结果, 当超参数  $\alpha$  设置为  $0.1$ ,  $\beta$  设置为  $0.01$ , 主题数  $K = 5$  时, 主题聚类效果相对较好。“胖猫事件”主题聚类表详见表 2、表 3。

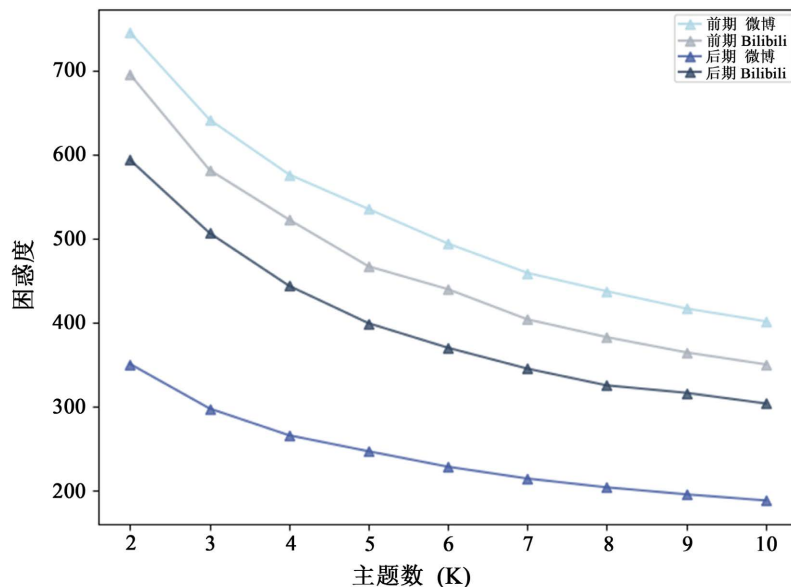


Figure 3. Perplexity  
图 3. 困惑度

Table 2. Preliminary topics  
表 2. 前期主题

主题	占比	微博		Bilibili	
		特征词(Top 15)	占比	特征词(Top 15)	占比
爱情与自我	24.28%	女人、男人、真心、世界、生命、爱人、爱情、爱你、女朋友、感情、下辈子、时间、生活、花钱、游戏	25.05%	感情、父母、女人、家庭、爱情、对方、朋友、社会、家人、经历、爱人、爱你、对象、男人、生活	
流量与热度	23.83%	姐姐、证明、男子、行政处罚、博主、弟弟、网民、证据、聊天记录、律师、流量、照片、警方、网络、直播	17.69%	姐姐、女方、诈骗、聊天记录、男方、官方、评论、流量、网络、法律、时间、警方、观点、舆论、通报	
性别话题	22.92%	樱子、女性、女人、热度、性别、男性、男女、评论、舆论、网民、世界、受害者、女权、中国、流量	22.51%	女性、社会、男性、男女、中国、婚姻、爱情、男人、教育、性别、资本、国家、矛盾、彩礼、女人	
事件感悟	17.84%	社会、感情、爱情、问题、对方、关系、家庭、男人、女方、情感、网络、生活、情绪、价值、生命	16.80%	女人、男人、中国、世界、颜值、花钱、人家、标准、价值、社会、爱情、唐诗、宋词、定义、拜金	
外卖事件	11.12%	网民、空包、重庆、麦当劳、茶百道、商家、浪费、曝光、女友、中国、门店、游戏、品牌、悼念、订单	17.94%	浪费、版本、战士、世界、评论、同情、贡品、食物、麦当劳、网民、热度、受害者、缅怀、男女、重庆	

在“胖猫事件”前期, 微博和 Bilibili 用户的讨论可以归纳为五个主题: 爱情与自我、流量与热度、性别话题、事件感悟和外卖事件。

表 2, 对“胖猫事件”的讨论涉及情感与性别的话题比例几乎占到一半。爱情与自我是两个平台共同关注且讨论量最大的话题, 存在三种观点: 1) 声讨谭某欺骗感情; 2) 心疼“胖猫”遭遇; 3) “胖猫”咎由自取, 不值得同情。此外, Bilibili 还探讨了“爱情”与“家庭”的关系, 指出离异家庭可能造成“胖猫”性格的原因之一。性别话题中, 微博讨论围绕“女性”和“女权”议题展开, 以及与“舆论”、“流量”和“热度”的联系; Bilibili 的讨论更深入, 探讨了性别问题在“社会”、“婚姻”和“资本”中的表现。

流量与热度中, 微博和 Bilibili 都包含“姐姐”、“聊天记录”、“流量”、“网络”和“警方”。微博上存在吸引流量的现象, 关于流量和热度的讨论集中在具体人和事上, 如: 冒充男女双方账号蹭热度等。Bilibili 话题则是对舆论的质疑更多, 如“诈骗”、“官方”、“法律”、“舆论”和“通报”等。事件感悟话题, 微博和 Bilibili 都涉及“男人”、“女人”、“社会”、“价值”、“爱情”等词, 反映出对性别角色以及事件背后的社会现象和价值观的关注。微博讨论更多围绕男女间的情感和家庭关系展开, “感情”、“家庭”和“价值”表明用户关注如何平衡现实生活中情感与生活压力; Bilibili 则关注社会对“女人”和“男人”的标准及文化内涵, “唐诗宋词”和“拜金”反映大家对现代社会中爱情、价值观和择偶标准的思考, 强调经济实力在现实生活中的重要性, 而非单纯的文化追求。外卖事件话题中, 微博用户聚焦于“空包”的问题和对商家的曝光和谴责, 如“麦当劳”、“茶百道”, 也有不少人表达了对“胖猫”的“悼念”; Bilibili 用户则高度关注外卖“浪费”的问题。

Table 3. Subsequent topics

表 3. 后期主题

微博			Bilibili		
主题	占比	特征词(Top 15)	主题	占比	特征词(Top 15)
金额细节	31.66%	姐姐、花店、股东、出资、细节、女人、男人、家庭、弟弟、女方、家人、父母、舆论、原生、网民	金额细节	18.98%	女方、男方、差额、姐姐、聊天记录、通报、双方、警方、花店、通告、诈骗、生活、时间、父亲、金额
网络舆论	10.05%	网络、谣言、细节、账号、流量、央视、暴力、人肉、姐姐、舆论、误导、新闻、平台、警方、信息	网络舆论	25.37%	评论、网络、网民、官方、通报、舆论、站队、节奏、流量、互联网、热度、新闻、信息、媒体、冲锋
道德与法律	18.85%	社会、法律、女性、警方、道德、经济、性别、细节、男性、男人、官方、舆论、女人、男女、情感	道德与法律	19.62%	姐姐、聊天记录、法律、道德、违法、官方、警方、诈骗、通报、事实、疫苗、犯法、舆论、潘金莲、合法
真相反应	24.53%	网民、网络、细节、舆论、官方、警方、通告、重庆、通报、评论、流量、微信、公告、账号、言论	现实情感	18.26%	女人、男人、社会、感情、爱情、生命、法律、金钱、情绪、花钱、世界、利益、网络、生活、彩礼
通报细节	14.91%	警方、通报、姐姐、事件调查、重庆、细节、诈骗、调查、网民、男子、舆论、事实、双方、犯罪、财物	“胖猫”家庭	17.78%	姐姐、家人、家庭、网民、弟弟、原生、女人、女友、父母、吃一堑、家里人、血馒头、原因、感情、死者



警方通报后两平台讨论焦点出现了差异。微博和 Bilibili 用户在金额细节、道德与法律和网络舆论这三个主题上仍表现出相似的关注,但在另外两个主题上则出现了显著差异。

金额细节话题,两个平台都关注金钱和家庭的交织(“家庭”、“父母”、“父亲”),“胖猫”和谭竹的金钱关系(“花店”、“股东”、“出资”)。Bilibili 主要讨论了“女方”和“男方”双方的“差额”,具体财务差异,以及其中是否涉嫌“诈骗”。网络舆论话题,微博讨论集中在“网络”、“谣言”、“细节”、“账号”、“流量”等,强调了对信息误导、暴力、人肉搜索等负面影响的关注,更侧重于网络暴力;Bilibili 特征词包括“评论”、“网民”、“站队”、“节奏”和“媒体”等,聚焦于用户互动、网络舆论的整体情况和媒体报道影响。道德与法律话题,“法律”、“道德”、“官方”、“警方”、和“舆论”在两个平台上共同出现,表明用户讨论的核心要素涉及法律问题、道德讨论、官方声明、警方介入以及舆论发展。微博侧重讨论社会现象和性别关系中的道德和法律问题;Bilibili 更关注具体案例、违法行为以及官方通报。

不同的是,微博用户侧重事件的社会影响和舆论情况,讨论了在真相公布后“网民”关于“网络”“舆论”的态度变化(真相反应),对“警方”“通报”“细节”的关注(通报细节)。Bilibili 用户注重事件对现实生活的反思,延续了前期对现实情感问题和“胖猫”家庭影响的关注。

### 3.3. 情感分析

首先,从整体上比较微博和 Bilibili 用户关于“胖猫事件”前、后期的情感倾向,发现总体上以负面的态度为主,微博的负面情感比例均高于 Bilibili,相差约 9%,见图 4。事件细节的公布并没有降低网民的负面态度。舆情和事实之间的偏差,导致在真相发布后两个平台的负面情感显著增加[19],微博和 Bilibili 的负面情感分别增加了 7.3%和 6.42%,中性情感在两个时期都相对稳定,正面情感在后期都有不同程度的下降。

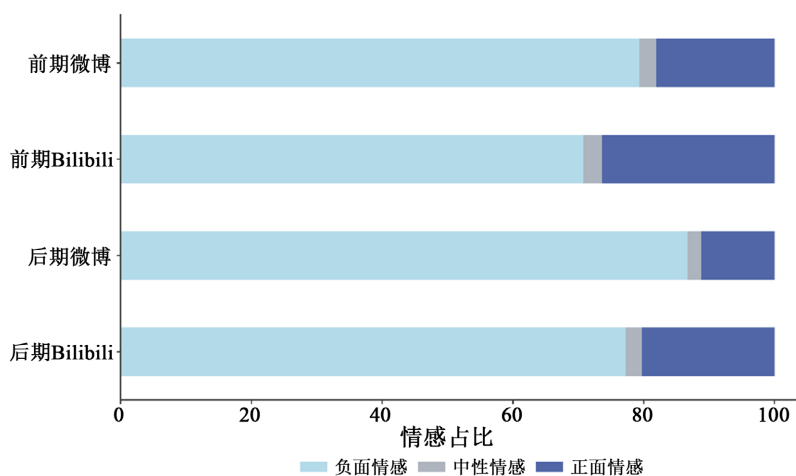


Figure 4. Percentage of emotions in different periods

图 4. 不同时期情感占比

对前期和后期 10 个主题中出现的 8 个相同主题的情感进行比较分析,前期 5 个主题,“爱情与自我”、“流量与热度”、“性别话题”、“事件感悟”、“外卖事件”,后期 3 个主题,“金额细节”、“道德与法律”和“网络舆论”,如图 5。

在所有相同主题中,中立情感的比例普遍较低,说明用户在这些话题上更倾向于表达明确的情感态度,而不是保持中立。微博的负面情感比例普遍超过 Bilibili,其中,“爱情与自我”和“金额细节”这

两个话题基于事件中的具体事实,具有一定的共识基础,情感差距较小。“流量与热度”是微博和 Bilibili 负面情感最高的话题,这说明两个平台上的网民都对网络谣言、蹭热度行为表现出不满。公布警情通报后,微博的“网络舆论”主题负面情感达到了 97%, 主要是因为网民知道真相后,对微博上存在的虚假信息和盲目跟风的现​​象表示反感。微博以文本为主,发布门槛低,用户和营销号在热点事件出现时,能够通过迅速发布简单的图文蹭热度。即时且情绪化的表达方式放大了用户的情绪波动,容易激发其对流量的抵触情绪,形成负面情感的集聚效应。相对而言,由于视频的制作成本较高, B 站的蹭流量现象相对较少,用户多针对问题在发布的视频下开展讨论,后期“网络舆论”的负面情绪相对较低。

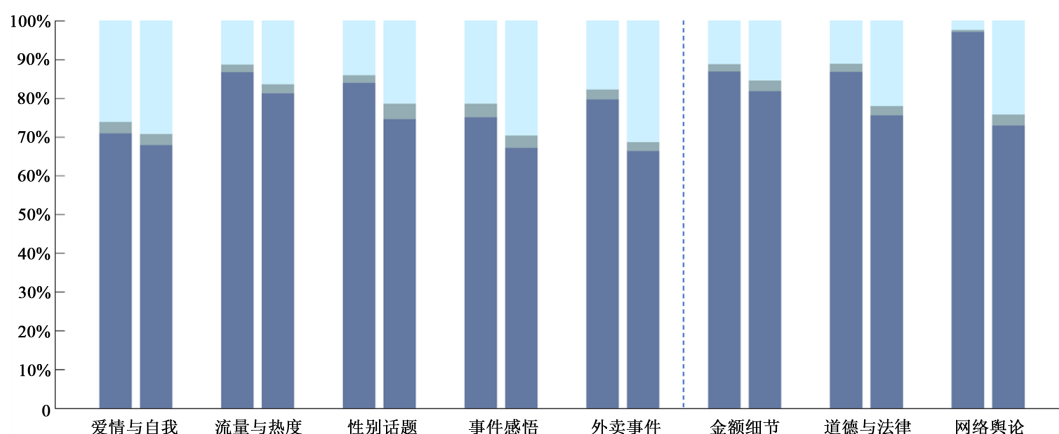


Figure 5. Comparison of thematic sentiment between Weibo (left) and Bilibili (right)

图 5. 微博(左)和 Bilibili(右)主题情感对比

#### 4. 讨论

互联网使得舆情事件在各社媒平台上迅速发酵。当热点事件发生,不同类型平台都会成为舆情发酵的主要场所,但舆论焦点和情感表现会有所不同。通过对“胖猫事件”微博帖子 and Bilibili 评论的特征词提取、舆情主题分析和情感计算,得出以下结论。

(1) 从主题内容来看,特征词表明微博用户讨论更分散, Bilibili 用户相对集中。主题聚类结果表明微博和 Bilibili 用户在“胖猫事件”中的讨论既存在相似之处,前期的讨论主要聚焦在情感与性别的话题。后期真相揭露后,“金额细节”、“道德与法律”、“网络舆论”三个主题也是相同的,另外两个主题则呈现出不同的内容。此外,研究发现微博蹭流量行为的频率更高。这是由于微博上文本内容的制作成本较低,用户容易利用热点事件来吸引流量。而 Bilibili 用户利用热点事件蹭流量的难度相对较大,大部分用户倾向于围绕事件本身进行讨论,评论较集中,也会透过事件去讨论更深层的社会原因。

(2) 从情感结果来看,总体上微博和 Bilibili 用户中立情感的比例普遍较低,均表现出明显的情绪极化,负面态度占据主导地位;且在真相公布后负面情感均上升。在各个话题上,微博的都呈现出更高的负面情绪,但是关于事实的话题,如“爱情与自我”“金额细节”等两个平台的情绪比较接近,显示出网友认知的一致性。而“网络舆论”的话题情绪差距最大,微博用户更多因流量导向和蹭热度行为产生抵触情绪,负面情绪的表达更集中。

微博的讨论门槛较低,通过简短的文字和图片即可快速分享信息,因此用户通常会迅速发表观点,拥有较强的话题延展性,容易出现广泛的互动和传播和情感化的表达。相比之下, Bilibili 用户更多地通过视频、弹幕和评论进行互动,讨论通常围绕事件的具体问题展开,探讨事件背后的社会原因和影响,讨论和交流的内容更集中。通过对比分析微博和 Bilibili 上关于“胖猫事件”讨论的话题类型、讨论态度、

情感表达, 可以揭示不同平台用户在舆论焦点、情感表达以及用户互动方式上的差异。不仅有助于更好地理解不同社交媒体平台的舆情特征, 还能为社会媒体管理和舆情监控提供方法和策略, 以更好地应对和引导公众舆论, 促进社会的理性讨论和进步。

## 5. 总结

本研究对比了以文本内容为主的微博和视频平台 Bilibili 在“肥猫事件”中舆情的异同。基于用户生产数据, 从内容和情感两个维度剖析网络舆情, 揭示了这两个平台用户在讨论焦点和情感表现上的显著差异, 提供具有实证性的分析结果。通过文本挖掘方法对两类社交媒体的舆情进行了定量分析, 为我们理解不同类型平台用户观点态度和情感反应提供了重要的量化依据。未来研究将进一步探讨导致差异出现的原因, 深入分析平台环境、平台用户特性、平台推送机制、内容属性和信息传播模式等因素如何影响用户的反应。这将有助于我们更全面地理解不同类型平台的用户行为, 为控制舆情和管理用户情感等提供理论支持和实践指导。

## 基金项目

国家社会科学基金一般项目: 新冠疫情中“信息疫情”的形成机理及应对策略研究(22BGL240)。

## 参考文献

- [1] Alfarrarjeh, A., Agrawal, S., Kim, S.H. and Shahabi, C. (2017). Geo-Spatial Multimedia Sentiment Analysis in Disasters. 2017 *IEEE International Conference on Data Science and Advanced Analytics*, Tokyo, 19-21 October 2017, 193-202. <https://doi.org/10.1109/dsaa.2017.77>
- [2] Md Suhaimin, M.S., Ahmad Hijazi, M.H., Moug, E.G., Nohuddin, P.N.E., Chua, S. and Coenen, F. (2023) Social Media Sentiment Analysis and Opinion Mining in Public Security: Taxonomy, Trend Analysis, Issues and Future Directions. *Journal of King Saud University—Computer and Information Sciences*, **35**, Article 101776. <https://doi.org/10.1016/j.jksuci.2023.101776>
- [3] Bai, H. and Yu, G. (2016) A Weibo-Based Approach to Disaster Informatics: Incidents Monitor in Post-Disaster Situation via Weibo Text Negative Sentiment Analysis. *Natural Hazards*, **83**, 1177-1196. <https://doi.org/10.1007/s11069-016-2370-5>
- [4] 王冲, 张雅君, 王娟. 社会大众如何看待生成式人工智能在教育中的应用?——对 B 站 ChatGPT 话题弹幕文本的舆情主题与情感分析[J]. 图书馆论坛, 2024(5): 1-12.
- [5] Toussaint, P.A., Renner, M., Lins, S., Thiebes, S. and Sunyaev, A. (2022) Direct-to-Consumer Genetic Testing on Social Media: Topic Modeling and Sentiment Analysis of Youtube Users' Comments. *JMIR Infodemiology*, **2**, e38749. <https://doi.org/10.2196/38749>
- [6] Chandrasekaran, R., Konaraddi, K., Sharma, S.S. and Moustakas, E. (2024) Text-Mining and Video Analytics of COVID-19 Narratives Shared by Patients on Youtube. *Journal of Medical Systems*, **48**, Article No. 21. <https://doi.org/10.1007/s10916-024-02047-1>
- [7] 薛海平, 张小莉. “双减”政策执行中的网络舆情监测研究[J]. 教育发展研究, 2024, 44(8): 50-60.
- [8] 罗怡然. 基于“小红书滤镜景点舆情事件”分析媒介景观的建构与反叛[J]. 科技传播, 2022, 14(6): 111-114.
- [9] 刘威, 侯林林. “拟剧理论”下用户使用微博与微信的对比研究[J]. 新闻知识, 2019(7): 53-56.
- [10] Wang, Z., Huang, W. and Liu-Lastres, B. (2022) Impact of User-Generated Travel Posts on Travel Decisions: A Comparative Study on Weibo and Xiaohongshu. *Annals of Tourism Research Empirical Insights*, **3**, Article 100064. <https://doi.org/10.1016/j.annale.2022.100064>
- [11] 胡媛, 蒋天森, 古林鑫, 等. 跨平台社交媒体用户生成内容差异研究[J]. 情报科学, 2024, 42(4): 69-78.
- [12] 李萌. 不同社交平台用户评论行为的对比研究[D]: [硕士学位论文]. 武汉: 武汉大学, 2020.
- [13] 王艺静, 钱清. 基于微博舆情事件的主题情感分析——以“H&M 禁用新疆棉”为例[J]. 新闻研究导刊, 2022, 13(5): 37-39.
- [14] 张晓君, 黄微, 宋先智. 基于心理距离的网络舆情传播倾向性研究[J]. 情报资料工作, 2017(6): 19-24.

- 
- [15] Hu, Z., Dai, H., Yang, G., Yi, X. and Sheng, W. (2022) Semantic-Based Multi-Keyword Ranked Search Schemes over Encrypted Cloud Data. *Security and Communication Networks*, **2022**, 1-15. <https://doi.org/10.1155/2022/4478618>
- [16] Kayest, M. and Jain, S.K. (2022) Optimization Driven Cluster Based Indexing and Matching for the Document Retrieval. *Journal of King Saud University—Computer and Information Sciences*, **34**, 851-861. <https://doi.org/10.1016/j.jksuci.2019.02.012>
- [17] Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) Latent Dirichlet Allocation. *Journal of Machine Learning Research*, **3**, 993-1022.
- [18] Ying, S. (2024) Guests' Aesthetic Experience with Lifestyle Hotels: An Application of LDA Topic Modelling Analysis. *Heliyon*, **10**, e35894. <https://doi.org/10.1016/j.heliyon.2024.e35894>
- [19] 刘焕. 公共事件网络舆情偏差及影响因素研究述评[J]. 情报杂志, 2018, 37(11): 96-102.