

· 研究论文 ·

基于时空视角的舆情反转事件情感演化特征研究

何婕君 李 阳

(南京大学信息管理学院,南京,210023)

[摘 要] 舆情反转事件是一类特殊的舆情事件,研究舆情反转事件的情感演化特征,对复杂舆情事件的综合治理以及网络生态稳定具有重要意义。通过划分特定舆情反转事件的生命周期,结合事件主题内容特征进行情感分析,从时间和空间两个维度刻画舆情反转事件的情感演化特征。结果表明,网民在舆情反转事件各生命周期的情感焦点、情感倾向呈现出复杂的演化态势,不同时间段、不同区域的舆情关注度和情绪状态存在差异。与普通舆情事件相比,舆情反转事件在反转前后的情感变化具有鲜明的特殊性,其情感演化过程与舆情反转事件发展态势密切相关。

[关键词] 舆情反转 微博 主题分析 情感分析 时空演化

[中图分类号] G250 [文献标识码] A [文章编号] 2095-2171(2022)02-0088-13

DOI: 10.13365/j.jirm.2022.02.088

Research on the Emotional Evolution Characteristics of Public Opinion Reversal Events Based on Time and Space Perspective

He Jiejun Li Yang

(School of Information Management, Nanjing University, Nanjing 210023)

[Abstract] Public opinion reversal events are a special kind of public opinion events. Studying the emotional evolution characteristics of public opinion reversal events is of great significance to the comprehensive management of complex public opinion events and the stabilization of the network ecology. By dividing the life cycle of a specific public opinion reversal event and combining with the subject content characteristics, we carried out the sentiment analysis to describe the emotional evolution features of the public opinion reversal event from time dimension and the space dimension. The results show that the emotional focus and emotional tendency of netizens in each life cycle of public opinion reversal events show a complex evolutionary trend, and there are differences in public opinion attention and emotional state in different time periods and regions. Compared with ordinary public opinion events, the emotional changes of public opinion reversal events before and after the reversal have distinct characteristics, and the emotional evolution process is closely related to the development trend of public opinion reversal events.

[Keywords] Public opinion reversal; Microblog; Topic analysis; Sentiment analysis; Temporal and spatial evolution

[基金项目] 本文系国家自然科学基金青年项目“面向应急管理的情报工程服务机制建设研究”(18CTQ021)的成果之一。

[作者简介] 何婕君,硕士生,研究方向为社交媒体信息分析、数据挖掘;李阳(通讯作者),副教授,硕士生导师,研究方向为应急情报、数据治理,Email: liyang@nju.edu.cn。

本文引用格式:何婕君,李阳.基于时空视角的舆情反转事件情感演化特征研究[J].信息资源管理学报,2022,12(2):88-100.

1 引言

舆情反转事件是一类特殊的舆情事件,近年来,各类舆情反转事件频繁发生,如重庆万州公交车坠江女司机事件、王凤雅小朋友事件、山东环卫工被顶替教师岗位事件等,引发网络舆论的高度关注和热议。舆情反转事件的显著特征就是舆情走向呈现波浪式特征,公众对事件的舆情表达与认知会随着时间推移而发生不断变化,这个过程中公众的情绪观点、情感态度往往会发生戏剧性的逆转或偏移,使公众容易感到混乱,滋生惊讶、被欺骗、被愚弄等情绪。

舆情反转事件的发生有其背后的原因,一方面是舆情反转事件本身具有复杂性,事件“真相大白”往往比较缓慢,而新闻媒体和社会公众获取真实信息的来源较少,“后真相化”导致渠道信息的可靠性无从考证;另一方面则是“抢流量”效应的存在,新媒体时代信息时效性变得极其重要,一些信息发布者为了抢占信息发布先机,往往在信息未经证实时就发布不确定信息,甚至在信息发布时带有明显的个人或机构情感倾向,对公众和社会产生误导。对舆情反转这类现象如果处置不当,容易扰乱公众对舆情事件及其相关议题的理性判断,短时间内催生和加深社会消极负面情绪,舆情严重时还可能从虚拟空间的情绪积累上升为社会运动。目前来看,学术界、舆情分析机构等对舆情反转这一特殊事件类型的关注度不够,而“反转”这一新现象、新特征在社会矛盾交织、“流量为王”、大众情感需求复杂化等环境下变得更加凸显,因此,积极探索舆情反转事件相关原因尤其是情感演化特征具有重要意义。

在舆情反转事件发生过程中,网络空间在掀起一场舆论风暴的同时也留下了反转前、反转中、反转后的数据痕迹。用户在各类社交媒体上的行为活动反映了他们对舆情反转事件的看法或在事件中的真实需求,而利用这些舆情数据可以帮助我们了解大众的情绪行为模式,为舆情反转现象的分析和应对策略的制定提供支持,进而有效监测网络空间活动和现实环境的动态变化。为此,本

研究旨在探索舆情反转事件情境下的情感演化特征与规律,通过利用网络爬虫获取特定案例事件在社交媒体平台中的数据信息,借助情感分析等方法对舆情反转事件各阶段的网民情感特征进行多维度分析,进而讨论舆情反转事件的网络舆情态势演变特征。本文的研究意义在于通过主题与情感融合、时间与空间结合的视角展现真实舆情反转事件中网民情感变化过程,加深对舆情反转事件情感演化所体现出的网络生态和社会矛盾问题的理解,为舆情反转事件的治理路径和应对策略提供参考。

2 相关研究回顾

2.1 舆情反转相关研究现状

国内对舆情反转相关问题的研究主要集中在新闻传播学、情报学等领域,学者们围绕舆情反转的成因、传播规律、演化机制、预测监督等方面展开探索。主要可划分为三个方向:

第一,舆情反转事件特征及成因。由于舆情反转事件的特殊性,有学者对其特征进行了总结,如谭玲珑^[1]认为舆情反转具有舆情走向的戏剧性、公众参与的广泛性、网民表达的情绪化等;倪万^[2]认为,舆情反转现象反映出信息传播中时间和空间上的不对称;田文利^[3]认为舆情反转根本原因在于虚假信息的发展和传播,相关部门、媒体、公众都应负有责任;Chen等^[4]的研究确定了个人内部特征和外部干预信息是影响舆论逆转的内部和外部因素;汪明艳等^[5]研究发现,群体认同、选择性接触、群际情绪等是影响舆论反转事件中网民群体行为意愿的重要因素。

第二,舆情反转传播演化机制。观察舆情反转过程不仅可以探索其传播过程的形态类型^[6],还可以结合相关理论分析舆情反转的演化机制。王玉龙^[7]发现群体极化通过“倾向携带”“协同过滤”“过程凝聚”隐性操控着舆情反转;岳甜等^[8]、张敏等^[9]均以“罗一笑事件”为实证案例,分别构建了反转新闻的舆论演变机制和情绪反转形成机理的理论模型;张兆曙^[10]从信息控制主体和舆情推动主体双向行动过程解释了网络舆论反转的

社会机制；关琮严^[11]从媒体和民众的角度认为，叙事机制、群体思维社会心理机制以及话语权竞争机制是舆情反转的三大机制。

第三，舆情反转演化仿真模型。部分学者利用模型进行仿真实验，揭示舆情反转的规律和成因。如陈一新等^[12]利用改进的Hegselmann-Krause模型，通过意见领袖观点值改变刻画舆论反转；Jiang等^[13]基于流行病模型提出了两阶段的SPNR舆论反转动态模型。还有学者通过模型评估舆情反转效应，助力舆情反转的预测监督。如袁野等^[14]借助HMM模型进行分类预测，得出四类反转网络舆情及其性质特征和应对策略。夏一雪等^[15]基于微分方程和信息生命周期理论，构建网络舆情反转机理模型进行反转舆情预测；王楠等^[16]借助改进的SMOTE算法和机器学习，构建神经网络模型对舆情反转事件进行预测。

2.2 舆情情感分析相关研究

情感分析是对带有情感色彩的文本进行分析和处理，以快速、准确获取文本数据中的情感态度特征，该方法也被广泛应用于舆情分析、评论分析、内容推荐等领域^[17-18]。目前，舆情情感分析研究可以按照所采用的情感分析方法划分为基于情感词典的研究、基于传统机器学习的研究、基于深度学习方法的研究等。

其中情感词典的方法相对简单且容易理解，只需要事先构建好情感词典，匹配文本语料中的情感词再通过加权计算，就可以判断文本的情感倾向性^[19]，这类方法在舆情情感态势研究中被广泛采用。如王丹等^[20]结合超网络和动态网络方法去识别舆情中的关键节点，借助知网词典Hownet判定关键节点的情感倾向，为舆情思想引领策略选取提供参考；杜毅贤等^[21]采用大连理工大学的情感词汇本体库^[22]和Hownet判定舆情情感，并借助交互信息图表来呈现不同平台、不同省份的舆情情感态势。

基于传统机器学习的研究，通过爬取数据，训练模型，再利用模型预测结果，模型训练常用算法有贝叶斯算法、KNN、支持向量机等^[23]。如王晰巍等^[24]构建了移动环境下的朴

素贝叶斯情感分析模型，对词频、地域和时间三个维度的情感演化过程进行研究。邓君等^[25]则采用Word2Vec和SVM方法构建情感分类模型，对微博舆情情感时序特征和舆情主体对象情感演化进行分析。

基于深度学习的研究，将神经网络应用到舆情情感分析中，试图获取效果更优的情感分析结果，国内外也有研究证明卷积神经网络(CNN)^[26]、循环神经网络(RNN)^[27]和递归神经网络模型(ANN)^[28]、双层注意力与Bi-LSTM模型^[29]可以有效挖掘舆情等文本的情感特征，王晰巍等^[30]基于卷积神经网络构建了社交网络舆情用户情感主题分类模型，通过实验对比，也证明该方法较之SVM和Time-LSTM模型分类效果更好。

舆情的情感分析研究往往分析网络受众对具体事件的情感态度演化，还常与主题分析、文本聚类等方法结合使用，来更好地刻画用户情感转变背后的指示性内容。例如，李超雄等^[31]提出了一种动态主题情感混合模型，通过追踪不同时间片的主题与情感变化，实现主题和情感的演化分析；安璐等^[32]以寨卡事件为例，采用情感词典方法结合主题分析方法揭示网络舆情情感演化规律，生成阶段情感演化图；朱晓霞等^[33]也提出，传统静态式的情感倾向判断仅仅局限于情感分类的结果，不能追溯整个情感演化过程及影响因素，为此，作者提出了一种动态主题-情感演化模型的网络舆情信息分析方法，并通过实验研究证明了其有效性。

随着情感分析对细粒度及其背后原因分析的需求逐渐提升，建立“话题”与“情感”之间的联系，已经成为舆情情感分析领域的重要前沿方向。

2.3 评述小结

基于上述分析可以发现，第一，目前虽然一些研究已经关注到舆情反转事件的特征、成因、演化机制、预测模型等内容，但基于相关事件社交媒体数据的价值挖掘比较少见，对于其深层次的情感特征变化揭示也缺乏应有的关注，而大众在舆情反转这一特殊事件情境下，其情感往往呈现出复杂的特

征,因此有必要深入探究。本研究拟基于舆情反转事件相关社交媒体用户数据痕迹,通过多维视角来解释舆情反转事件的情感演化特征。第二,在舆情情感分析方面,目前主要采用基于情感词典、基于传统机器学习、基于深度学习等研究方法,在不同情境下相关方法需要进行合适性的选择。与此同时,将主题与情感进行融合分析,逐渐成为舆情情感态势分析的前沿方向,对于深入挖掘情感结果及其背后的影响因素与特征具有重要意义。本研究拟以舆情反转事件作为研究对象,通过主题与情感的协同分析,从时间与

空间双重视角来刻画相关情感演化规律,以更加全面和真实地反映事件情感倾向的变化及其原因。

3 舆情反转事件情感演化分析框架构建

3.1 研究框架设计

本研究以微博平台的舆情反转事件为例,参考安璐等^[32]、张琛等^[34]相关舆情情感分析研究工作,构建舆情反转事件情感演化探索的分析框架,总体研究思路如图1所示,包含数据采集与预处理、舆情周期划分与主题聚类、情感变化时空分析三个部分。

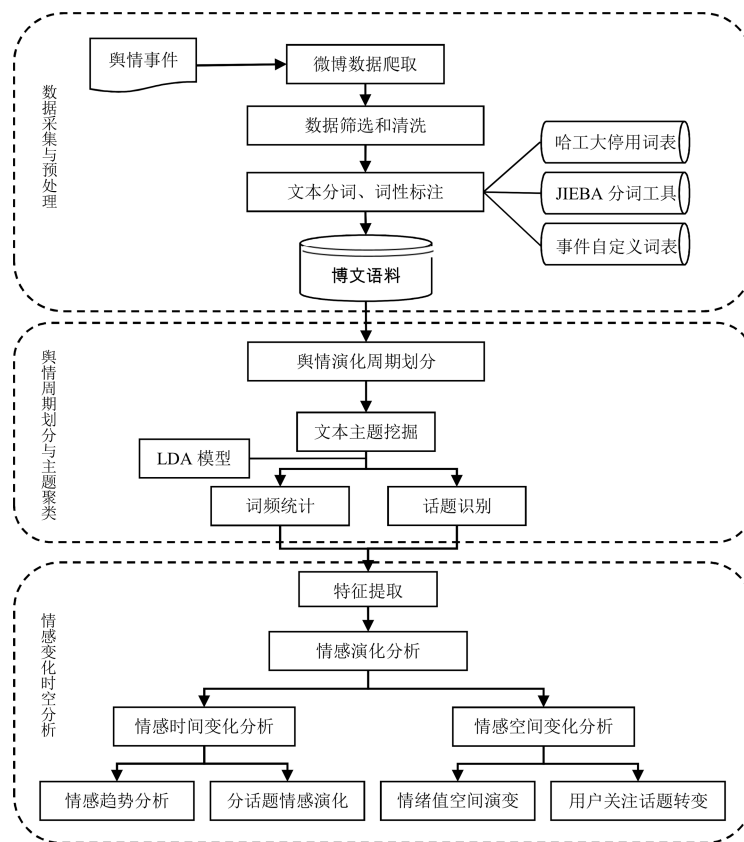


图1 舆情反转情感演化研究框架

第一部分,获取某一舆情反转事件的博文及评论数据。根据博文数量高峰个数划分舆情周期。

第二部分,采用LDA主题聚类方法对舆情演化周期各阶段的博文进行主题聚类分析。

第三部分,借助情感分析方法,从时间维和空间维两个角度分析舆情反转事件各阶段的情感特征和演化规律。在时间维分析中,为了

更细粒度提取文本情感,实现多情感分类,本文参考国内较为成熟的大连理工大学情感词汇本体库^[22]用于文本情感分类;同时为了更好地呈现空间维情感演化差异,采用SnowNLP直接计算文本对应情感值。

3.2 研究方法与技术

3.2.1 LDA主题模型

隐含狄利克雷分布,简称LDA(Latent

Dirichlet allocation), 是一种主题模型。该模型由 Blei 等^[35]于 2003 年在概率潜在语义分析 (PLSA) 模型的基础上提出。其假定文档集中的所有文本均共享一定数量的隐含主题, 将整个文档集特征化为隐含主题的集合, LDA 的目的就是为每一篇文档推测其主题分布, 包括计算各个主题出现的概率大小。由于 LDA 模型能有效挖掘文档中所包含的主题信息, 因而被广泛应用于信息检索、文本分类、评论主题词抽取等领域, 也是网络舆情主题识别的热门研究方法^[36]。

LDA 模型原理是将每篇文档被表示为 K 个隐含主题的混合分布, 每个主题是在 W 个词语上的多项分布, ϕ 表示 LDA 中主题-词语的概率分布, θ 表示文档-主题的概率分布, α 和 β 分别表示 ϕ 和 θ 所服从的 Dirichlet 先验分布的超参数 (见图 2)。

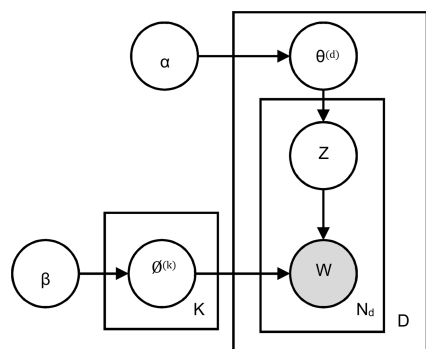


图 2 LDA 概率图表示

3.2.2 情感分类方法

就文本粒度而言, 情感分类可以实现文档级、句子级不同粗细层次的分析, 基于句子的情感分类就是识别粒度较细但具有整体意义的句子的情感倾向。本研究采用的文本语料情感分类基本流程如图 3 所示。

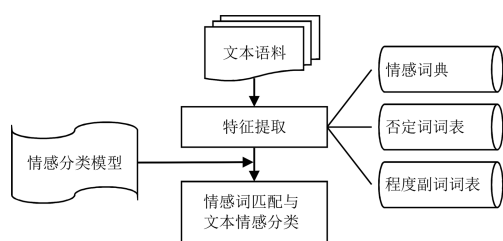


图 3 情感分类流程

通过提取的特征计算每类情绪的情感强

度, 将强度最大的情绪类别作为单条微博文本的情绪类别。本文选择大连理工大学的情感词汇本体库^[22]作为情感词典, 该词典将情感分为“乐”“好”“怒”“哀”“惧”“恶”“惊”7 个大类和 21 个小类, 其情感词的初始情感强度被设置为 1、3、5、7、9 五档, 9 表示强度最大, 1 为强度最小。每个词在每一类情感下都对应了一个极性。其中, 0 代表中性, 1 代表褒义, 2 代表贬义, 3 代表兼有褒贬两性。词汇的情感值公式为:

$$s(w) = v(w) p(w) \quad (1)$$

式 (1) 中, $s(w)$ 表示词汇的情感值, $v(w)$ 表示词汇的情感强度, $p(w)$ 表示词汇的情感极性。由于否定词和程度副词影响情感判别的结果, 在参照已有否定词词典和程度副词词典的基础上, 本研究补充建立了适用于分析语料的否定词词典和程度副词词典。根据中文语言习惯, 当情感词前后出现奇数个否定词时, 对情感词极性进行偏转, 参考杜振雷^[37]关于否定词的分析, 本研究收录了 33 个否定词, 对 7 类情感进行否定迁移后的情况如表 1 所示。

表 1 否定词修饰后的情感迁移

原始情感类别	否定迁移后的情感类别
乐	哀
好	恶
怒	无
哀	a * 乐
惧	无
恶	b * 喜好
惊	无

注: 其中 a 和 b 是 [0, 1] 的实数, 表示情感强度被减弱, 实验测定 a=0.1, b=0.2 时, 情感识别准确率最高。

此外, 为提高情感特征判别的准确度, 加入程度副词并人工判别对情感词的影响权重值形成程度副词词典。统计发现, 将情感词前后滑动窗口值设为 2 时, 能更准确地判别情感类别。语料中情感词的强度经程度副词修饰后需要重新计算, 即:

$$\text{WordSentiment} = v(w) * v(adv) \quad (2)$$

式中, $v(adv)$ 为程度副词的权重, $v(w)$ 为情感词的强度值。

对每一条文本内容, 计算各类情感对应

情感词经修饰后的强度值，按照类别求和，将情感强度最大的情绪类别作为文本最终的情感分类，以此刻画时间维度视角下的舆情反转事件情感演化规律。

3.2.3 情感值计算

SnowNLP 的情感值计算以朴素贝叶斯分类算法为理论基础，属于生成式分类算法，计算公式为：

$$P(\text{类别} | \text{词}) = \frac{P(\text{词} | \text{类别})P(\text{类别})}{P(\text{词})} \quad (3)$$

计算之前，由于朴素贝叶斯分类易受分类任务影响，需要根据微博语料数据导入正面样本和负面样本重新训练情感分析模型。本文选用了正负各 6000 条情感标注已完成的新浪微博数据集进行重新训练，模型训练好后就可以计算情感值。在数据处理过程中，我们将情感区间从 $[0, 1]$ 转换为 $[-0.5, 0.5]$ ，使得单条微博文本情感值越接近 -0.5 越消极，情感值越接近 0.5 越积极，并借助 SnowNLP 计算得到的情感值来刻画空间视角下的舆情反转事件情感演化规律，力图更加精准反映相关区域的细节变化。

4 舆情反转事件情感演化分析

4.1 舆情反转事件的选取

“苟晶事件”是 2020 年教育热点事件之一，涉及职业操守、教育公平、媒体公信力和体制信任等多方面。2020 年 6 月 22 日，苟晶向山东省教育厅实名举报，表示自己曾在 1997 年和 1998 年山东高考中连续两年被冒名顶替，其中 1997 年顶替者是高三班主任的女儿。2020 年 7 月 3 日 13 时 47 分，人民日报公布官方调查结果^[38]，表明苟晶本人未填报 1997 年志愿，其个人身份及当年的高考成绩被高三班主任邱印林女儿邱小慧冒用，后者在当年以苟晶的身份被北京煤炭工业学校录取。1998 年的高考录取中不存在苟晶被他人冒名顶替上学的问题。苟晶以“学霸”“高考顶替”“退学”“改变命运”“被威胁”等关键词的发声在事件之初获取网友的同感和愤慨，而随着调查结果的公示，网民对苟晶提出质疑，认为其有“故意撒谎煽动群众”“利用舆论玩弄公众”的过错，因此事件真相前后形成了巨

大的舆情偏差。“苟晶事件”作为一个社会性话题，具有强社交舆情、情绪唤起强烈、受关注度高、公众舆情反转等特征，因此本文将其作为实证对象。

4.2 数据收集和预处理

本文选取新浪微博平台，以“苟晶”为关键词，收集 2020 年 6 月 24 日 0 时至 2020 年 7 月 14 日 0 时期间相关微博以及对应的用户信息，采用 Python 撰写代码，构建网页爬虫爬取相关微博数据。数据获取的具体步骤为模拟登录新浪微博，以关键字为“苟晶”进行高级搜索，由于微博的限制，一次搜索的结果最多显示 50 页，因此设定每次搜索时间范围为 2 小时，获取每条微博的用户 id、用户认证类型、用户名、发布时间、对应收藏次数、转发次数、评论人数和点赞人数，以及微博文本内容；再对相关微博数据简单去重和初步清洗后得到 24584 条数据。

同时根据博文数据中的 mid 字段调用微博 API 接口获取评论数据，由于频次限制，获取每条微博的前 50 条评论，包含评论 id、评论对应微博 mid、评论创建时间、评论文本内容、用户 id、用户名、用户省份、用户所在城市、用户所在详细地址、用户简介、用户性别、用户的粉丝数和关注数等 12 个字段，共计 80238 条评论数据，其中位置信息非空的评论数据共计 57264 条，占比 71.4%。

在预处理过程中，采用代码和人工检查的方式对获取数据进行筛选和清洗，并借助 jieba 分词工具和哈工大停用词表完成分词操作，在此过程中，添加用户词典导入分词系统以提升分词效果。

4.3 舆情周期划分

统计相关微博在时间序列上的分布特征，如图 4 所示，“苟晶事件”微博发布趋势出现两个顶峰，呈现出明显的“双峰”特征。根据贾亚敏等^[39]提出的双峰事件网络信息传播规律，对“苟晶事件”的演化周期进行阶段划分，得到舆情起始阶段、爆发阶段、第一次衰退阶段、第二次增长阶段、第二次衰退阶段、平息阶段。

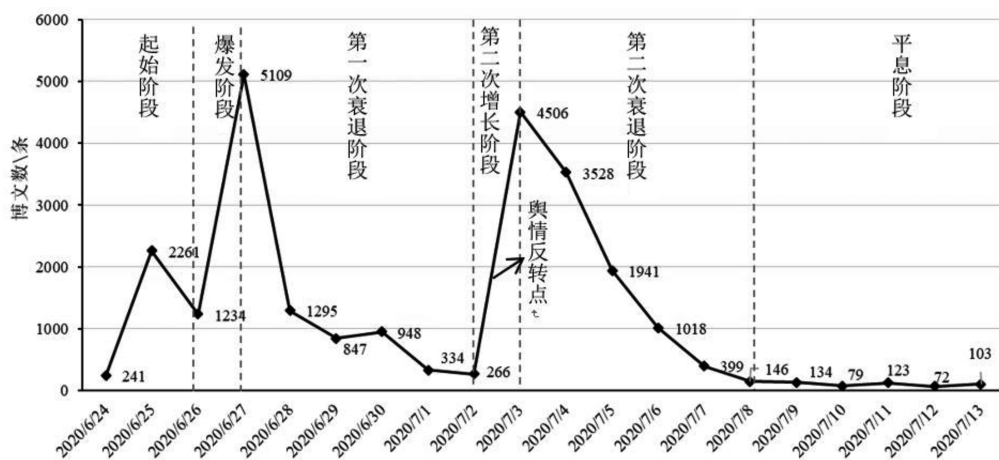


图4 “苟晶”事件微博时序演变图

查看两次舆情高涨前后的博文信息得知，第一次峰值产生的原因是各大新闻媒体对事件的报道，事件同时出现“苟晶八位同学也怀疑自己高考有问题”“高中班主任见苟晶遭拒画面曝光”“苟晶第二次参加的是不是真高考”等话题，引发网民们的强烈关注；第二次峰值产生的原因是2020年7月3日13时47分，人民日报公布官方调查结果并被广泛转发扩散，引发网民对舆情反转真相的热议，这一节点也是该舆情事件的反转点。

4.4 主题演化结果分析

利用LDA主题聚类方法对各阶段微博主题进行建模，抽取不同传播周期各话题下出现频率较高的词作为主题特征词，得到舆情反转事件不同阶段的话题概括，如表2所示。

(1) 起始阶段。话题主要围绕事件参与者、涉事部门展开报道和讨论，南方周末等新闻媒体相继转发跟进“苟晶事件”，包括实名举报、当事人回应与行动举措、涉事校方回应等，网民就此发表讨论，引发相关部门重视并介入。

(2) 爆发阶段。事件进一步曝光，引发对社会公平、高考制度等的热烈讨论，舆论不断发酵，网民在微博上的情感宣泄增多。

(3) 第一次衰退阶段。主题延续了对社会正义和高考公平性的讨论，人民日报声明官方将严查事件真相，绝不姑息腐败黑暗操作，相关部门单位组成调查工作专班联合校方等深入调查，力图还原事件真相；而此时苟晶删帖，班主任道歉，事件热度逐渐衰退。

(4) 第二次增长阶段。该时间段内，由于山东省教育厅回应介入调查，通报了初步调查结果，2020年7月2日下午公布核实调查结果并对涉事人员进行处罚，对比调查结果与当事人言论，关于苟晶到底是受害者还是造谣者的讨论使舆论再攀高峰。

(5) 第二次衰退阶段。主题包括高考公平，指责苟晶夸大事实，批评其带来的恶劣影响等方面。这一阶段出现“舆论世界的真相被情绪左右，只看到自己想看到的一半”“法纪不张和道德沦丧哪个更可怕”等辩论话题，引发热议。

(6) 平息阶段。这一阶段微博数量较少，主要围绕事件真相公布，网民反感情绪宣泄、道德法律层面的争论与分析等话题展开，公众冷静思考、理智发言。

根据上述主题聚类结果可知，在舆情反转事件起始阶段话题主要为事件参与者（苟晶、班主任、当地招生办、校方）、事件类型（教育公平）、地域（山东省）等相关信息的传播扩散。爆发阶段的网民则倾向于事件性质（Topic II -3、Topic II -4）的探讨与情感的宣泄（Topic II -2），舆情高涨达到顶点后逐渐回落。而由于反转事件具备特有的舆情反转点，随着人民日报公布调查结果（Topic IV -1、Topic IV -2、Topic IV -3）使得现有舆情内容发生改变，公众认知产生冲突，导致舆情再次攀升，转向对当事人（Topic V -1）和事件本质（Topic IV -4）的探讨，民众关注点延续的同时

表 2 “荀晶”事件不同发展阶段主题分布

舆情周期	主题编号	主题特征词(部分)	主题概括
萌芽阶段	Topic I -1	命运、教育、改变、受害者、道歉、正义、权利、威胁、原谅、农家女、疑遭、无奈、可怕、欺负	网民情感宣泄
	Topic I -2	疑点、学籍、录取、曝光、道歉、操作、发现、复读、持续、恐怖、可恶	事件讨论
	Topic I -3	调查组、济宁、回应、教育、发帖、前往、年前、农家女、发帖人、希望、人生、道歉信、联合、调查核实	涉事学校发布声明
	Topic I -4	荀晶、冒名顶替、班主任、女儿、两年、连续、大学、高考、录取、实名、山东省教育厅、忏悔书、身份、媒体、院校、南方日报	荀晶实名举报高考被顶替上学事件
	Topic I -5	班主任、见面、济宁、女子、退休、发酵、人民政府、部门、介入、专程	班主任登门来访，部门介入调查
爆发阶段	Topic II -1	顶替、高考道歉、跨省、见面、浙江、大汉、拒绝、曝光	事件进一步曝光，调查组面见荀晶
	Topic II -2	顶替、长相、选中、怀疑、发凉、照片、公平、人生、教育	事件舆论发酵，网民情感宣泄
	Topic II -3	高考、山东、家庭、冒名顶替、大学、社会、公平、命运、底层、博士、学生、落榜	高考制度讨论
	Topic II -4	人生、调查组、女子、真相、链条、严查、教育、学子、寒门、改变命运、考上、遭遇、工作、努力	荀晶回应引发社会公平性热议
	Topic II -5	相似、怀疑、长相、照片、选中、发凉、打听、老家、住址	涉事人长相讨论与拍客
第一次 衰退阶段	Topic III -1	全民震怒、铲断、作恶、链条、人民日报、严查、赔偿、说法	人民日报报道将严查真相
	Topic III -2	工作专班、处分、档案、录取、聊城、责任、山东省、单位	相关部门单位组成调查工作专班
	Topic III -3	高考、顶替、公平、人生、善良、希望、正义、命运、农村、寒门、底层、公正、机会、学子、发声、曝光、道歉、怀疑、选中、发凉、受害者	高考公平性讨论
	Topic III -4	班主任、道歉、电话、找到、害怕、威胁、实名、见面、删帖	荀晶删帖，班主任道歉
	Topic III -5	开展调查、处理、及时、公布、反映、举报、绝不姑息、知情	工作专班进一步跟进调查
	Topic III -6	教育、招生办、操作、农家女、不敢发帖、复杂、学籍	涉事教育部门和招生办的质疑
第二次 增长阶段	Topic IV -1	冒名顶替、社会、部门、调查、追究、责任、人民日报、法律责任、违法行为、回应、山东省教育厅、调查组、公平	山东省教育厅回应介入调查
	Topic IV -2	问题、调查结果、山东、通报、分数线、理科、调剂、学校、黄冈、统招	初步调查结果情况的通报
	Topic IV -3	公安局、派出所、党内处分、犯罪、开除、警告、职务、法律法规	调查结果核实与涉事人员处理
	Topic IV -4	事实、人生、受害者、陈春秀、夸大、舆论、真相	真相讨论与对政府公信力的质疑
第二次 衰退阶段	Topic V -1	真相、公平、正义、维权、夸大其词、热度、夸大、恶心、质疑	高考公平，指责荀晶夸大事实
	Topic V -2	顶替、受害者、高考、陈春秀、损失、撒谎、道义、处罚、欺骗	网络造谣及其恶劣影响的讨论
	Topic V -3	调查结果、班主任、档案、邱小慧、志愿、分数线、委培、填报	荀晶回应调查结果讨论
平息阶段	Topic VI -1	高考、顶替、真相、网络、撒谎、反转、消费、情绪、大众	事件真相公布，反感情绪宣泄
	Topic VI -2	受害者、网友、社会、损失、同情心、事实、道德、伤害、支持、舆论	道德法律层面的争论与分析
	Topic VI -3	邱小慧、北京工业职业技术学院、处理、依纪、依法、档案、煤炭	涉事学校的回应和处理

也扩散转移到网络空间话语权 (Topic V-2)、社会道德与媒介素养 (Topic VI-2) 等话题上。在舆情衰退平息阶段, 公众逐渐趋于理性, 回归对事件本质的关注, 并呼吁大众合理发声, 正确运用网络平台, 在遵守法律法规基础上杜绝虚假发言与传播报道。

4.5 基于时间维的网民情感变化分析

本部分结合舆情反转事件周期划分与主

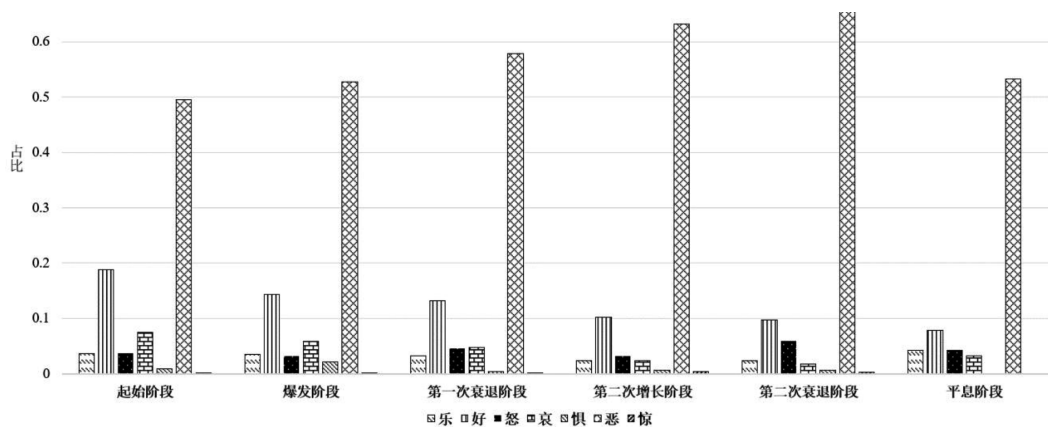


图5 “苟晶事件”情感演变图

首先, 从整个舆情生命周期对应微博情感来看可以发现, 舆论主要由“恶”和“好”两种情绪主导, 表明舆论出现高度情绪集聚现象。其中, 起始阶段和爆发阶段舆论情感类别分布较为相似, 多为网民表达自身愤怒与生气, 如“被恶心到发抖”“看得我气死了”等言论。网民认为苟晶被顶替上大学反映了社会的黑暗 (Topic II-4)、高考制度弊病 (Topic II-3) 以及当时政府与招生办以权谋私, 进而产生怀疑政府媒体公信力等消极不满情绪; 同时对当事人表示深切同情, 体现在对当事人的鼓励、安慰与祝福等正面情感上 (Topic I-1)。第一次衰退阶段, 官方介入调查, 人民日报表示会严查事件真相 (Topic III-1), 工作专班成立并跟进调查 (Topic III-2、Topic III-5), 使得舆情态势得到和缓, 但鉴于尚未查实事件真相, 相关涉事人员也并未得到处理, 网民负面情绪并没有消解, “怒”这一情绪仍在加强, 网民对涉事班主任与学校表示强烈的谴责和愤怒 (Topic III-4), 也贬责当地政府教育部门腐败 (Topic III-6), 质疑高考的公平性。到第二次增长阶段, 官

方公布回应核实调查真相并做出人员惩治等举措 (Topic IV-1~3), 政府部门及时有效的介入在一定程度上化解了公众部分消极情绪。但随着事件真相公布, 舆论导向也开始发生逆转, 因为公众发现调查结果与苟晶所述不符, 认为其故意撒谎, 夸大事实, 是在网络造谣与恶意炒作, 消费大众同情心, 激发了大众的口诛笔伐 (Topic IV-4)。事件反转之后, 公众对调查结果感到震惊的同时也产生被愚弄的情感, 如一位网民说道: “网友发现自己被骗后也会生气的, 苟晶当时有多让网友群情激动, 被发现撒谎后就会对她有多失望, 作为山东济宁人, 我认为苟晶的老师和苟晶都抹黑了这个地方。”因此, 这一阶段网民情绪仍然以“恶”主导, 且负向情绪大幅上涨, “惊”“惧”两种情绪明显增多。在第二次衰退阶段, “恶”“惧”这两种情感强度有所回落, 尽管讨论热度逐天衰退, 但关于社会信任危机与网络造谣恶劣影响的讨论 (Topic V-1, Topic V-2) 使得网民的负面情绪依旧占据主导地位。平息阶段内“惊”“惧”两种情感消失, 公众开始冷静思考这一舆情反转事件,

开始了对高考信息冒用性质的讨论 (Topic VI-1), 对舆情反转事件道德法律层面的争论与分析 (Topic VI-2), 公众认为面对碎片化信息

需要客观分析, 保持清醒的头脑, 这一阶段“乐”“好”情绪回升, 舆论也逐渐结束。

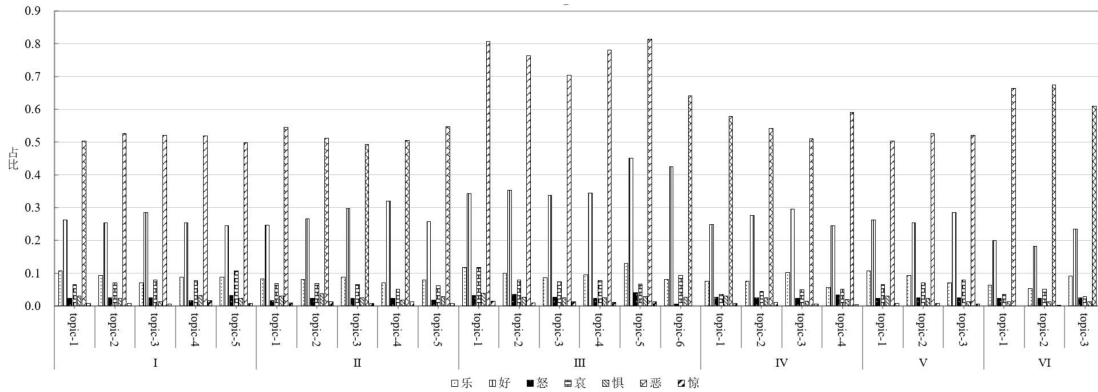


图6 “苟晶事件”微博评论话题情感分布图

4.6 基于空间维的网民情感状态分析

舆情反转是民众情绪变化的客观反映。研究表明, 舆情反转事件中, 用户往往受投射效应、近因效应以及传播失序等因素影响而产生认知偏差, 且不同受众感知到的心理刺激与影响的强度和持续性也不同^[34,40]。本文参考前人研究, 将舆情反转背景下微博用户所在地情感均值作为地区情感评价指标, 来进行网民情感状态空间分析。

基于评论用户的地理位置信息, 统计相关用户的空间分布情况见表3中用户比例所示, 以反映不同地区用户对“苟晶事件”的关注程度。统计发现, 关注用户覆盖全国各省级行政区, 说明目标事件在网络空间引起了广泛的舆论参与。其中, 所在地为北京、广东、山东的用户占比分别高达14.5%、11.3%、11%, 其次为江苏、上海、浙江地区, 评论人数占比均超过5%。以上这些地区的微博用户活跃程度较高, 参与话题数量远超其他地区, 对事件的讨论和关注度更高。

为进一步考察不同地区用户的情感特征, 笔者统计省际尺度下舆情周期内各阶段微博评论整体情感均值及主要关注主题, 其中用户情感均值空间分布如表3所示。在整体情感值的计算中, 为了减小各省样本数量差异对情感分析结果的影响, 采用情感均值分析各省网民情感差异, 情感状态值较大的省份代表其整体情感状态更积极, 情感状态

值较小的省份其整体情感状态则较为消极。

总体来说, 舆情反转前, 不同地区之间情感值整体差异较大, 舆情反转后, 情感状态值整体降低, 用户更加倾向于负面情感, 地区之间差异减小。在舆情反转之前, 青海、宁夏、江西、贵州等地区的情绪状态值较低, 说明这些地区的评论用户情绪普遍低落, 心情愉悦度较低, 情感值较高的地区为黑龙江、吉林、新疆、西藏等地区, 说明这些地区评论用户情绪较为积极, 消极情绪较少; 而用户活跃度较高的省市如北京、广东、山东等地, 用户总体情绪状态处于中游。在舆情发生反转之后, 公众情绪总体转向消极, 省际差别减小。其中贵州、宁夏等地用户情感值略高, 表明这些地区用户总体更为积极, 而黑龙江、江西等地区总体转为消极负面。舆情反转之后, 公众负向情绪增多, 心情愉悦度整体降低, 情感平均值由0.1182下降为0.0782。

由于舆情反转事件直接指向山东省, 研究将其作为舆情涉事地, 而将山东以外的地方作为舆情非涉事地, 统计舆情反转下各周期阶段涉事地与非涉事地的用户主要关注话题与情感值变化(图7)。研究发现, 山东作为舆情事件涉事地, 在舆论反转期间呈现较大的情感波动, 情感值经历了两次的先升后降, 由起始阶段的0.1334最终下降为0.0527; 而非涉事地舆情在舆情反转期间情感呈总体下降趋势, 在平息阶段略有回升。在话题讨

表3 舆情反转下各阶段各省级行政区的情感状态值(按照省份代码排序)

省级行政区	用户比例	各阶段情绪值	省级行政区	用户比例	各阶段情绪值
北京	14.50	0.117/0.157/0.116/0.088/0.070/0.078	湖南	2.05	0.114/0.158/0.101/0.060/0.066/0.229
天津	1.71	0.198/0.116/0.159/0.134/0.036/0.119	广东	11.29	0.117/0.130/0.116/0.086/0.069/0.072
河北	2.93	0.125/0.134/0.095/0.081/0.040/0.020	广西	1.48	0.165/0.128/0.143/0.098/0.056/0.137
山西	1.49	0.140/0.111/0.151/0.079/0.057/-0.054	海南	0.54	0.157/0.114/0.092/0.077/0.075/-0.098
内蒙古	0.84	0.084/0.116/0.125/0.096/0.065/0.248	重庆	1.90	0.124/0.064/0.066/0.092/0.081/0.172
辽宁	2.55	0.160/0.097/0.068/0.072/0.070/-0.041	四川	4.13	0.098/0.153/0.125/0.114/0.069/0.090
吉林	1.00	0.107/0.199/0.106/0.106/0.068/0.010	贵州	0.62	0.014/0.126/0.147/0.296/0.055/0.096
黑龙江	1.36	0.192/0.131/0.096/0.086/0.035/-0.010	云南	0.98	0.110/0.166/0.140/0.118/0.073/0.024
上海	7.29	0.155/0.108/0.106/0.101/0.054/0.039	西藏	0.35	0.189/0.118/0.207/0.030/0.113/0.166
江苏	7.59	0.119/0.111/0.100/0.091/0.068/0.108	陕西	2.39	0.134/0.127/0.091/0.083/0.05/-0.052
浙江	5.96	0.137/0.117/0.093/0.113/0.059/0.129	甘肃	0.50	0.072/0.253/0.191/0.182/0.077/0.180
安徽	2.58	0.145/0.152/0.097/0.089/0.044/0.073	青海	0.15	0.035/0.032/0.050/-0.119/0.120/0.120
福建	2.80	0.131/0.129/0.119/0.147/0.078/0.032	宁夏	0.29	0.089/0.046/0.139/0.142/0.126/-0.031
江西	1.34	0.084/0.091/0.028/0.082/0.033/0.032	新疆	0.51	0.119/0.277/0.175/0.004/0.082/-0.075
山东	10.97	0.133/0.139/0.086/0.101/0.061/0.053	台湾	0.18	0.061/0.114/0.220/0.217/0.140/0.484
河南	3.31	0.143/0.133/0.122/0.104/0.064/0.041	香港	0.50	0.145/0.080/0.116/0.111/0.105/-0.006
湖北	3.69	0.160/0.139/0.161/0.109/0.051/0.060	澳门	0.21	-0.157/0.095/0.068/0.107/0.073/-0.194

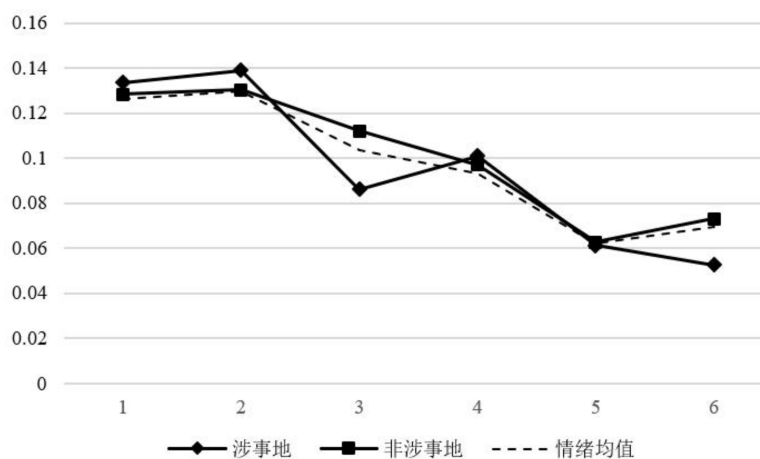


图7 舆情涉事地与非涉事地各阶段情绪值变化图

论上, 舆情涉事地与非涉事地在反转前后的话题关注点也存在着差异。舆情反转之前, 涉事地和非涉事地共同关注高考制度公平性 (Topic III-3)、事件进展 (Topic II-1、Topic III-4)、官方回应 (Topic III-1) 以及情感宣泄 (Topic I-1) 等话题, 但舆情涉事地还发布了较多关于提醒调查结果核实与涉事人员处理 (Topic IV-3)、持续为事件发声 (Topic I-4) 的内容, 舆情非涉事地则更多关注事件背后所体现的高考制度与社会公平性 (Topic II-3、Topic II-4) 内容。舆情反转之后, 指责苟晶夸大事实 (Topic V-1) 以及对事件当事人和官方处理结果 (Topic V-3、Topic IV-4) 进行讨论是舆情涉

事地和非涉事地的共同关注点, 山东作为舆情涉事地, 侧重关注事件真相调查的进展、本地官方政府是否有效作为 (Topic IV-1、Topic IV-2), 而舆情非涉事地的关注话题高度集中, 其中 Topic V-1 话题占比高达 42.01%, 说明非涉事地更聚焦事件性质本身, 而较少关注涉事地点。

5 结论与展望

舆情反转事件涉及到的舆情信息裂变传播牵扯到用户情感“一波三折”的变化, 研究舆情反转事件的情感演变对于把握“后真相时代”的舆情监测和导控工作具有重要意义。本文选取特定舆情反转事件, 划分舆情演化

生命周期，融合主题分析与情感分析，从时间维度和空间维度两方面探讨舆情反转事件中的情感演化特征及相关规律。研究发现：

第一，发生一次反转的舆情反转事件可以划分为起始阶段、爆发阶段、第一次衰退阶段、第二次增长阶段、第二次衰退阶段、平息阶段六个生命周期发展阶段。本文中即是如此，而发生多次反转的则基于上述阶段叠加演化。另外，舆情反转事件往往有关键反转节点，这一节点一般以官方或正式组织、个人的公告、公示、证明、说明等为标志，本例中就是以官方调查结果公示为标志。

第二，舆情反转事件的情感演化在时间和空间上都有不同表征。不同时间片的舆情主题和情感会随着事件发展的变化而发生变化，转折前的情感与转折后的情感往往形成鲜明的对比，它们可能在情感结果上偏向一致（本例中即是如此），但其内容却明显不同，这主要是由反转事件本身的特点所决定。不同地区对舆情反转事件关注程度不同，情感状态值也存在差异。如本例中在事件反转前后，舆情涉事地与非涉事地在关注话题上也体现出一定的差异，舆情涉事地显然更容易引起用户的关注和情感变化，同时会更加关注与本地区域特征相关的一些内容。

通过上述研究与发现认为，针对舆情反转这一特殊事件类型，判定事件是非仍然需要“让信息再飞一会儿”，而不能完全遵循“第一反应”。从政府角度看，政府需要加

强对舆情反转事件的监测管控，发挥官方公信力，及时公示事件真相，缓解舆情反转导致的极端情绪波动，减少由事件延伸出的动摇社会信任、污染网络环境的消极话题。从媒介角度看，相关舆情信息传播主体需要提升媒介素养，正确运用网络空间话语权，不唯“流量为王”，不唯“首创首发”，消除舆情初期的虚假报道；不同地区的媒介主体可以加强信息整合，规避信息延迟和误传，有效发挥正向传播作用。从用户角度看，广大网民要树立舆情研判的理性自觉，同时匡正相关表层认知，客观理性发声，避免被流量误导，避免消极过激情绪泛滥。

本文的研究可进一步丰富舆情反转相关理论体系，并为舆情反转监测管控实践提供参考。然而本文也存在以下研究不足，首先，本文仅选取微博平台作为数据获取来源，以单一舆情反转事件进行探索分析，研究发现可能不全面；其次，将主题分析与时空维度情感分析深度融合还值得进一步优化思考，关于时空视域下微博文本情感分析结果的差异性有待方法论层面的优化，在不同类型舆情反转事件情境下也有待进一步的综合考察。后续相关研究可以采用更多类型网络媒体平台的数据，借助其他领域的可视化方式，对不同类型舆情反转事件的情感演化趋势进行更全面详细的分析，以掌握其共性和个性的特征与规律。

参考文献

- [1] 谭玲珑. 网络舆情反转中的框架效应研究[D]. 长沙: 湖南大学, 2018.
- [2] 倪万. 信息不对称理论下舆情反转乱象及媒体责任[J]. 青年记者, 2017(21): 45-46.
- [3] 田文利. 从网络舆情反转谈如何防范网络虚假新闻[J]. 当代电视, 2016(11): 91-92.
- [4] Chen T, Wang Y, Yang J, et al. Modeling public opinion reversal process with the considerations of external intervention information and individual internal characteristics[J]. Healthcare, 2020, 8(2): 160.
- [5] 汪明艳, 余丽彬, 译冰. 舆论反转中群体极化效应的影响因素研究[J]. 情报杂志, 2018, 37(9): 106-112, 119.
- [6] 黄远, 刘怡君. 网络舆论反转效应研究[J]. 管理评论, 2016, 28(8): 71-78.
- [7] 王玉龙. 舆情反转: 突发事件过程中网络群体极化的角色研究——基于“天价鱼事件”的分析[J]. 电子政务, 2016(5): 16-24.
- [8] 岳甜, 王新童, 刘悦. 反转新闻的网络舆论演变机制研究——以“罗一笑事件”为例[J]. 当代传播, 2018(2): 92-94, 98.
- [9] 张敏, 刘雪瑞, 张艳. 重大疾病在线救助事件社交舆情的受众情绪反转形成机理分析, [J]. 情报杂志, 2018, 37(11): 122-130.
- [10] 张兆曙. 网络舆论的“反转”何以可能? ——基于发生结构的视角[J]. 学海, 2018(4): 150-157.

- [11] 关琮严. 网络舆论反转机制及其治理策略探究[J]. 中国出版, 2018(14): 37-40.
- [12] 陈一新, 陈馨悦, 吕妍, 等. 基于改进 Heggelmann-Krause 模型的微博舆论反转研究[J]. 情报理论与实践, 2020, 43(1): 82-89.
- [13] Jiang G, Li S, Li M. Dynamic rumor spreading of public opinion reversal on Weibo based on a two-stage SPNR model [J]. Physica A: Statistical Mechanics and its Applications, 2020, 558: 125005.
- [14] 袁野, 兰月新, 张鹏, 等. 基于系统聚类的反转网络舆情分类及预测研究[J]. 情报科学, 2017, 35(9): 54-60.
- [15] 夏一雪, 兰月新, 刘荣, 等. 大数据环境下网络舆情反转机理与预测研究[J]. 情报杂志, 2018, 37(8): 92-96, 207.
- [16] 王楠, 李海荣, 舒孺. 基于改进 SMOTE 算法与集成学习的舆情反转预测研究[J]. 数据分析与知识发现, 2021, 5(4): 37-48.
- [17] 王婷, 杨文忠. 文本情感分析方法研究综述[J]. 计算机工程与应用, 2021, 57(12): 11-24.
- [18] 张刚强, 刘群, 纪良浩. 基于序贯三支决策的多粒度情感分类方法[J]. 计算机科学, 2018, 45(12): 153-159.
- [19] Chen L C, Lee C M, Chen M Y. Exploration of social media for sentiment analysis using deep learning[J]. Soft Computing, 2020, 24(11): 8187-8197.
- [20] 王丹, 张海涛, 刘雅姝, 等. 微博舆情关键节点情感倾向分析及思想引领研究[J]. 图书情报工作, 2019, 63(4): 15-22.
- [21] 杜毅贤, 徐家鹏, 钟琳颖, 等. 网络舆情态势及情感多维特征分析与可视化——以 COVID-19 疫情为例[J]. 地球信息科学学报, 2021, 23(2): 318-330.
- [22] 徐琳宏, 林鸿飞, 潘宇, 等. 情感词汇本体的构造[J]. 情报学报, 2008, 27(2): 180-185.
- [23] 钟佳娃, 刘巍, 王思丽, 等. 文本情感分析方法及应用综述[J]. 数据分析与知识发现, 2021, 5(6): 1-13.
- [24] 王晰巍, 张柳, 文晴, 等. 基于贝叶斯模型的移动环境下网络舆情用户情感演化研究——以新浪微博“里约奥运会中国女排夺冠”话题为例[J]. 情报学报, 2018, 37(12): 1241-1248.
- [25] 邓君, 孙绍丹, 王阮, 等. 基于 Word2Vec 和 SVM 的微博舆情情感演化分析[J]. 情报理论与实践, 2020, 43(8): 112-119.
- [26] Zhou Y, Xu R, Gui L. A sequence level latent topic modeling method for sentiment analysis via CNN based diversified restrict boltzmann machine[C]//2016 International Conference on Machine Learning and Cybernetics (ICMLC). IEEE, 2016: 356-361.
- [27] 张海涛, 王丹, 徐海玲, 等. 基于卷积神经网络的微博舆情情感分类研究[J]. 情报学报, 2018, 37(7): 695-702.
- [28] 梁军, 柴玉梅, 原慧斌, 等. 基于极性转移和 LSTM 递归网络的情感分析[J]. 中文信息学报, 2015, 29(5): 152-159.
- [29] 曾子明, 万品玉. 基于双层注意力和 Bi-LSTM 的公共安全事件微博情感分析[J]. 情报科学, 2019, 37(6): 23-29.
- [30] 王晰巍, 邢云菲, 韦雅楠, 等. 大数据驱动的社交网络舆情用户情感主题分类模型构建研究——以“移民”主题为例[J]. 信息资源管理学报, 2020, 10(1): 29-38, 48.
- [31] 李超雄, 黄发良, 温肖谦, 等. 基于动态主题情感混合模型的微博主题情感演化分析方法[J]. 计算机应用, 2015, 35(10): 2905-2910.
- [32] 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.
- [33] 朱晓霞, 宋嘉欣, 孟建芳. 基于动态主题 - 情感演化模型的网络舆情信息分析[J]. 情报科学, 2019, 37(7): 72-78.
- [34] 张琛, 马祥元, 周扬, 等. 基于用户情感变化的新冠疫情舆情演变分析[J]. 地球信息科学学报, 2021, 23(2): 341-350.
- [35] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation[J]. Journal of Machine Learning Research, 2003, 3: 993-1022.
- [36] 张柳, 王晰巍, 黄博, 等. 基于 LDA 模型的新冠肺炎疫情微博用户主题聚类图谱及主题传播路径研究[J]. 情报学报, 2021, 40(3): 234-244.
- [37] 杜振雷. 面向微博短文本的情感分析研究[D]. 北京: 北京信息科技大学, 2013.
- [38] 人民日报. 关于苟晶反映被冒名顶替上学等问题调查处理情况的通报[EB/OL]. [2021-05-16]. https://weibo.com/ttarticle/p/show?id=2309404522599614120045#_0.
- [39] 贾亚敏, 安璐, 李纲. 城市突发事件网络信息传播时序变化规律研究[J]. 情报杂志, 2015, 34(4): 91-96, 90.
- [40] 林振明. 舆情反转的应对策略与媒体责任[J]. 科技传播, 2021, 13(12): 114-116.

(收稿日期: 2021-09-13)