

新冠疫情相关社交媒体谣言传播量化分析

陈慧敏¹ 金思辰² 林 微¹ 朱泽宇¹ 仝凌波^{2,4} 刘一芄² 叶奕宁² 姜维翰³ 刘知远² 孙茂松²
金兼斌¹

¹(清华大学新闻与传播学院 北京 100084)

²(清华大学计算机科学与技术系 北京 100084)

³(北京航空航天大学软件学院 北京 100191)

⁴(美国圣母大学心理学系 美国印第安纳州南本德 46556)
(huimchen@mail.tsinghua.edu.cn)

Quantitative Analysis on the Communication of COVID-19 Related Social Media Rumors

Chen Huimin¹, Jin Sichen², Lin Wei¹, Zhu Zeyu¹, Tong Lingbo^{2,4}, Liu Yipeng², Ye Yining²,
Jiang Weihuan³, Liu Zhiyuan², Sun Maosong², and Jin Jianbin¹

¹(School of Journalism and Communication, Tsinghua University, Beijing 100084)

²(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

³(School of Software, Beihang University, Beijing 100191)

⁴(Department of Psychology, University of Notre Dame, South Bend, Indiana, USA 46556)

Abstract The outbreak of the COVID-19 pandemic is accompanied by numerous rumors spreading on the social media platform, which seriously affects the stability of society and the safety of public. Existing quantitative analyses of COVID-19 related social media rumors only focus on single element of communication, such as content, while ignoring other basic elements of communication, including communicator, audience, and effect. Besides, compared with the real social media rumor data, the rumor data of these studies have distribution bias and lack of information. Therefore, we conduct a more comprehensive quantitative analysis on the communication of COVID-19 related social media rumors based on the Sina Weibo platform. Specifically, we first analyze the communication content of rumors, including the analysis of the topic, involved regions, event tendency and sentiment. Further, we investigate the users engaged in rumor communication and divide the users into three categories, namely, rumor posters, rumor spreaders, and rumor informers. We explore the basic attributes, topic preferences, individual sentiments, and self-network characteristics of the engaged users. Finally, we study the public opinion triggered by rumors, including the overall sentiment distribution, its correlation with topics, keywords and regions, as well as the evolution of sentiment. To conclude, this study first quantitatively analyzes COVID-19 related social media rumors from the perspective of different basic elements in communication. It provides a more comprehensive and profound understanding of COVID-19 related social media rumors and is of great value for both research and management of rumor in public emergencies.

Key words COVID-19; social media; rumor communication; communication structure; quantitative analysis

收稿日期:2020-10-10;修回日期:2021-01-18

基金项目:国家自然科学基金项目(13&ZD190)

This work was supported by the National Social Science Foundation of China (13&ZD190).

通信作者:刘知远(liuzy@tsinghua.edu.cn)

摘要 新冠肺炎疫情的爆发伴随着大量的谣言在社交媒体平台传播,对网络秩序和社会稳定产生了不良影响.已有的疫情相关社交媒体谣言传播量化分析研究仅对谣言内容等单一传播要素展开分析,而忽略了构成信息传播的其他基础要素,包括传播者、受众以及传播效果等.同时,这些研究的谣言数据与真实的社交媒体谣言数据也存在分布偏差和信息缺失.因此,基于新浪微博平台对新冠疫情相关社交媒体谣言的传播展开更加全面的量化分析.具体而言,首先对谣言传播内容进行分析,包括其主题分析、涉及地区分析、事件倾向性分析以及情感分析;进一步对谣言参与用户进行分析,将参与用户分为3类:造谣者、传谣者和辟谣者,并分别对其基础属性、关注主题、个体情绪以及自网络属性进行探究;最后对谣言引发舆情进行分析,探究其情感的整体分布、与主题、关键词和地区的关系、以及情感的演变规律.该研究首次从信息传播的各个基础要素层面对疫情相关的社交媒体谣言传播展开量化分析,不仅对新冠肺炎疫情相关谣言传播有了更全面深刻的认识,同时对突发公共事件的谣言研究和谣言治理也具有十分重要的价值.

关键词 新冠肺炎;社交媒体;谣言传播;传播结构;量化分析

中图法分类号 TP391

谣言一般指没有事实依据而捏造的言论^[1],是社会科学和计算机科学的重要研究问题^[2-5].谣言最初主要以口耳相传形式流传,随着近年来互联网的发展,诸如微博、微信、Twitter等社交媒体成为谣言传播的新平台.由于社交媒体平台的开放性和便捷性,大量虚假言论在社交媒体平台上广泛传播,对网络秩序和社会稳定产生了不良影响^①.

重大公共事件出现时往往伴随着谣言的产生,如2014年的埃博拉瘟疫^[6]、2016年的美国总统大选^[7].而2019年出现的新冠肺炎(COVID-19)疫情发展至今,已经构成了全球性“大流行”(Pandemic)^②.与之相伴的则是信息瘟疫(Infodemic)^③的到来,大量谣言在社交媒体平台上衍生并广泛传播,这些谣言不仅严重危害到公众的生命安全,同时给社会带来了极大的恐慌和不稳定因素^④.与此同时,已有研究表明,类似于这次新冠肺炎疫情的突发公共事件相关谣言,其传播与社会常态时谣言的传播具有十分不同的特性^[8-11].因此,对新冠肺炎疫情相关社交媒体谣言传播进行量化分析不仅具有深远的社会意义,同时具有十分重要的研究价值.

目前已经有一些研究者针对新冠肺炎疫情相关社交媒体谣言传播展开量化研究.Islam等人^[12]收集了来自Twitter和Facebook等社交媒体平台的

2311篇疫情相关谣言和阴谋论报道,并对其进行内容分析.Cinelli等人^[13]收集了来自6个国外社交媒体平台的疫情相关谣言,并对其传播模式进行对比分析.钱迎飞^[14]和江苏佳^[15]则收集了国内的腾讯较真查证平台和丁香园辟谣平台的疫情相关谣言,并展开内容分析.然而,目前这些工作仅对疫情相关社交媒体谣言传播的单一要素展开量化分析,如内容特征.而根据拉斯韦尔的经典信息传播模型^[16],构成信息传播的基本要素包括传播者(communicator)、传播内容(message)、传播媒介(medium)、受众(audience)以及传播效果(effect),其中对谣言的传播者、传播媒介、受众以及传播效果的研究仍未涉及.同时,已有工作基于的谣言数据大多来自于人工归纳整理,相对于真实的社交媒体平台谣言数据,其仍然存在一定的分布偏差和信息缺失:其分布偏差体现在人工归纳收集的谣言数据多为广泛传播的谣言,而社交媒体平台中的谣言只有极少数具有广泛的传播范围^[17].其信息缺失体现在人工归纳收集的谣言往往只保留谣言内容,而其引发的舆情和传播信息以及参与用户信息则没有保留.

为了解决分布偏差和信息缺失问题,本文基于国内最大的微博服务平台——新浪微博,对新冠疫情相关社交媒体谣言传播展开全面地量化分析.我们

① http://www.xinhuanet.com/2020-01/27/c_1125506246.htm

② <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19%E2%80%949411-march-2020>

③ <https://www.who.int/news-room/feature-stories/detail/immunizing-the-public-against-misinformation>

④ <https://www.who.int/news-room/detail/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation>

首先构建了一个新冠肺炎疫情相关谣言数据集,其中包括谣言及其转发、评论以及参与用户信息的数据.本文对疫情期间的社交媒体谣言研究聚焦于对其传播结构和传播要素的分析,不对谣言的真实性做探讨^①.进一步,我们结合计算机和深度学习方法,对构成谣言传播的各基本要素分别展开分析:首先我们针对传播内容展开谣言内容分析^②,紧接着我们针对传播者和受众展开谣言用户分析,最后我们针对传播效果展开谣言舆情分析.通过对新冠肺炎疫情相关谣言传播的量化分析,我们得出了3个有意义的结论:

1) 在谣言内容方面,疫情防控类是出现最多的谣言类型,其次是海外相关类;谣言在疫情发展的不同阶段具有不同的主题特征;谣言涉及地区多为疫情严重地区,且地区相关谣言多集中产生在疫情早期阶段;谣言中“坏消息”居多,且传达的情绪也以负面为主.

2) 在谣言参与用户方面,造谣者原创微博及其发微博字数更多,造谣者则转发微博更多;造谣者在上午更为活跃,倾向于在上午发表言论,且其影响力远大于造谣者和辟谣者;辟谣者在历史微博中表现出了对新冠相关科学研究主题更多的关注,且其自

网络结构更加聚集;造谣者和造谣者的个人情绪则更为负面.

3) 在谣言所引发的舆情方面,近半数谣言评论持中性态度,在此之外,消极评论则整体多于积极评论;疫情相关奇闻异事和名人轶事类谣言热度最高,但疫情防控类谣言评论最正面,病毒传播类谣言评论则最负面,其中“封城”相关谣言引发大量积极舆论;谣言评论情感往往向负面演变,其中辟谣信息的出现也是其中原因之一.

据我们所知,本文是较早全面地从信息传播的各基础构成要素角度对新冠肺炎疫情相关社交媒体谣言传播展开量化分析的工作.通过本文的研究,可以对疫情相关社交媒体谣言传播有更全面深入地认识,对突发公共事件的谣言传播研究和谣言治理具有重要意义.

1 新冠肺炎相关谣言数据

本文使用的新冠肺炎疫情相关谣言数据来源于新浪微博社区管理中心和举报处理大厅^③,该中心接受微博用户对于不良信息的举报,我们选择其中受举报并已公示的不实信息数据作为谣言数据.图1



Fig. 1 Web page of Weibo misinformation processing platform

图1 微博社区管理中心不实信息处理公示页面

- ① 数据主要来源于新浪微博社区管理中心中被举报和核查的不实信息,不排除其可能存在覆盖度不全和人工判断的误差.同时谣言的判断具有不确定性,随着更多信息的引入可能会发生变化.
- ② 由于是基于新浪微博的谣言数据分析,故我们省去对传播的基本要素——传播媒介——的分析.
- ③ <https://service.account.weibo.com/?type=5&status=0>

展示了疫情期间一条不实信息的处理公示页面。

具体来说,我们爬取了 2020-01-22—2020-07-27 的不实信息数据,并人工筛选出其中疫情相关的谣言数据,共 448 条^①,其数量随时间变化如图 2 所示,可以看出,疫情谣言主要集中在 2020 年 1—4 月,这一时期是国内疫情爆发和国外疫情开始阶段,疫情谣言涉及的话题涵盖了病毒信息、疫情防控与国际轶事,同时后期 6 月中旬至 7 月出现了疫情谣言小高峰,主要涉及的话题与当时北京疫情小范围 2 次爆发有关。

进一步地,为了对谣言传播中的传播者和受众

进行量化分析,我们爬取了谣言传播中参与用户的用户数据,包括其基本属性数据(用户认证、关注数和粉丝数等)、和历史微博数据(时间为 2019-07-31—2020-07-31)。我们将谣言传播参与用户分为 3 类:辟谣者(举报人)、造谣者(被举报人)、传谣者(转发用户)。经过统计,疫情相关谣言的造谣者共 468 人^②,传谣者共 1321 人,辟谣者共 1076 人。

最后,为了分析谣言的传播效果,我们爬取了谣言的转发和评论数据,共 45 638 条(统称为评论)。图 3 展示了疫情谣言评论数量随时间的变化。

由图 3 可知,谣言评论主要集中在 2020 年 1—

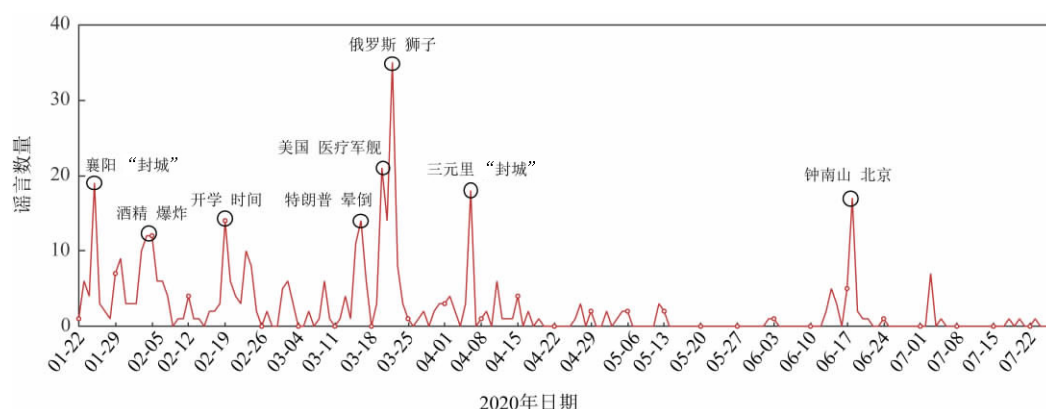


Fig. 2 Number of COVID-19 related rumors over time (2020-01-22—2020-07-22)

图 2 疫情谣言数量随时间变化(2020-01-22—2020-07-22)

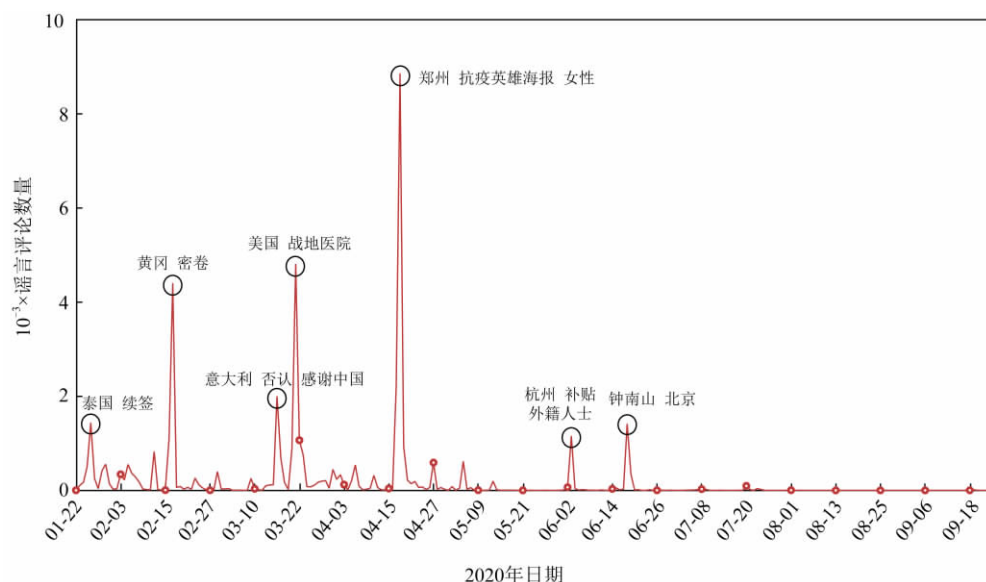


Fig. 3 Number of COVID-19 related rumor comments over time(2020-01-22—2020-09-18)

图 3 疫情谣言评论数量随时间变化(2020-01-22—2020-09-18)

① 数据集已公开于 <https://github.com/thunlp/COVID19-Social-Datasets/tree/master/rumor>

② 由于部分内容已被删除的谣言发布信息,所以较前文的谣言数量更多。

4月,即国内疫情爆发的主要阶段和国外疫情流行初期.2020年6月出现了评论数量的小高峰,与疫情2次爆发相关.进入2020年7月后,基本未再出现新的疫情相关谣言,评论数也随之趋向于零.

2 谣言内容分析

内容是信息传播过程的核心.本节我们首先对疫情相关谣言的内容进行分析,具体维度包括主题分析、涉及地区分析、事件倾向性分析和情感分析.

2.1 谣言主题分析

我们基于爬取的疫情谣言数据,结合已有研究对谣言的划分^[14-15,18],将疫情谣言分为5类,如表1所示.我们招募了2名新闻与传播专业的学生作为标注员对数据进行标注.正式标注之前,我们首先对

标注员进行培训,要求其阅读包含规则和案例的标注指南,并标注测试样本.最终经过培训,2名标注员的标注结果一致占比为95%,并进一步完成所有数据的标注.

谣言主题分布.由表1可以看到,数量最多的疫情谣言类型是疫情防控类谣言,共计226条,占总量的一半以上,显示了疫情期间谣言热点主要与防疫抗疫相关事件相关.疫情期间,预防方法、防控措施与民生息息相关,涉及生命安全的保障,这反映出谣言主要围绕在公众所关心的话题领域.主题分布第2位是国际相关类谣言,共计206条.此类谣言大部分涉及“中国人”“国籍”“美国”等民族主义相关词语^[19].此类谣言主要凭借受众的民族主义情绪获得讨论度和关注度^[19].此外,这类谣言也善于通过迎合人们的认知来获取认同感、博取可信度.

Table 1 Topic Categories of Rumors

表1 谣言主题分类

谣言主题类别	谣言内容	谣言数量
病毒传播	确诊个例出逃、聚集性爆发、小范围二次爆发相关谣言	35
病毒信息	病毒传播途径、病毒溯源、病毒危害相关谣言	16
疫情防控	“封城”、旅行禁令、复工复产、医院相关谣言	226
国际相关	国际双边政策、对外抗疫行动、坊间外籍轶事、国际轶事相关谣言	206
其他	坊间奇闻异事、名人轶事	81

谣言阶段性分布.我们进一步根据《抗击新冠肺炎疫情的中国行动》白皮书^①将疫情按照国内外形势划分为4个阶段,分别分析了这4个阶段的谣言主题.这4个阶段分别是:疫情爆发阶段(2020-01-20至2020-02-20)、国内疫情得到控制阶段(2020-02-21至2020-04-28)、疫情全球大流行阶段(2020-03-11以来)和疫情常态化阶段(2020-04-29以来).如图4所示,通过对比发现,谣言具有明显的阶段性分布特点,具体表现为:

1) 病毒传播类谣言主要分布的阶段包括:疫情爆发阶段和疫情控制阶段.在疫情爆发阶段,传播类谣言涉及各地确诊隐瞒出逃的谣言事件;在国内疫情得到控制这一阶段,随着国内疫情得到控制,对于复工复产的猜测、聚集性爆发的谣言增加;在疫情全球大流行阶段,病毒传播类谣言又出现了关于境外输入性传播的小高峰.

2) 病毒信息类谣言主要分布的阶段包括:疫情爆发阶段和疫情控制阶段早期.在疫情爆发阶段,诸

如“人工病毒”“犬类传播”之类的谣言盛行,主要是由于疫情开始爆发时公众对病毒的认识不足,缺少权威介绍,此时对病毒信息的猜测较多;而在国内疫情得到控制初期出现了一些有关病毒后遗症和传播途径导致二次爆发的病毒信息类谣言.

3) 疫情防控类谣言在各个阶段都有分布,属于民生相关、民众持续关注类谣言,主要涉及事件包括:疫情爆发阶段关于“封城”等防控措施的虚假信息;国内控制阶段对于复工复产的谣传;疫情全球大流行阶段对于国外防控措施的谣言.

4) 国际相关类谣言主要分布在疫情全球大流行阶段早期,随着疫情震中从国内转移到全球范围,海外相关类谣言数量开始增加,包括一些国际轶事,国际疫情防控相关的谣言.在疫情全球大流行阶段后期,谣言的总体数量回落,疫情谣言主要包括海外相关、疫情防控,例如境外输入与国内小范围爆发联系、签证政策等.

从不同主题谣言的时序分布占比可以看出:谣言

① http://www.gov.cn/zhengce/2020-06/07/content_5517737.htm

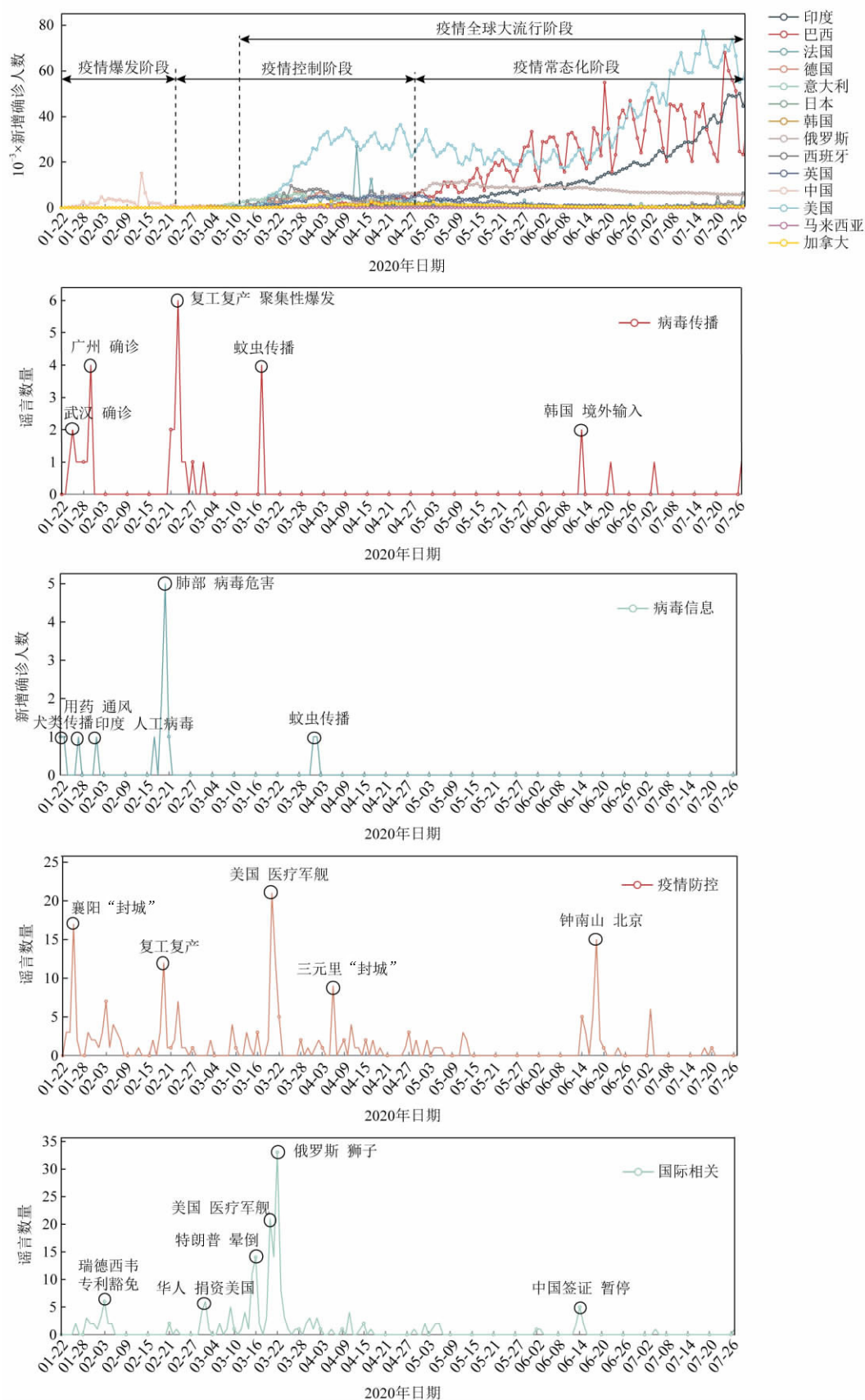


Fig. 4 Numbers of different categories of rumors (2020-01-22—2020-07-26)

图4 谣言分类数量随疫情阶段变化(2020-01-22—2020-07-26)

的出现具有明显的阶段性,其主题随疫情的形势变化而不断演变.这对于谣言的治理和预防有重要的借鉴意义,在疫情不同阶段,社会关切不同,应该加强相应信息的公开性、透明度.

2.2 谣言涉及地区分析

本节主要研究海外国家和国内具体省市谣言讨论与该地区疫情关系.首先我们通过人工标注提取谣言中出现的地名,我们招募了地理信息科学专业的标注员进行标注,标注员必须阅读包含规则和案例的标注指南,而后才可开始标注.标注结束后,我们根据提取出来的地名集合将谣言进行归类,与相应讨论地区关联起来.根据标注结果统计,谣言讨论度排名前5的地区分别是湖北、美国、俄罗斯、日本、北京.

谣言涉及地区多为疫情严重地区.我们进一步将谣言讨论地区进行归类,得出谣言涉及的不同国家的谣言数目、国内不同省份的谣言数目、国内不同城市的谣言数目.根据统计,我们发现疫情涉及的国

内城市谣言数排行榜中,武汉、北京最多,分别为47条和40条.谣言涉及的省份主要是湖北省,包含76条.在涉外谣言数目分布中,美国、俄罗斯、日本排名前三,涉及谣言数目分别为104条、39条和23条.可以发现,不论是国内还是国外,疫情谣言讨论的热点地区都是疫情比较严重的地区.

地区相关谣言集中产生在地区疫情早期阶段.基于统计数据,我们进一步分析谣言讨论地区随该地区疫情变化特征.如图5所示,对于各地区本身的谣言,大多出现在该地区疫情刚刚兴起的时候.其可能的原因包括:首先是疫情初期信息的不确定性^[20],即在疫情初期,大多数公众对疫情的认识十分模糊,对新冠肺炎信息有很多的误解,给谣言的传播以可乘之机.其次是疫情初期人心恐慌,不安定因素加剧了谣言传播^[21].最后是新闻的助推效应^[22],即疫情初期,新冠肺炎成为新闻报道焦点,公众对其关注程度陡然上升,也会助长相关谣言的传播.

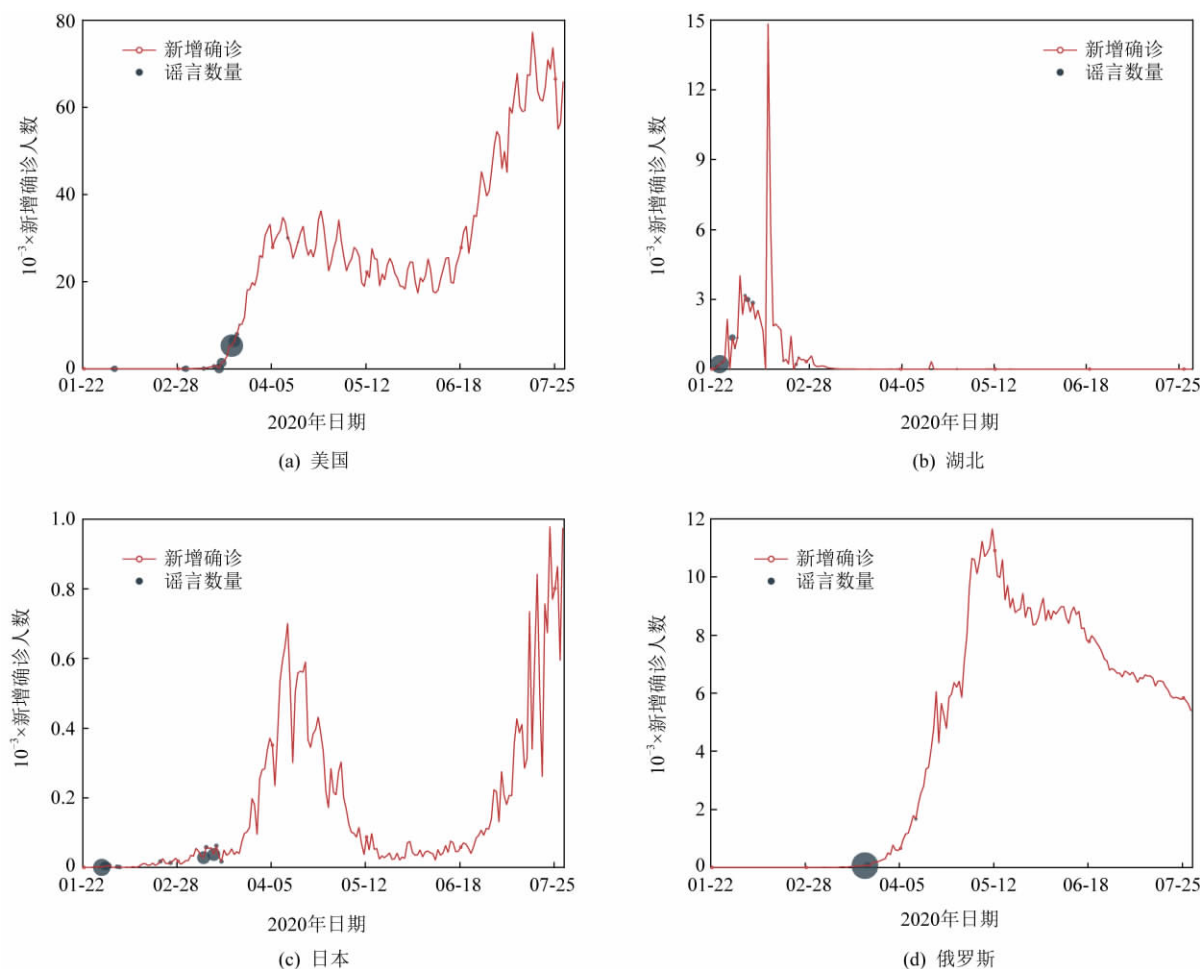


Fig. 5 Relevance between number of rumors and its related regions' confirmed cases (2020-01-22—2020-07-25)

图5 谣言讨论地与该地区新增确诊病例关系(2020-01-22—2020-07-25)

2.3 谣言事件倾向性分析

事件倾向性分析是关于谣言涉及事件正负面的研究.正面性是指文本内容释放的事件消息是积极的,负面性是指文本内容涉及的事件消息是消极的.我们通过人工标注的方法将谣言涉及事件分为正面、中性、负面 3 类.标注方法与 2.1 节中主题标注相同.

“坏消息”居多.通过对疫情谣言的正负面分类分析,我们发现负面谣言在疫情谣言中数量最多,为 269 条,占比 60%;正面谣言 111 条,占比 25%;中立谣言最少,69 条,占比 15%.这一方面体现出谣言往往伴随具有倾向性和话题度的事件产生;另一方面表现出负面事件更容易引发谣言,这可能是由于负面信息更容易引发公众的恐慌心理,而当人们身处惶恐紧张情况下,负面信息再经过加工渲染后更易受到关注并广泛传播.

此外,我们还探究了不同倾向谣言的内容分布.正面的疫情谣言主要包括国际援助、复工复学、疫情防控工作、医护人员等相关的事件和消息;负面的疫情谣言主要包括病毒传播、病毒危害、恐怖事件和国际关系等相关事件和消息.

2.4 谣言情感分析

谣言的情感分析主要包括情感分类和用词特征分析,我们参考 Allport 等人^[3]和阮璋琼等人^[23]的谣言情绪分类类型,疫情谣言按照所投射出来的心理将谣言情感分为 5 类,分类标准如表 2 所示.我们采用人工标注的方法,对疫情谣言进行情感分类标注.标注方法与 2.1 节中主题标注相同.

Table 2 Different Sentiment Types Rumor

表 2 谣言情感分类

谣言情感类型	谣言内容
愿望型	对抗疫形势的期待和愿望
恐惧型	具有恐慌色彩的疫情信息
善意型	对医护人员的赞赏、对抗疫行动的支持
敌意型	带有愤怒、讽刺等攻击色彩的疫情相关言论
中立性	没有明显的情绪倾向的疫情谣言

1) 消极情感谣言占上风.如表 3 所示,可以发现,充满消极情感的恐惧类型和敌意类型的谣言总数多于充满积极情感的善意类型谣言和愿望类型谣言的总数,一定程度上反映人们在疫情压力之下,易寻找负面情绪的宣泄口,并可能促进分裂与焦虑情绪在社交媒体上进一步蔓延.

① 由于中立型谣言不包含感情色彩,以及愿望型谣言数量过少,其特征词的分析予以省略.

Table 3 Cross Comparison Between Rumor Content Categories and Rumor Sentiment Types

表 3 谣言主题类型与情感类型的交叉分析

谣言主题类型	谣言情感类型				
	愿望型	恐惧型	善意型	敌意型	中立型
病毒传播	0	19	1	3	12
病毒信息	0	12	0	2	2
疫情防控	22	51	88	20	45
国际相关	0	24	21	76	85
其他	0	30	12	13	26
总计	22	136	122	114	170

但值得注意的是,对比来看,善意类型的疫情谣言占比也并不低,例如一则关于钟南山出征北京参与新冠肺炎治疗的谣言一时间引起很高的关注度和转发量,可见人们在面对疫情的压力时,不仅存在恐慌、宣泄的心理状态,也存在寻求希望寄托、自我安慰的心理状态.

2) 谣言内容类型与情感类型的交叉分析.如表 3 所示,我们通过分析不同谣言内容类型与其情感类型之间的联系,发现在病毒传播类型的疫情谣言中恐惧型谣言最多,主要涉及内容包括确诊病例隐瞒潜逃等;在病毒信息类谣言中也是恐惧类谣言最多,主要包括病毒传播途径和病毒对人体造成的危害相关谣言;在疫情防控谣言中善意类型谣言最多,主要包括对管控城市的祈福、对复工复产的期待和愿望;在国际相关类谣言中敌意和中立类型的谣言最多,敌意类型的国际谣言主要包括一些病毒溯源阴谋论、二次输入的猜测,中立型谣言主要包括对国外疫情防控措施的讨论和调侃.这反映了不同主题类型谣言背后的社会心理,对于病毒信息、病毒传播类谣言,出于恐慌心理,人们往往“宁可信其有,不可信其无”,是一种恐慌心理下急于求证的自我保护机制^[24].对于疫情防控类谣言人们面对疫情长期的压抑,倾向于寻求心理慰藉,因此出现较多对于复工复产、医护人员奔赴前线的善意类谣言.而对于国际类谣言往往充斥着攻击性言论,善于挑起民族主义情绪^[19],因此出现大量敌意型情感.

3) 谣言情感类型与特征词.我们进一步利用词云分别展示不同情感下的疫情谣言的词语分布,分析不同情感疫情谣言的关键词和用词特点^①.具体分析为:

① 如图 6(a)所示,敌意类型谣言主要包括“口罩”

“绿卡”“女子”“捐赠”“物资”“顺丰”等主题词.对比文本,我们发现情绪的表达式大都使用类似“这么猖狂”“这么无耻”“根本没有考虑我们的公民”“没有任何隔离措施”“你们都干了些什么?”“你们的良心过得去?”的语句来加强语气、增加情绪激烈程度、增强

情绪感染力.可以看出,敌意类型的谣言一方面内容上通过与国家利益关系、群体冲突、物资等敏感话题和涉及普通民众安全的话题相关联,另一方面语言上利用具有强烈、夸张感情色彩的词语加深语言的情感张力,煽动民众情绪.

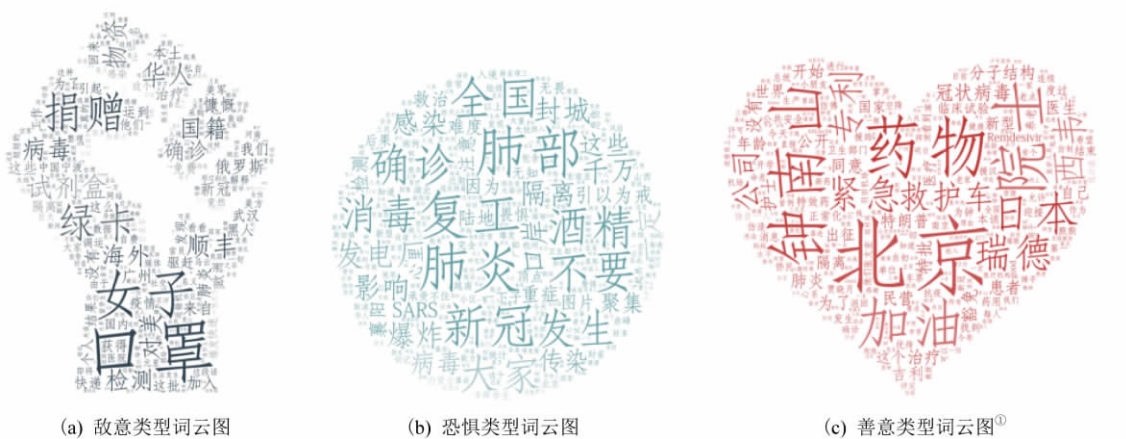


Fig. 6 Word cloud of different sentiment types rumor

图6 不同情感谣言词云图

② 如图6(b)所示,恐惧类型谣言主要包括“复工”“酒精”“新冠肺炎”“肺部”“确认”等主题词.为了强调恐慌情绪,谣言中常使用程度副词,例如“一定不要”“千万注意”“引以为戒”等,以达到渲染恐慌氛围、引起人们注意的作用.可以发现,恐惧类型的谣言一方面内容上大多涉及死亡相关的恐吓性新闻,另一方面用词上利用警戒性词语引起关注,耸人听闻,达到扩散的目的.

③ 如图6(c)所示,善意类型谣言主要包括“北京”“钟南山”“加油”“药物”“瑞德西韦”“日本”“救护车”等主题词,内容主要涉及对医护人员的赞赏、对国际援助的感谢等,多利用“加油”等正能量的词语来带动积极情绪.善意类型谣言的产生大多是在信息未经证实的情况下,网友对于不实信息的误传,对社会治安不会造成很大程度的威胁,但是此类谣言的出现也反映了人们在面对疫情时寻求心理寄托和安慰的现象.

3 谣言参与用户分析

在拉斯韦尔的传播模型中,传播者与受众位于传播过程的两端.而微博上的谣言传播是一个复杂的过程.除了造谣者作为原始的传播者之外,当谣言

的受众转发这则谣言,也在无形之中成为了造谣者.此外,辟谣者也是谣言传播的重要角色.正是辟谣者的参与,抑制了谣言的传播.本节我们对造谣者、造谣者与辟谣者3类人群的谣言接触历史、基础画像、个体情绪、自网络属性以及历史微博内容进行分析,以期对参与谣言传播的各类人群获得更为深入的认识.

本节我们采用降采样的方式,对微博不实信息举报平台上新冠疫情相关谣言所涉及的3类用户,即造谣者、造谣者和辟谣者,分别随机抽样200人,利用其用户属性数据和历史微博数据,开展谣言用户分析.

3.1 3类人群的基础属性分析

本节基于此前得到的用户数据,我们对用户在过去一年中的原创微博数量、转发微博数量、原创微博平均字数、转发微博平均字数、用户认证情况、用户影响力指标(用户被关注数与关注数的比值)、发博时间分布等这些基础属性指标展开分析.需要注意的是,由于微博的隐私访问限制,我们只能抓取每位用户关注列表的前20页,这可能导致对用户的关注列表获取不全.但我们认为用户近期的关注者,正反映了用户近期的关注导向,故用此数据进行后续分析.

1) 造谣者原创微博更多,造谣者转发微博更

① 图中将“瑞德西韦”折词成“瑞德”“西韦”2词.

多.在我们随机抽样得到的各 200 位造谣者、造谣者和辟谣者中,在过去一年时间段内,造谣者发表原创微博数量的中位数为 293 条,造谣者为 18 条,辟谣者为 5 条,如图 7(a)所示.方差分析结果显示 3 组人群的差异显著($p < 0.000$).事后检验显示,造谣者的原创发博量显著高于另外两者,而造谣者和辟谣者之间则没有显著差异.对于转发微博数量而言,造谣用户的转发微博数量则是显著高于另外两者($p < 0.000$),造谣者在过去一年中转发微博数的中位数

为 399 条,造谣者为 35 条,辟谣者为 8 条,如图 7(b)所示.

2) 造谣者发微博字数更多.如图 7(c)所示,三者的原创微博平均字数有显著差异($p < 0.000$),造谣者的原创微博字数(中位数为 81)高于造谣者(中位数为 42),后者又高于辟谣者(中位数为 32).根据图 7(d)进一步发现,辟谣者在转发微博的平均字数上也显著低于另外两者($p < 0.001$),其中位数为 101,而造谣者为 124,造谣者为 116.

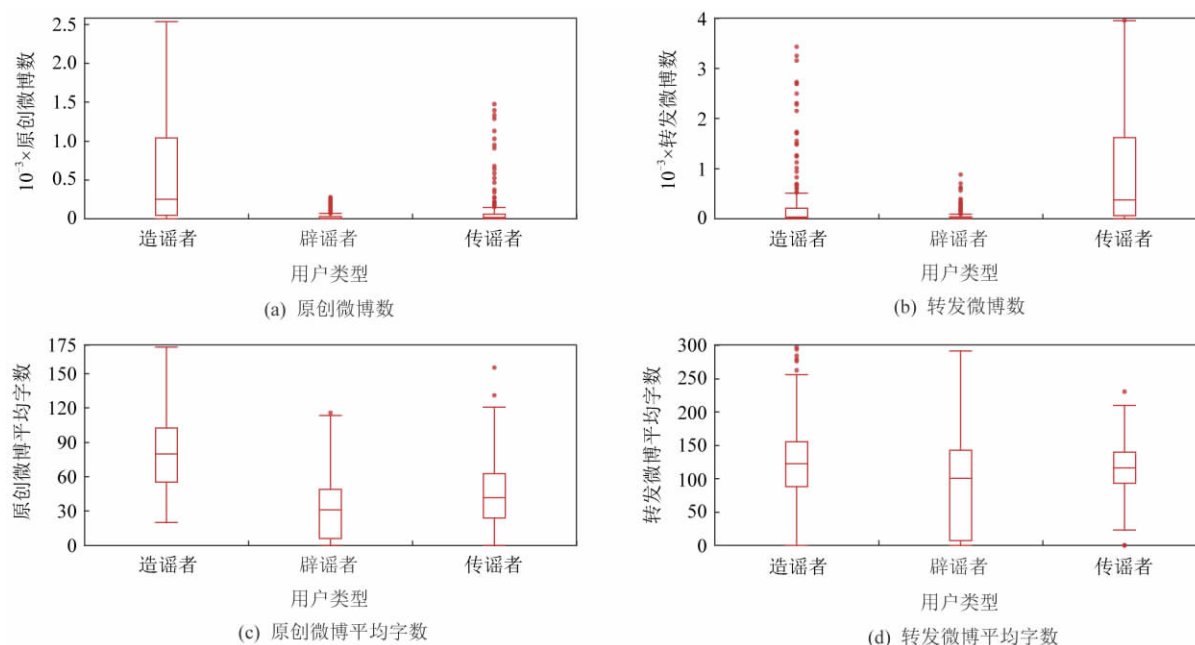


Fig. 7 The number of original Weibo and forwarded Weibo, the average length of words of original Weibo and forwarded Weibo

图 7 原创微博数和转发微博数、原创微博平均字数和转发微博平均字数

综合图 7 可知,造谣者偏好发表原创微博,且原创与转发的博文数字较多;而过往更偏好转发的用户,则在接触了此次新冠疫情相关谣言后,以转发行为再一次参与到了谣言的传播之中.与这二者相比,辟谣者发博数量、发博字数都更少,展现出一种“沉默寡言”的用户形象和“冷静谨慎”的发言姿态.

3) 造谣者上午更活跃.从用户在一天的发微博比例分布图 8 中可以看出,造谣者在上午 6 点到 11 点最为活跃.相比之下,造谣者在午后 13 至 15 时阶段性活跃后,在傍晚再度活跃起来,并在深夜 24 点左右达到发微博数量的峰值.辟谣者的活跃时间有 2 个小峰值,分别在 12 时至 14 时与 22 时至次日

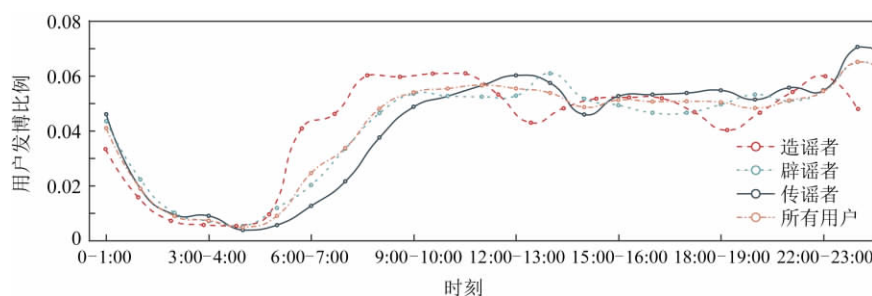


Fig. 8 The distribution of the creation time of Weibo

图 8 3 类用户在一天的发微博比例分布图

凌晨1时.这一定程度反映造谣者倾向于在早上发表言论,其可能有引导当日舆论的意图.

4) 造谣者影响力远大于其他人群.从图9所示的3类人群被关注数和关注者的比例看来,较造谣者、辟谣者来说,造谣者的相对影响力是显著更大的.

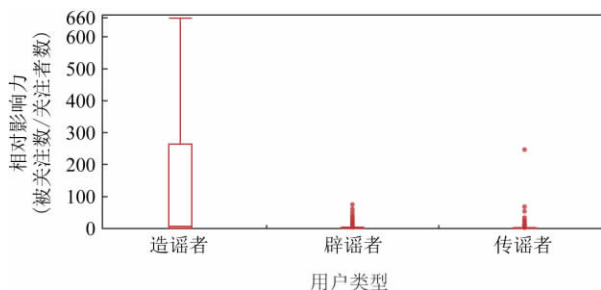


Fig. 9 The boxplot of the ratio of followers and followees

图9 3类用户被关注数与关注者数的比例图

基于分析可以看出,微博虽然日益被诟病为“泛娱乐平台”,但其中意见领袖的引导机制本质上没有发生变化,有影响力的造谣者仍然是那些本身把握信息影响力优势的人,他们或出于有意“蹭热点”,或是无意地轻信了错误信息而将谣言引入微博空间之中.

3.2 3类人群的关注主题分析

本节基于用户的历史发博数据,包括原创博文

内容和转发微博的转发内容,我们进行微博主题层面的分析.

我们利用STM^[25-26] (structural topic model)模型对用户自疫情发生以来的疫情相关微博进行主题分类.STM模型是一种无监督文本主题模型,可以通过纳入协变量在主题建模中取得更好地效果.为了更好地训练STM主题模型,我们爬取微博平台中2020年1—5月疫情相关的微博数据^①,随机抽取其中10%的微博作为训练文本,共1211760条微博.我们设置发博时间作为协变量,将疫情相关微博的主题划分为“缅怀与致敬”“国内疫情形势”“控制疫情扩散”“医疗救治”“复工复产”“海外疫情形势”“全球疫情影响”“新冠科学研究”“社会互助行动”“生活日常记录”10个类别.进一步,我们同样通过关键词抽取3类用户的疫情相关微博,分别得到26515条造谣者微博、1006条辟谣者微博和25229条传谣者微博,并利用本节在大规模疫情相关微博中训练的STM模型计算每位用户的微博隶属于各个主题的概率分布.

辟谣者更偏好新冠科学研究.如图10所示,我们发现辟谣者发布的微博中“新冠科学研究”主题占比最高,且显著高于造谣者和传谣者.由此可以推测,辟谣者对新冠科学研究话题的偏好,显示出相对更高的科学素养,对谣言有更强的甄别能力.

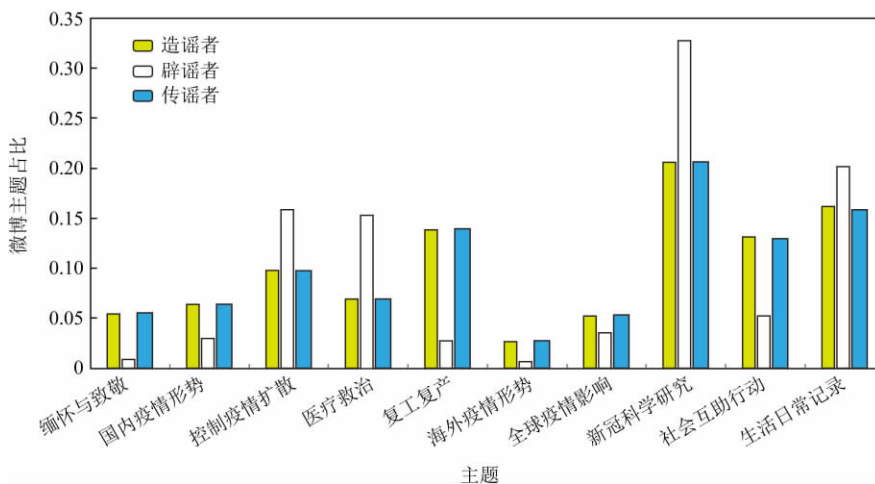


Fig. 10 The bar of topic distribution of users on Weibo

图10 3类用户的微博主题分布图

3.3 3类人群的个体情绪分析

本节我们将针对用户的历史微博数据对3类人群的个体情绪进行探究.

我们使用基于BERT^[28] (bidirectional encoder representations from transformers)与TextCNN^[29]结合的方式对3类用户的历史微博数据进行情感

① 我们参考官方报道^[27]并结合微博平台在不同阶段对新型病毒的命名,选取“肺炎”“SARS”“新冠”“冠状”以及“COVID-19”作为疫情相关微博的抽取关键词.

分类.BERT 是谷歌团队提出的预训练语言模型,被广泛用于文本分类、问答等下游任务中^[30-31].TextCNN 模型由 Kim 提出,其核心思想是利用卷积神经网络捕获文本的局部特征并用于最终的文本分类.由于疫情相关文本的领域特殊性,与日常微博表达不同,我们使用北京市经济和信息化局和中国计算机学会等单位发布的“疫情期间网民情绪识别”数据集^①训练情感分类模型.该数据集包含 10 万条 2020-01-01—2020-02-20 期间的疫情微博,每条微博被人工标注为 3 类情感之一,3 类情感分别为积极、中性和消极.我们使用这 10 万条数据进行训练和交叉验证,训练得到的模型在比赛测试集上的预测 F1 值为 73.6%,相比于只基于 BERT 的模型提升了 4.1%.

造谣者和传谣者情绪更为负面.基于这一模型,以用户发表的负面微博数占总微博数的比例作为对用户负面情绪倾向程度的测量标尺,我们对 3 类人群的负面情绪倾向进行统计分析.3 类人群负面情绪倾向分布如图 11 所示.就中值而言,造谣者负面情绪的比例高于传谣者,传谣者负面情绪的比例高于辟谣者.方差分析结果显示,造谣者的负面情绪程度显著高于传谣者($p=0.0169<0.05$).

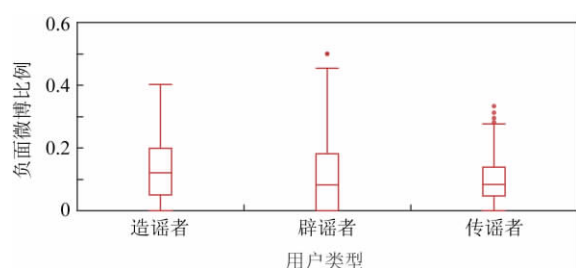


Fig. 11 The boxplot of negative emotion of users on Weibo

图 11 3 类用户的负面情绪倾向图

既有的研究已经讨论了不同情感在微博中传递的效果差异^[32],愤怒相较于喜悦、甚至悲伤都更有传染性.一方面,造谣者由于其本身的负面情绪,可能使得其产生的谣言更具感染力.另一方面,根据 Na 等人的研究^[33],在公共健康危机中,如果流言所激发的情绪与受众本身所处的情绪一致,则受众会更倾向于相信这一流言.面对此次新冠肺炎疫情,长期身处于负面情绪的社交媒体用户则更为轻易被负面谣言所欺骗和感染,并成为其中的一名传播者.

① <https://www.datafountain.cn/competitions/423/datasets>

3.4 3 类人群的自网络属性分析

本节我们将基于用户的关注列表数据,对 3 类人群构建自网络来探究其所处网络结构.自网络(Ego network)由中心节点(Ego,“自我”)、与之直接连接到的节点(Alter,“他者”)以及“他者”之间的联系(如果这些连边存在)组成.因此,我们抓取了 3 类用户的关注列表以及这些受关注用户的下一级关注列表来构建其二级网络结构.

大量既有的研究从社会网络视角研究了谣言传播现象,其考察的网络属性包括连边强度和节点属性等基础网络属性^[34-37].相比之下,Burt 经典的结构洞理论^[38]对这一问题提供了更为深层次的洞见.根据这一理论,处于结构洞位置的节点为网络中的其他节点提供了非冗余的联系,因而拥有更大的信息优势.基于包括脸书、推特在内的 7 个真实网络开展的实验表明,若是移除结构洞位置的重要用户,将会阻断多达 24% 的信息传播^[39].这表明,结构洞理论在社会化媒体上依旧适用.因此,本文将基于结构洞理论,探究在本次疫情谣言传播中,3 类用户在社交媒体中所处网络结构的差异.

辟谣者的自网络聚集系数更高.我们以局域聚集系数(local clustering coefficient)作为测量用户自网络结构的关键指标^[40],就网络属性对谣言用户参与行为的影响展开探究.如图 12 所示,对辟谣者、造谣者和传谣者的局域聚集系数分布进行对比,结果显示,就中值而言,辟谣者的聚集系数高于造谣者,造谣者的聚集系数高于传谣者.事后检验显示,辟谣者和传谣者的网络结构聚集系数存在显著差异($p=0.0078<0.01$).

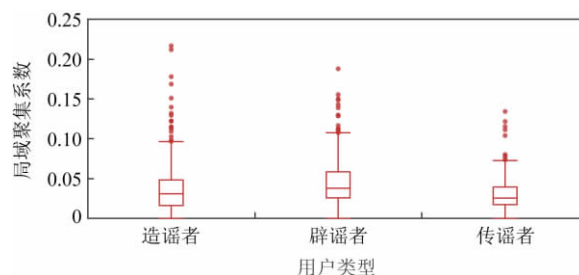


Fig. 12 The boxplot of local clustering coefficient of users

图 12 3 类用户的局域聚集系数分布图

这扩充了我们对于结构洞的认识,即结构洞属性高的节点尽管拥有信息优势,但也容易受到虚假

信息的影响.辟谣者的自网络聚集系数更高反映出辟谣者所处网络更聚集稳定,其关注者之间往往也相互关注,形成一个较为闭合的关注网络.而造谣者和传谣者的自网络结构则更为松散,这可能使得他们的信息获取渠道更加开放,但也同时失去了信息来源之间较为稳定的信任关系,失去信息“交叉验证”的质量保证.

4 谣言舆情分析

传播效果指的是来自传播者的讯息对受众产生的效果.本文谣言微博所引发的舆情,无疑是谣言传播的一种重要效果.在本节中,我们通过对谣言的评论进行情感分析来研究谣言引发的舆情,并探索其与谣言主题、关键词以及涉及地区的关系,最后进一步探索公众舆情演变的原因.

4.1 谣言评论情感整体分析

本节我们使用 3.3 节中所述模型对谣言的评论进行情感分类.图 13 展示了谣言评论情感倾向的整体分布.

近半数谣言评论持中性态度.在剔除“转发微博”等微博自动生成的关键词后,中性评论仍然在 3 类情感中占据主导地位,达到了 48%.通过归纳整理,

我们将中性评论概括为 3 种主要类型:冷静辟谣型评论、质疑打探型评论和无关评论.中性评论占比较大反映出相当一部分网络用户对于谣言具有一定的辨别能力,在消息未证实前持观望态度,在辟谣后会要求博主进行澄清或删除,以防止更多用户被误导.

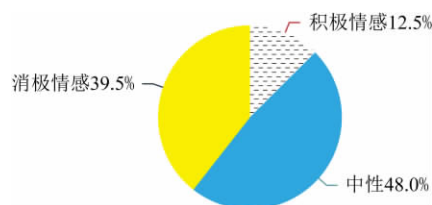


Fig. 13 Sentiment distribution of COVID-19 rumor comments

图 13 疫情谣言评论情感分布

消极评论多于积极评论.消极评论可概括为恐慌、矛盾争议和激烈辟谣 3 类,积极评论则主要包括加油祝福、赞美英雄和善意调侃 3 种类型.除中性评论外,消极情感评论数量整体领先于积极情感评论,前者数量约为后者的 3 倍.图 14 展示了每周积极和消极的评论在该周总评论数中的占比随时间的变化.可以看出,消极情感在疫情各个阶段均基本占据主导地位.

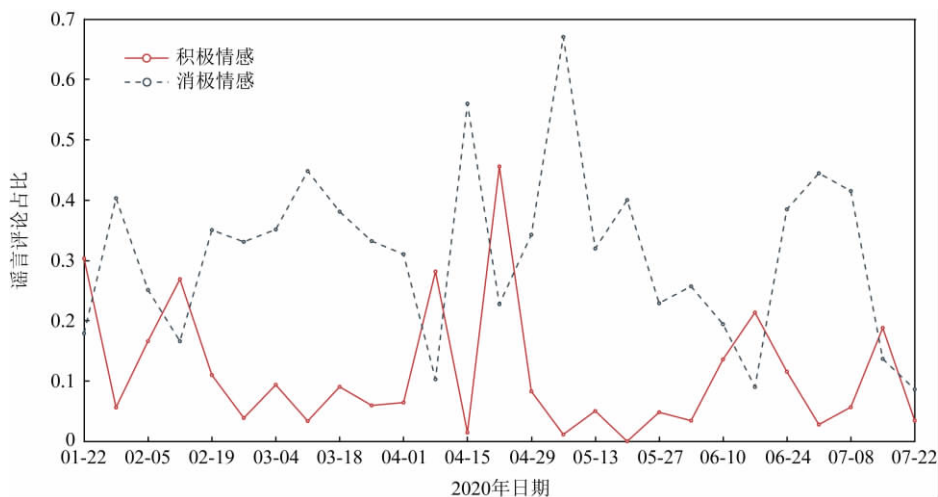


Fig. 14 Sentiment of COVID-19 rumor comments over time

图 14 疫情谣言评论情感随时间变化

4.2 谣言主题与评论情感分析

本节我们对第 2 节划分的 5 类主题谣言的评论进行统计和分析(分类标准如表 2 所示).5 类主题中,疫情防控、国际相关和其他类主题谣言获得的评论数量较多,分别为 33 616,17 283 和 27 132 条.病

毒传播和病毒信息类谣言的评论数量较少,分别为 591 和 249 条.

奇闻异事和名人轶事热度最高.为了进一步了解用户对各个主题谣言的关注程度,我们对五大主题谣言的平均评论数进行统计.结果显示,其他类主题

谣言的平均评论数最高,平均每条谣言获得了 335 条评论,这表明用户对于其他类中占主要部分的坊间奇闻异事、名人轶事十分热衷,其次是疫情防控主题谣言,平均每条谣言被评论 146.6 次,这表明人们对于疫情的防控情况较为关注,会通过评论来对一些表达利好消息的谣言表达喜悦和支持,对描述违反防控相关政策的谣言予以谴责。

疫情防控最正面,病毒传播最负面,5 类主题的谣言评论情感分布如图 15 所示。其中疫情防控相关谣言的积极情感评论占比最多,为 12.99%,这反映了大众对防疫政策的理解和支持。病毒传播相关谣言的消极情感评论占比较多,达 44.16%,原因可能是人们对于新冠病毒的传播持有警惕心理,此类谣言容易引发公众对于疫情蔓延的恐惧和担忧。

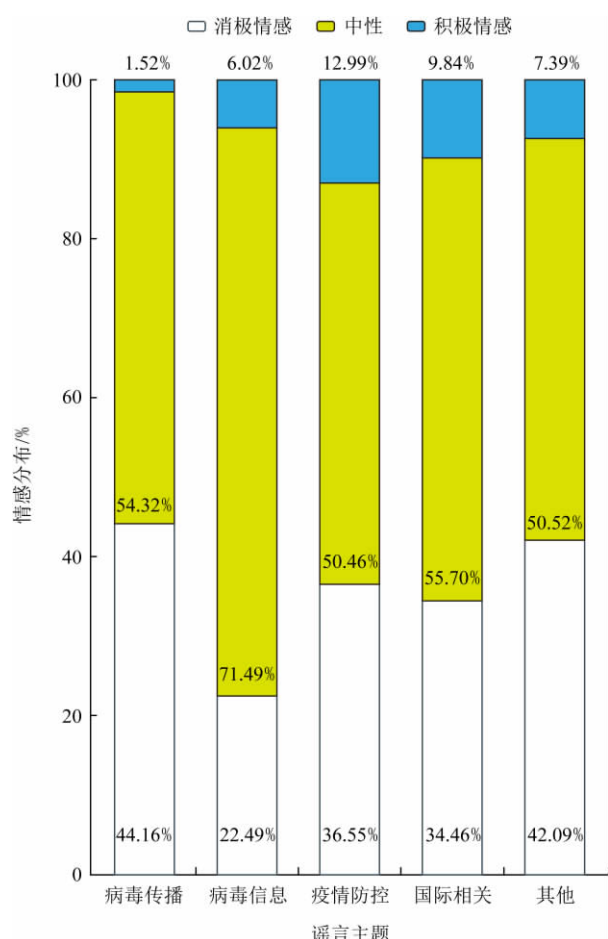


Fig. 15 Sentiment distribution of comments under different topics

图 15 5 类主题对应的评论情感分布

4.3 谣言关键词与评论情感分析

随着新冠疫情的出现,一些疫情相关的特色词汇成为了公共讨论的焦点,也为谣言制造者提供了

素材。本节中我们统计了在谣言中出现频率位居前 30 且具有疫情特色的关键词对应的谣言评论数量和情感指数,如图 16 所示。在选取过程中对近义词进行了合并处理,例如“瑞德西韦”“双黄连”等词被归入药品类。这里,我们将情感指数定义为谣言评论情感倾向(消极为-1,中性为 0,积极为 1)的均值。

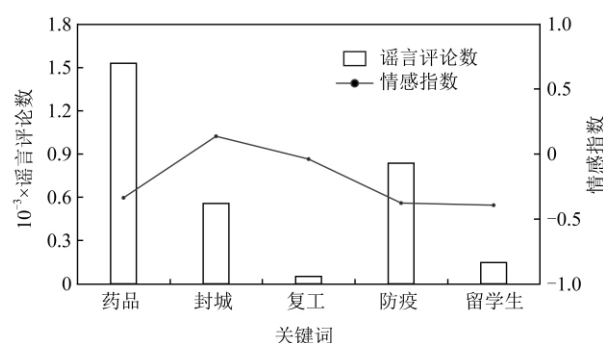


Fig. 16 Sentiment distribution of rumors with different keywords

图 16 含有不同关键词谣言的评论情感分布

从数量上看,“药品”和“防疫”谣言最受关注。在药品相关的谣言中,获得转评最多的谣言大多与瑞德西韦有关,这些转评的高峰均出现在 1 月底到 2 月初,彼时人们对于疫情了解较少,更容易轻信治疗病症的谣言。防疫则一直是公众关心的热点,相关谣言主要围绕国内外的疫情防控形势展开。

“封城”谣言引发积极舆论。从情感指数上看,“封城”谣言引发了最多的正面情感。“封城”相关的谣言评论主要来自 1 月 25 日“襄阳封城”谣言,突出湖北为防控疫情做出的牺牲和奉献,唤起公众的感动情绪。因此尽管“封城”意味着抗疫形势愈发严峻,网友仍然留下大量“感动祝福”类型的评论。这一定程度上反映了谣言的情绪感染作用。

4.4 谣言涉及地区与评论情感分析

本节中我们对涉及不同地区的谣言评论进行统计和分析。图 17 为武汉和国内 4 个一线城市的谣言评论数量对比。通过柱形图可以看出,武汉作为 1~3 月疫情的中心,获得了最多的关注,之后依次为北京、上海、广州、深圳,这表明地区谣言受到的关注不仅与其受影响程度,同时与地区的政治、经济等发展水平具有相关性。

从不同地区的情感指数对比可以看到,北京相关的谣言评论平均情感指数最高(-0.03),广州相关的谣言评论平均情感指数最低(-0.53)。与北京相关的积极评论主要来自疫情防控相关的利好消息。公众对疫情防控中英雄人物的敬佩和崇拜某种

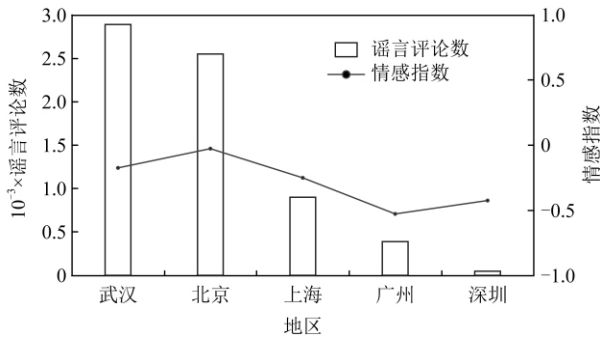


Fig. 17 Comment sentiment distribution regarding different cities

图 17 不同城市相关谣言评论情感分布

程度上使其对“救世”类的谣言更宽容,在评论时会降低语言的攻击性.与广州相关的负面转评则主要围绕防疫形势恶化消息,此类谣言利用了普通民众对疫情、种族等问题的恐慌情绪,在短时间内获得大量传播,一度成为新闻热点.偏负面的谣言评论情感也反映出一些潜在的社会问题,例如这里涉及的民众对广州外国人聚居现象的观点之争,值得引起相关部门的重视.

4.5 谣言评论中的情感演变

在谣言不断传播,评论数量不断增长的过程中,谣言引发的舆情往往会发生时序的变化^[41-42].因此,本节从情感倾向随时间的变化入手,剖析疫情期间谣言的情感演变.

谣言评论情感倾向于向负面演变.根据谣言评论的初始情感和最终情感,可以将谣言划分为 4 种类型,即始终保持积极和消极、初始为积极最终为消极以及初始为消极最终转向积极谣言数量,其分布如表 4 所示.由表 4 可知,疫情相关谣言中初始评论为负面、最终转化为正面谣言的比例只占了全部初始为负面谣言的 8.05%,远低于初始正面向负面转化的比例 56.60%,表明谣言引发的评论情感倾向于向负面演变.

Table 4 Number of Comments with Different Initial or Final Sentiment Orientation

表 4 不同初始和最终情感倾向的评论数量

初始\最终情感倾向	积极	消极
积极	23	30
消极	12	137

4.6 辟谣对评论情感演变的影响

已有研究^[41]证明辟谣对谣言引发舆情演变具

有重要影响,其发现辟谣后公众情感通常会由负面转向正面或者中立.因此,本节我们进一步探究辟谣对于疫情相关谣言的评论情感演变的影响.

在所有微博谣言评论中,存在着一类特殊的评论,这些评论的发布者试图通过发布评论表达“此条微博是谣言”的含义,我们称之为辟谣评论.辟谣评论的出现能够使之后的评论者意识到该微博可能为不实信息,从而影响后继评论的情感倾向.通过标记出辟谣评论的出现时间,可以将谣言评论在时间上分为 2 个类别:辟谣前评论与辟谣后评论.我们使用关键词筛选的方式提取辟谣评论,选取的关键词如表 5 所示.如果一条评论中包含任意一个关键词(keywords),且不含有任何的停用词(stop words)时,我们就认为这条谣言评论为辟谣评论.

Table 5 Keywords of Refutation Comments

表 5 辟谣评论关键词

类别	词或短语
关键词	假的 辟谣 举报 网警 造谣 传播 谣言
停用词	假的? 真的假的

通过这方法,我们对所有评论数量大于 10 条的谣言提取其辟谣评论.根据谣言中辟谣评论的出现时间,我们将谣言分为 3 类:发布至今未出现辟谣评论的谣言、第 1 条评论就是辟谣评论的谣言、其他(即一般谣言).3 类的谣言的数量如表 6 所示(注:受限于微博时间的表示形式,评论的发布时间只精确到分钟).通过表 6 可以发现,大部分谣言(74.11%),即“一般谣言”,都经历了在公众中发酵之后被辟谣的过程,只有少数谣言(9.14%)一经发布即被公众识别.这一方面体现了公众对于疫情相关谣言的辨别力不高,但同时也说明了谣言检测中谣言评论的重要作用,可以辅助公众对信息的甄别和自动化谣言检测.

Table 6 Number of Three Types of Rumors Divided According to the Time that Refutation Comments Occur

表 6 根据辟谣评论出现时间划分的 3 类谣言数量

谣言类型	无辟谣评论	第 1 条评论为辟谣评论	其他
谣言数量	33	18	146

辟谣使得谣言评论情感向负面演变.对于“一般谣言”类,我们记录其对应的最早辟谣评论和最晚辟谣评论,比最早辟谣评论更早的评论被划分为辟谣前的评论,比最晚辟谣评论更晚的评论被划分为辟谣后的评论.通过对辟谣前后评论的情感倾向进行

统计分析得到表 7,可以发现,与 Zeng 等人^[41]的研究不同,谣言的评论中正面情感的占比从辟谣前的 18.3%下降到了辟谣后的 11.9%,而负面情感的占比从辟谣前的 31.2%上升到了辟谣后的 33.9%。通过卡方检验验证,在置信水平为 0.005 的条件下,是否辟谣与谣言评论的情感演变具有显著相关性,这种相关性表现在正面评论的显著减少和负面评论的显著增加。

Table 7 Number of Comments of Different Sentiment Types Before and After the Refutation

表 7 辟谣前后不同情感类型的评论数量

谣言情感	积极	中立	消极
辟谣前	1 462	4 029	2 489
辟谣后	145	663	414

辟谣导致舆情向负面变化的原因也是直观的:用户在识破谣言后感受到被欺骗,倾向于表达负面情感。在新冠疫情的背景下,这一点尤为突出,谣言被辟谣后造谣者和传谣者常常会被大量的负面评论攻击,引发负面舆情,一定程度上污染网络信息环境。因此,对于突发公共新冠疫情类似的公众事件而言,在辟谣的同时,也应注意疏导谣言被揭发后引发的负面情绪级。

5 总 结

本文基于新浪微博社交媒体平台,从信息传播的基本要素出发,对新冠肺炎疫情相关谣言传播展开较为全面的量化分析,包括谣言的传播内容分析、谣言的参与用户分析、谣言引发舆情分析。据我们所知,这是首次从信息传播的各个基本构成要素角度对新冠肺炎疫情相关谣言传播展开量化分析,对新冠肺炎疫情相关的谣言传播有了更全面深入的认识。此次新冠肺炎疫情是一次全球性的突发公共事件,针对突发公共事件的谣言传播展开研究不论是对于学术研究还是社会治理都具有重要意义。

本文开展的疫情相关社交媒体谣言分析基于新浪微博社区管理中心中被举报和核查的不实信息数据,其难免存在覆盖度不全和人工判断的误差。同时对谣言的判断具有不确定性,随着更多信息的引入可能会发生变化。因此,本文对疫情期间的社交媒体谣言研究聚焦于对其传播结构和传播要素的分析,不对谣言的真实性做探讨。未来可以进一步探索不同传播平台的谣言传播差异,实现对谣言不同传播

途径的分析。同时可以进一步将本文的量化分析结论应用到谣言的自动检测技术中,当突发公共事件再次来临时,助力于社交媒体谣言的自动检测。

作者贡献声明:陈慧敏负责文章选题、实验设计及文章撰写工作;金思辰主要参与谣言内容分析工作,林微、朱泽宇主要参与谣言用户分析工作,全凌波主要参与谣言舆情分析工作。金思辰、林微、朱泽宇、全凌波 4 位排名不分先后,同等贡献;刘一凡主要负责数据的收集和整理;姜维翰参与谣言舆情分析工作;叶奕宁参与谣言用户分析工作;刘知远、孙茂松、金兼斌老师为论文开展提供了宝贵的经验,并对论文进行修改。

参 考 文 献

- [1] Institute of Linguistics, Chinese Academy of Social Sciences. Contemporary Chinese Dictionary [M]. Beijing: Commercial Press, 2012 (in Chinese)
(中国社会科学院语言研究所词典室. 现代汉语词典[M]. 北京: 商务印书馆, 2012)
- [2] Cai Jing. Rumor: Social Transmission in the Shadow [M]. Beijing: China Radio and TV Press, 2008 (in Chinese)
(蔡静. 流言: 阴影中的社会传播[M]. 北京: 中国广播电视出版社, 2008)
- [3] Allport G W, Postman L. The Psychology of Rumor [M]. Oxford: Henry Holt, 1947
- [4] Peterson W A, Gist N P. Rumor and public opinion [J]. American Journal of Sociology, 1951, 57: 159-167
- [5] Zubiaga A, Aker A, Bontcheva K, et al. Detection and resolution of rumors in social media: A survey [J]. ACM Computing Surveys, 2018, 51(2): No.32
- [6] Isaac C H F, Fu K W, Chan C H, et al. Social media's initial reaction to information and misinformation on Ebola, August 2014: Facts and rumors [J]. Public Health Reports, 2016, 131(3): 461-473
- [7] Grinberg N, Joseph K, Friedland L, et al. Fake news on Twitter during the 2016 US presidential election [J]. Science, 2019, 363(6425): 374-378
- [8] Zhao Junfeng, Jin Taijun. Survival Logic of rumors in public crisis—An analytical framework of rumors [J]. Jiangsu Social Sciences, 2013 (1): 130-135 (in Chinese)
(赵军锋, 金太军. 论公共危机中谣言的生存逻辑——一个关于谣言的分析框架[J]. 江苏社会科学, 2013 (1): 130-135)
- [9] Wang Canfa. Construction and destruction of rumor communication patterns in public emergencies [J]. Modern Communication, 2010 (6): 45-58 (in Chinese)
(王灿发. 突发公共事件的谣言传播模式建构及消解[J]. 现代传播, 2010 (6): 45-48)

- [10] Zhou Xiaohong. The Mis-transmission: A social psychological analysis of the prevalence of rumors and gossip about SARS [J]. Sociological Studies, 2003 (6): 43-54 (in Chinese)
(周晓虹. 传播的畸变——对“SARS”传言的一种社会心理学分析[J]. 社会学研究, 2003 (6): 43-54)
- [11] Guo Xiaolan, Dong Tiance. Rumors, medium and collective actions: A case study of three rumors causing panics [J]. Modern Communication, 2013 (9): 58-92 (in Chinese)
(郭小安, 董天策. 谣言, 传播媒介与集体行动——对三起恐慌性谣言的案例分析[J]. 现代传播, 2013 (9): 58-62)
- [12] Islam M S, Sarkar T, Khan S H, et al. COVID-19-related infodemic and its impact on public health: A global social media analysis [J]. The American Journal of Tropical Medicine and Hygiene, 2020: 103(4): 1621-1629
- [13] Cinelli M, Quattrocioni W, Galeazzi A, et al. The covid-19 social media infodemic [J]. ArXiv preprint arXiv: 2003.05004, 2020
- [14] Qian Yingfei. Rumors during COVID-19: An analysis based on a sample of 437 rumors from Tencent Fact Check platform [J]. Radio & TV Journal, 2020 (8): 132-133 (in Chinese)
(钱迎飞. 新冠肺炎疫情期间的谣言研究——基于“较真”辟谣平台 437 条样本的分析[J]. 视听, 2020 (8): 132-133)
- [15] Jiang Sujia. Infodemic: Study on the spread of and response to rumors about COVID -19 [J]. Study on Science Popularization, 2020, 15(1): 70-78(in Chinese)
(江苏佳. 信息疫情: 新冠肺炎疫情谣言传播及应对研究[J]. 科普研究, 2020, 15(1): 70-78)
- [16] Lasswell H D. The structure and function of communication in society [J]. The Communication of Ideas, 1948, 37(1): 136-139
- [17] Liu Zhiyuan, Zhang Le, Tu Cunchao, et al. Statistical and semantic analysis of rumors in Chinese social media [J]. Scientia Sinica Informations, 2015, 45(12): 1536-1546 (in Chinese)
(刘知远, 张乐, 涂存超, 等. 中文社交媒体谣言统计语义分析[J]. 中国科学: 信息科学, 2015, 45(12): 1536-1546)
- [18] Yin Xiaochun. A case study on the spread of epidemic rumors based on new media [D]. Shanghai: Shanghai Jiao Tong University, 2014 (in Chinese)
(殷晓椿. 基于新媒体的疫情谣言传播个案研究[D]. 上海: 上海交通大学, 2014)
- [19] Shi Aidong. The grammar of Internet rumors [J]. National Art, 2016(5): 109-115 (in Chinese)
(施爱东. 网络谣言的语法[J]. 民族艺术, 2016 (5): 109-115)
- [20] Zhu Songmei, Ren Yan. The causes of Weibo rumors and the mechanism of refuting rumors—Taking the rumors after the 2011 Japanese earthquake as an example [J]. Young Journalist, 2011 (17): 91-92 (in Chinese)
(朱松梅, 任雁. 微博谣言产生的原因和辟谣机制——以 2011 年日本震后谣言为例[J]. 青年记者, 2011 (17): 91-92)
- [21] Liu Jun. Promotion of communication art in public emergencies: Information and public opinion are based on COVID-19 in Wuhan [J]. Modern Insight, 2020 (2): 8-13 (in Chinese)
(刘俊. 突发公共事件中的“传播艺术”提升论要: 信息与舆情——基于武汉新冠肺炎疫情的示例[J]. 现代视听, 2020 (2): 8-13)
- [22] Yao Siyu. The authenticity verification and dissemination analysis of microblog hot news [D]. Dalian: Dalian University of Technology, 2016 (in Chinese)
(姚思宇. 微博热点新闻的真实性验证及传播分析[D]. 大连: 大连理工大学, 2016)
- [23] Ruan Zhangqiong, Yin Liangrun. Types and discourse of microblog rumors: Content analysis of 307 microblog rumors [J]. Contemporary Communication, 2014 (4): 77-78 (in Chinese)
(阮璋琼, 尹良润. 微博谣言的类型与话语焦点——基于 307 条微博谣言的内容分析[J]. 当代传播, 2014 (4): 77-78)
- [24] Xing Dan. Social Psychology Analysis: Why Believe in “Bad News”? [J] Party's Regular Activities(Heilongjiang), 2012 (6): 7-8 (in Chinese)
(邢丹. 社会心态透析: 为何偏信“坏消息”? [J]. 党的生活(黑龙江), 2012 (6): 7-8)
- [25] Roberts M, Stewart B, Tingley D. STM: R Package for Structural Topic Models [J]. Journal of Statistical Software, 2019, 91(2): 1-40
- [26] Roberts M, Stewart B, Tingley D, et al. Structural Topic Models for Open-Ended Survey Responses [J]. American Journal of Political Science, 2014, 58(4): 1064-1082
- [27] Hou Min, Teng Yonglin. A keyword-based description of China's fight against COVID-19 [J]. Chinese Journal of Language Policy and Planning, 2020, 5 (3): 50-61 (in Chinese)
(侯敏, 滕永林. 词述中国战“疫”[J]. 语言战略研究, 2020, 5 (3): 50-61)
- [28] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [C] //Proc of the 2019 Conf of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1. 2019: 4171-4186 arXiv: 1810.04805v1 [cs.CL]
- [29] Kim Y. Convolutional neural networks for sentence classification [C] //Proc of the 2014 Conf on Empirical Methods in Natural Language Processing (EMNLP). Stroudsburg, PA: ACL, 2014: 1746-1751
- [30] Qi Qin, Hu Wenpeng, Liu Bing. Feature projection for improved text classification [C] //Proc of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2020: 8161-8171
- [31] Wei Yang, Xie Yuqing, Lin Aileen, et al. End-to-End Open-Domain Question Answering with BERT [C] //Proc of the 2019 Conf of the North American Chapter of the Association for Computational Linguistics (Demonstrations). Stroudsburg, PA: ACL, 2019: 72-77

- [32] Fan Rui, Zhao Jichang, Chen Yan, et al. Anger is more influential than joy: Sentiment correlation in Weibo [J]. *PloS One*, 2014, 9(10): No.e110184
- [33] Na K, Garrett R K, Slater M D. Rumor acceptance during public health crises: Testing the emotional congruence hypothesis [J]. *Journal of Health Communication*, 2018, 23(8):791-799
- [34] Lai G, Wong O. The tie effect on information dissemination: the spread of a commercial rumor in Hong Kong [J]. *Social Networks*, 2002, 24(1): 49-75
- [35] Golbeck J. Trust and nuanced profile similarity in online social networks [J]. *ACM Transactions on the Web*, 2009, 3(4): 1-33
- [36] Ji Jiaojiao, Chao Naipeng, Ding Jieyu. Rumormongering of genetically modified (GM) food on Chinese social network [J]. *Telematics & Informatics*, 2019, 37: 1-12
- [37] Dey P, Pyne S, Roy S. Information spreading in Online Social Networks: A case study on Twitter network [C] // *Proc of the 18th ACM Int Symp on Mobile Ad Hoc Networking and Computing*. New York: ACM, 2017: 1-2
- [38] Burt R S. *Structural Holes: The Social Structure of Competition* [M]. Cambridge, MA: Harvard University Press, 2009
- [39] Xu Wenzheng, Li Tong, Liang Weifa, et al. Identifying structural hole spanners to maximally block information propagation [J]. *Information Sciences*, 2019, 505: 100-126
- [40] Meng Jingbo, Chung M, Cox J. Linking network structure to support messages: Effects of brokerage and closure on received social support [J]. *Journal of Communication*, 2016, 66(6): 982-1006
- [41] Zeng Runxi, Zhu Di. A model and simulation of the emotional contagion of netizens in the process of rumor refutation [J]. *Scientific Reports*, 2019, 9(1): No.14164
- [42] Sun Yan. Rumorstorm: Research on the crisis of network opinion after disasters [J]. *Journalism & Communication*, 2011 (5): 52-62 (in Chinese)
(孙燕. 谣言风暴: 灾难事件后的网络舆论危机现象研究[J]. 新闻与传播研究, 2011 (5): 52-62)



Chen Huimin, born in 1994. Postdoc at the School of Journalism and Communication, Tsinghua University. She received the PhD degree from the Department of Computer Science and Technology, Tsinghua University, China. Her main research interests include natural language processing and social computing.

陈慧敏, 1994 年生。清华大学新闻与传播学院博士后, 博士毕业于清华大学计算机系。主要研究方向包括自然语言处理和社会计算。



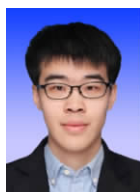
Jin Sichen, born in 1996. PhD candidate. Her main research interests include visualization, natural language processing.

金思辰, 1996 年生。博士研究生。主要研究方向为可视化、自然语言处理。



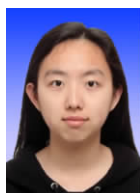
Lin Wei, born in 1998. Master candidate at the School of Journalism and Communication, Tsinghua University. Her main research interests include computational social science and media psychology.

林微, 1998 年生。清华大学新闻与传播学院硕士研究生。主要研究方向为计算社会科学和媒介心理学。



Zhu Zeyu, born in 1997. Master candidate at the School of Journalism and Communication, Tsinghua University. His main research interests include computational social science, communication effects and international communication.

朱泽宇, 1997 年生。清华大学新闻与传播学院硕士研究生。主要研究方向为计算社会科学、传播效果和国际传播。



Tong Lingbo, born in 1998. PhD candidate. Her main research interests include natural language processing, data mining and computational social science.

仝凌波, 1998 年生。博士研究生。主要研究方向为自然语言处理、数据挖掘和计算社会科学。



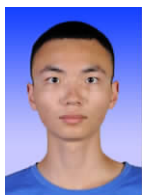
Liu Yipeng, born in 2002. Undergraduate of the Department of Computer Science and Technology, Tsinghua University. His main research interests include natural language processing and linguistics.

刘一芑, 2002 年生。清华大学计算机系在读本科生。主要研究方向为自然语言处理和语言学。



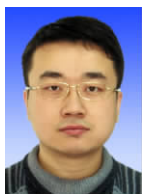
Ye Yining, born in 2001. Undergraduate of the Department of Computer Science and Technology, Tsinghua University. His main research interests include natural language processing, social computation and language generation.

叶奕宁, 2001 年生. 清华大学计算机科学与技术系在读本科生. 主要研究方向为自然语言处理、社会计算和语言生成.



Jiang Weihang, born in 2000. Undergraduate of Beihang University. His main research interests include sentiment analysis and machine learning.

姜维翰, 2000 年生. 北京航空航天大学在读本科生. 主要研究方向为情感分析和机器学习.



Liu Zhiyuan, born in 1984. Associate professor of the Department of Computer Science and Technology, Tsinghua University. His main research interests include natural language processing and social computation.

刘知远, 1984 年生. 清华大学计算机科学与技术系副教授. 主要研究方向为自然语言处理和社会计算.



Sun Maosong, born in 1962. Professor of the Department of Computer Science and Technology, Tsinghua University. His main research interests include natural language processing, Chinese computing, Web intelligence, and computational social sciences.

孙茂松, 1962 年生. 清华大学计算机科学与技术系教授. 主要研究方向为自然语言处理、中文计算、Web 智能和计算社会科学.



Jin Jianbin, born in 1968. Professor of the School of Journalism and Communication, Tsinghua University. His main research interests include empirical studies of media adoption, uses and effects, and science and risk communication.

金兼斌, 1968 年生. 清华大学新闻与传播学院的教授. 主要研究方向为媒介使用、使用与效果、科学与风险传播的实证研究.