



开放科学
(资源服务)
标识码
(OSID)

涉警网络舆情主题发现及情感分析研究

管雨翔 王娟 张鹏

中国人民警察大学网络舆情治理研究中心 廊坊 065000

摘要: [目的/意义] 研究涉警网络舆情主题特点及网民情感变化, 可以为构建良好的警民关系和警察形象提供参考, 具有重要的社会意义和实际价值。[方法/过程] 获取微博平台涉警网络舆情数据, 使用 LDA 主题模型及 TF-IDF 算法对网络舆情数据进行主题发现研究, 基于 Word2Vec 模型构建涉警网络舆情领域情感词典, 进而对网民情感进行分析。[局限] 在案例选取及分析方面不够全面。[结果/结论] 通过 LDA 主题模型及 TF-IDF 算法可以使主题划分更为明显, 基于领域情感词典的情感分析也较为准确, 更好地反映出舆情传播过程中热点话题及网民情感的变化。

关键词: 涉警舆情; 主题发现; 情感分析; LDA 主题模型

中图分类号: G35; TP391

Research on the Theme Discovery and Emotion Analysis of Police-related Internet Public Opinion

GUAN Yuxiang WANG Juan ZHANG Peng

Research Center for Network Public Opinion Governance of CPPU, Langfang 065000, China

Abstract: [Objective/Significance] Studying the theme characteristics of police-related network public opinion and the emotional changes of netizens can provide reference for building a good relationship between the police and the people and the police image, which has important social significance and practical value. [Methods/Processes] The online public opinion data related to police in Weibo platform were obtained, and the topic discovery of online public opinion data was studied by using LDA theme model and TF-IDF algorithm. Based on Word2Vec model, an emotional dictionary in the field of online public opinion related to police was constructed, and then the feelings of netizens were analyzed. [Limitations] The case selection and analysis are not comprehensive enough. [Results/Conclusions] The LDA theme model and TF-IDF algorithm can make the theme division more obvious, and the emotion analysis based on the domain emotion dictionary is more accurate, which better reflects the hot topics and the emotional changes of netizens in the process of public opinion dissemination.

Keywords: Police-related Internet Public Opinion; Theme Discovery; Emotional Analysis; LDA Theme Model

基金项目 国家社会科学基金项目“基于大数据的网络舆情全息建模与决策情报感知研究”(20BXW074)。

作者简介 管雨翔(1998-), 硕士研究生, 研究方向为网络舆情; 王娟(1979-), 博士, 副教授, 研究方向为数据警务技术, E-mail: wangjuan@cpcu.edu.cn; 张鹏(1981-), 博士, 副教授, 研究方向为网络危机管理。

引用格式 管雨翔, 王娟, 张鹏. 涉警网络舆情主题发现及情感分析研究[J]. 情报工程, 2023, 9(6): 105-116.

引言

随着互联网技术的快速发展,各种网络社交媒体平台不断涌现,这极大地满足了人们的社交需求。第51次中国互联网络发展状况统计报告指出,截至2022年12月,中国网民人数已经达到10.67亿,互联网普及率达到75.6%^[1]。海量的网民在各种社交媒体平台上发布各类信息、表达各种观点,这使得信息传播渠道增加、信息传播速度加快。随着社会的发展,各种社会矛盾和问题日益突出,社交平台上也充斥着大量的谣言和虚假信息,这使得网络舆情变得日益复杂。涉警网络舆情是网络舆情的重要组成部分,涉警网络舆情话题较为敏感、易引起网民群体的关注,对警察整体形象有着重要影响^[2]。

负面涉警网络舆情的传播无疑会降低政府

公信力,干扰警方的正常工作,对于和谐稳定的社会环境造成不良影响。因此,对于网民在涉警网络舆情中宣泄的情感进行监测,有利于警方及时了解舆情发展态势并对舆情进行引导。综上所述,通过收集和分析社交媒体平台上的涉警网络舆情数据,对涉警网络舆情数据进行主题发现及网民情感趋势研究,为引导和治理涉警网络舆情提供参考,对于构建良好的警民关系具有重要意义,有助于提高社会治理效能,保障社会安全稳定。

1 研究现状

通过文献计量分析软件对涉警网络舆情领域的文献进行分析,得出时间线图,如图1所示。国内对于涉警网络舆情的研究主要围绕舆情应对,舆情对公安机关、警民关系的影响等方面。

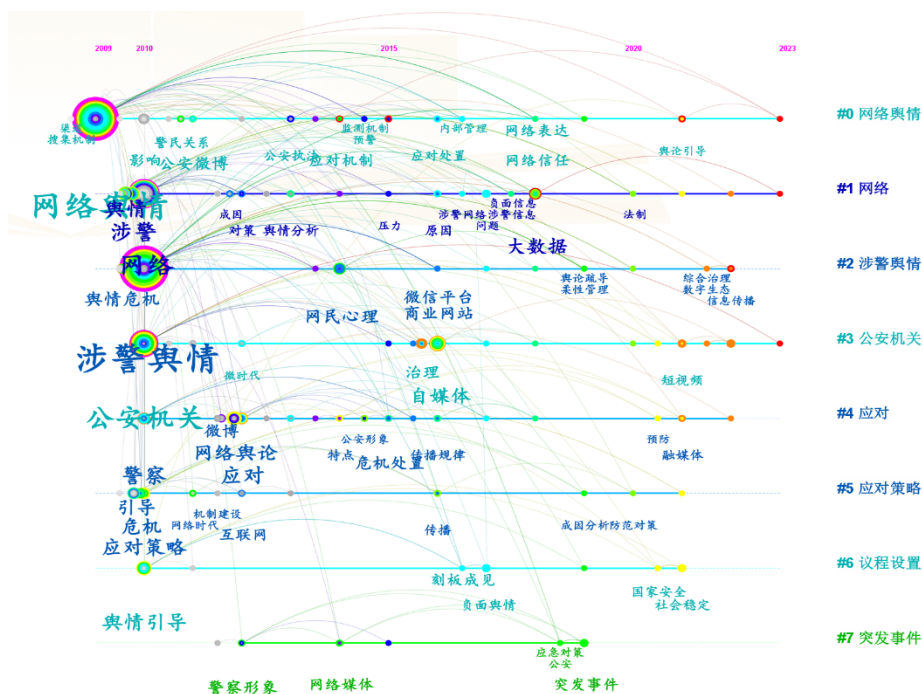


图1 2009—2023年涉警网络舆情领域时间线图

在涉警网络舆情的应对研究方面,学者们通过分析网络舆情产生的原因以及影响,提出相应的涉警网络舆情应对策略。王欢^[3]详细地研究了涉警舆情的生成背景、舆情的特点、警民和警媒关系的特点、舆情的发展趋势和治理取向。孙锦露等^[4]通过研究涉警网络舆情的相关概念和形成机制,对网络舆情的四个发展周期都提出了对应的处理措施。孙建等^[5]基于情景危机传播理论,分析了近年处理涉警网络舆情的实例,对公安机关处理舆情提出了几点建议。李若愚^[6]分析了短视频平台上的涉警网络舆情的演化周期,提出了公安机关处置舆情时的不足之处,并给出了3点对策。张鹏等^[7]分析微博平台涉警舆情的传播机制,利用系统动力学理论构建微博涉警舆情仿真模型,通过分析实例仿真结果,为警务部门应对微博涉警舆情提供了决策参考。张露露^[8]分析了涉警舆情事件的主要类型及影响,列举了中美两国应对舆情时所具有的优点,提出公安机关可以适当借鉴成熟经验,进一步提高舆情应对能力。

在研究涉警网络舆情引起的相关影响方面,学者们从不同的影响角度和对象进行分析。翟宇翔^[9]对负面涉警网络舆情进行分类,分析了网络舆情对社会稳定的影响及相应的应对措施。孙建^[10]分析了网民对涉警网络舆情的态度与安全感的关系以及网络舆情对网民安全感的影响,根据研究结果给出了相应的建议。徐梓童^[11]通过实证研究分析了公安机关处理涉警网络舆情的方式对公信力的影响,并提出了提升公安机关公信力的路径建议。

主题发现能够从大量文本数据中提取隐

含的、有价值的热点主题信息,从而帮助人们更好地理解 and 利用文本数据^[12]。目前,文本聚类算法和主题模型等相关主题发现方法和技术已经被广泛应用于主题发现研究领域中^[13],如邱均平等^[14]通过LDA主题模型提取大数据领域研究的热门主题,并进行可视化分析。徐红等^[15]使用LDA主题模型和共词分析法在学习强国平台上识别出了农村阅读推广举措的主题。

常用的情感分析方法根据所用到的技术可以分为基于情感词典的情感分析方法、基于机器学习的情感分析方法以及基于深度学习的情感分析方法^[16]。基于情感词典的情感分析方法是情感分析的基础。该方法易于理解和使用,在情感词典涵盖的情感词语较全并且情感极性、情感强度准确率较高的情况下,情感分类效果比较精确。目前,构建特定领域的情感词典研究在情感分析研究领域关注度颇高。崔彦琛等^[17]利用PMI-IR、SO-PMI等算法建立了由通用情感词典、消防领域情感词典和网络用语情感词典三种词典组合而成的消防舆情领域情感词典。周知等^[18]使用点互信息方法和规则组合构建了图书领域的情感词典。Bravo-Marquez等^[19]提出了一种方法,根据时间序列演化的分布词向量来训练增量词情感分类器,自动从Twitter的数据流中持续更新情感词汇。Aribowo等^[20]使用基于情感词典的方法分析在Covid-19大流行期间印度尼西亚语的推文,旨在研究公众情感以及七种感情下最常出现的词语。

综上所述,涉警网络舆情得到了学者的广泛关注,并取得了丰富的成果。梳理相关文献

发现, 现有研究大多关注涉警网络舆情的相关理论、产生原因和相应的治理措施。先划分阶段, 再区分主题, 在此基础上进行情感分析的相关研究并不常见, 导致对于舆情主题下的情感时序变化认识不清晰。情感时序变化研究可以帮助我们挖掘出网民对于该事件主题的情感倾向^[21], 更全面地捕捉到网民情绪的动态变化, 识别出情感发生重大变化的主题及时间段, 进一步分析舆情。因此, 本文采用 Word2Vec 模型、LDA 主题模型、TF-IDF 算法等方法, 拟对微博平台上的涉警网络舆情数据进行主题发现研究, 并通过构建领域情感词典分析网民情绪, 以期有关部门引导和治理涉警网络舆情提供决策依据。

2 舆情主题发现及情感分析

研究舆情主题及网民情感对涉警舆情研究有重要的意义, 可以帮助警方和政府更好地应对涉警舆情事件, 提高社会治理水平。分析舆情主题可以帮助警方和政府了解公众关注的重点, 从而针对性地采取措施来回应公众的需求和呼声。这有助于提高警方和政府公众心目中的形象和声誉, 提高执法效果和社会满意度。通过分析网民的情感, 可以直观地感受到公众对舆情事件的评价, 以及公众在舆情事件中对警方的看法。这有助于警方在执法过程中更好地满足公众的期待, 提高执法公信力。

本文的研究路线图如图 2 所示。

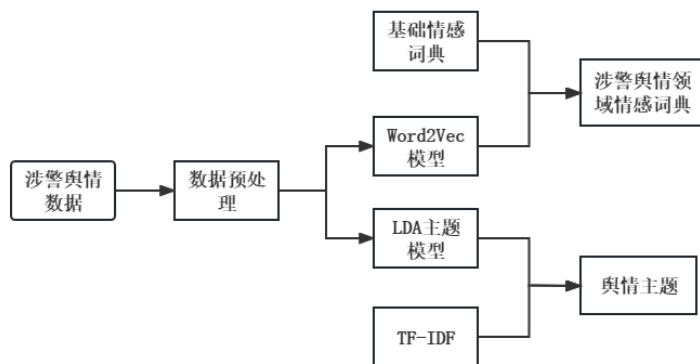


图 2 研究路线图

首先爬取涉警舆情事件数据, 进行数据预处理, 然后训练 Word2Vec 模型, 对基础情感词典进行扩充得到涉警舆情领域情感词典。通过 LDA 主题模型对主题进行划分, 再使用 TF-IDF 对特征词进行扩充, 归纳得到舆情主题。最后使用基于情感词典的情感分析方法对各个舆情主题下的网民情感进行研究。

2.1 涉警舆情数据获取

涉警舆情可以大致分为 2 类, 第一类是警

察勤务工作过程中直接引起的涉警舆情, 如警方在维护治安秩序、调查处理案件的过程中出现执法方式不当、执法力度不当等问题, 可能会引起公众的质疑和不满。第二类是由于从事警察职业的个人在日常生活中的言行而导致的涉警舆情。

微博平台用户众多、内容多样, 是最大的舆情产生及传播平台。选取“安徽一女生疑在警察注视下自杀溺亡引争议”和“安徽一妻子实名举报民警丈夫涉嫌违法违纪”这两个涉警

舆情事件为这两种舆情的典型代表。根据知微事见平台的数据统计，安徽女生自杀事件在微博平台的影响力高于75.4%的事件，总体影响力高于77%的事件，共有28家重要媒体参与，微博相关话题近40个，话题阅读量最高达2.7亿。安徽民警违法事件在微博平台的影响力高

于64.4%的事件，共有27家重要媒体参与，微博相关话题近20个，话题阅读量最高达三千多万。通过后羿采集器爬取微博平台上的相关网络舆情数据。时间范围分别为2020年12月5日—12日和2022年1月5日—12日，共计6080条文本数据，部分数据如表1所示。

表1 部分舆情数据

发布人	发布内容	发布时间
维达拉托里	#键盘侠# 很多键盘侠之所以会在网络上有恃无恐地进行批判，是因为他们躲在一块屏幕后面 - 他们知道他们的言语并不会给自己带来惩罚。如果是在现实生活中，他们绝不敢轻易批判别人，因为那可能招来别人的拳头。但在网络上就不一样了 - 没有人知道键盘侠们长什么样子、姓甚名谁，所以他们才敢肆无忌惮地发表自己的歪理邪说。就最近的#安徽17岁女孩在警察目视下溺水身亡##成都确诊病例孙女#在事实真相还没有完全浮出水面之前，很多人轻信了网上的舆论，在不了解事实的情况，迫不及待地对着警察，成都女孩一顿狂喷。盲目跟风，向来都是大众的心理特征之一。但凡这些键盘侠有一点独立思考的能力，也不至于妄下定论，伤害到别人。	2020-12-11 00:20
Y聪明小脑瓜33	绝了，乌烟瘴气，警察，消防员，医生都是特殊职业，有自己特殊的职责，评论里一堆抨击警察不下去救人的，那换个新闻，警察勇救自杀女孩，不幸双双殒命。你们又会说自杀的人为什么要去救，吧啦吧啦，又开始称赞警察。是非全是你们一张嘴，有想过警察不会游泳去救了，没救回来都死了，那是几个家庭的悲剧	2020-12-5 09:56
这糖保甜吗	这个是少数个例，警察很多都很辛苦的，节假日，巡逻的不都是警察吗。以前朋友宿舍进了贼，2天不到就抓到了，各行各业都有一两颗老鼠屎	2022-1-8 01:42
执伞画晴天	把警察群体的脸被这个人丢光了	2022-1-8 20:57

2.2 涉警舆情数据处理

在处理微博平台的数据时，需要去除文本中的话题标签等无意义信息，使文本内容更具有实际意义。使用正则表达式进行匹配去除微博话题及@用户名等内容，正则表达式匹配结果如表2所示。

2.3 涉警情感词典构建

基于情感词典的情感分析方法基本原理是使用情感词典提供的代表情绪类别的词语，来划分文本数据的情感极性。该方法首先输入文本数据，进行数据预处理操作（包括数据清洗、数据降噪、去除无效评论等），然后对文本数

据进行分词、去停用词操作，最后根据判断情绪类别的规则将文本数据的情感极性输出。

已有的情感词典大多为人工构建，人工构建情感词典需要花费大量的人力和时间总结已有的研究成果对情感词语的情感极性以及情感强度进行人工标注。现在主流的中文情感词典包括BosonNLP情感词典、知网HowNet情感词典、清华大学李军中文褒贬义词典和大连理工大学情感词汇本体库等。这些情感词典分别包含不同数量、不同情感强度的积极词语和消极词语。本文去重融合知网情感词典、大连理工情感词典和NTUSD情感词典构建基础情感词典，共有41738个情感词语。

表 2 正则表达式匹配结果

正则表达式	文本内容
无	【#安徽警方通报女生跳河自杀事件#】@安庆望江公安在线 消息，安徽省安庆市望江县公安局发布警情通报称，2020年12月4日13时许，望江县公安局110报警服务台接报警：望江县城吉水桥边有一女子欲投河自尽。接警后，处警民警立即赶到现场处置。在民警安抚劝导过程中，女子突然扑入深水区。处警民警展开施救，后将该女子打捞上岸，经抢救无效死亡。目前，死者自杀原因正在调查，相关善后工作正在进行。针对网友反映的民警、辅警在施救过程中存在的问题，我局第一时间成立调查组，对涉事民警、辅警作出停职接受调查决定。
#(.*?)#	【】@安庆望江公安在线 消息，安徽省安庆市望江县公安局发布警情通报称，2020年12月4日13时许，望江县公安局110报警服务台接报警：望江县城吉水桥边有一女子欲投河自尽。接警后，处警民警立即赶到现场处置。在民警安抚劝导过程中，女子突然扑入深水区。处警民警展开施救，后将该女子打捞上岸，经抢救无效死亡。目前，死者自杀原因正在调查，相关善后工作正在进行。针对网友反映的民警、辅警在施救过程中存在的问题，我局第一时间成立调查组，对涉事民警、辅警作出停职接受调查决定。
@(.*?)[:]	【】消息，安徽省安庆市望江县公安局发布警情通报称，2020年12月4日13时许，望江县公安局110报警服务台接报警：望江县城吉水桥边有一女子欲投河自尽。接警后，处警民警立即赶到现场处置。在民警安抚劝导过程中，女子突然扑入深水区。处警民警展开施救，后将该女子打捞上岸，经抢救无效死亡。目前，死者自杀原因正在调查，相关善后工作正在进行。针对网友反映的民警、辅警在施救过程中存在的问题，我局第一时间成立调查组，对涉事民警、辅警作出停职接受调查决定。

Word2Vec 是用来生成词向量的神经网络模型，Word2vec 依赖 Skip-grams 或 CBOW，Skip-grams 根据当前词预测上下文来进行训练，而 CBOW 根据上下文预测当前值^[22]。

本文将得到的语料与基础情感词典匹配，将重复的词语作为基准词语，共有 2246 个基准词语，其余词语作为新增词语，使用 Skip-grams 训练模型，计算得到与基准词语最相似的 10 个词语，并把其中的新增词语作为扩充词语。最后，扩充了 351 个词语，部分扩充词语如表 3 所示。

表 3 部分扩充词语

词语	极性	词语	极性
不怕牺牲	积极	徇私枉法	消极
高素质	积极	脱离群众	消极

2.4 文本主题发现

LDA(Latent Dirichlet Allocation) 是一种文

档主题生成模型，它基于贝叶斯理论，有着高效的计算方法和相对简单的模型结构^[23]。它将文本数据中的每一篇文档视为一个由若干主题组成的混合，并假设每一个主题都是由一组词构成的。LDA 可以自动地探测出文本数据集中的主题结构，并为每一篇文档分配主题。LDA 模型非常适用于文本数据分析，它能够帮助我们提取出文本数据中的主题信息，从而实现文本的自动分类和摘要。

LDA 主题模型最常用的两个评价方法是困惑度和主题一致性，困惑度 (Perplexity)^[24] 可以理解为对于一篇文章 a ，所训练出来的 LDA 主题模型对 a 属于哪个主题的不确定程度就是困惑度。困惑度越低，说明聚类的效果越好。困惑度的公式如下所示：

$$\text{perplexity}(D) = \exp\left[-\frac{\sum_{d=1}^M \log_2 p(W_d)}{\sum_{d=1}^M N_d}\right] \quad (1)$$

其中参数 D 表示文档集合，集合 D 中含有 M 篇文档， N_d 为文档 d 中的单词个数； $p(W_d)$

为文档 d 中单词 W_d 产生的概率。

主题一致性 (Topic Coherence)^[25] 是对于 LDA 生成的主题下词汇的语义一致性的定量评价指标, 较大的一致性指标代表主题下的词汇联系更紧密。因此, 一致性指标最高点即为最优主题数。主题一致性的公式如下所示:

$$Coherence(T) = \sum_{(i,j)} Score(i, j) \quad (2)$$

$$Score(i, j) = \log \frac{D(i, j)}{D(j)} \quad (3)$$

其中, T 代表某一主题, i, j 为该主题内的词汇, $D(i, j)$ 代表这两个词汇同时出现的文档数量, $D(j)$ 表示包含该词汇的文档数量。

TF-IDF 是一种基于词频和逆文档频率的统计方法, 可以用于挖掘文章中的关键词^[26]。计算公式如下:

$$TF = Count / Num \quad (4)$$

Count 是某个词语在特定文本中的出现次

数, Num 是该特定文本的总词数;

$$IDF = \log \frac{TextNum}{1 + texts} \quad (5)$$

TextNum 是语料库的文本总数, texts 是包含该词的文本数。

$$TF-IDF = TF * IDF \quad (6)$$

本文使用 TF-IDF 对 LDA 主题模型的特征词进一步丰富扩充, 能更好地归纳出主题。

3 实证研究

本文对“安徽一女生疑在警察注视下自杀溺亡引争议”这一事件进行实证分析, 根据如图 3 所示的事件累计评论数将事件分为两个阶段进行研究。2020 年 12 月 5 日—7 日为事件热度上涨期, 12 月 7 日后为事件热度消退期。

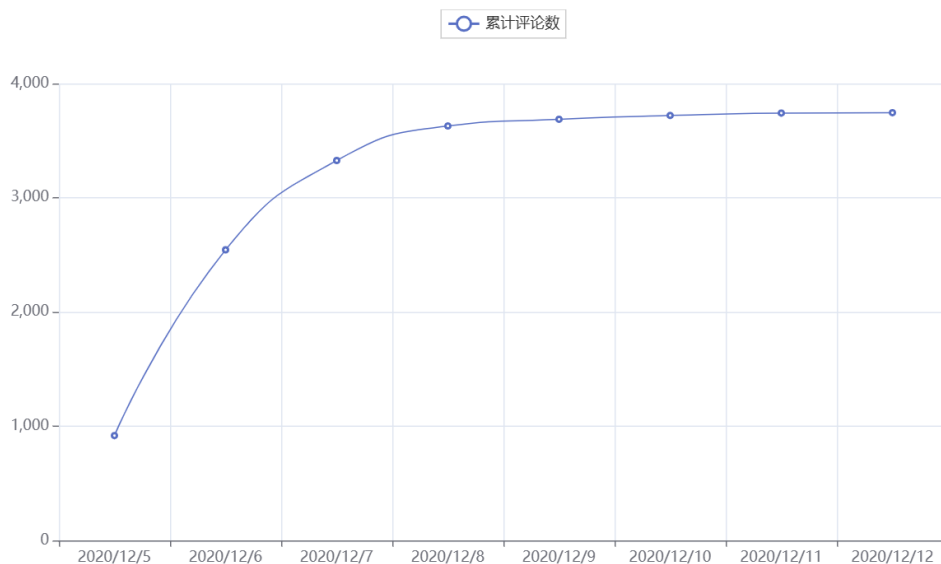


图 3 累计评论数

3.1 网民情感趋势

对舆情数据按天为单位进行情感分析, 图 4 为网民情感极性变化图。由图可知,

在网络舆情发生时, 网民的讨论往往带有强烈的情绪色彩, 并且很容易受到偏见性文章的影响, 造成负面舆情。在该舆情事

件中负面情绪占比较大，舆情事件经过一两天的发酵就到达了高峰，随着事件的进一步发展，关于事件的更多细节被披露，

引起了网民情感的波动。到了事件的热度下降时期，网民评论数也随之减少，网民情感极性数量趋于0。

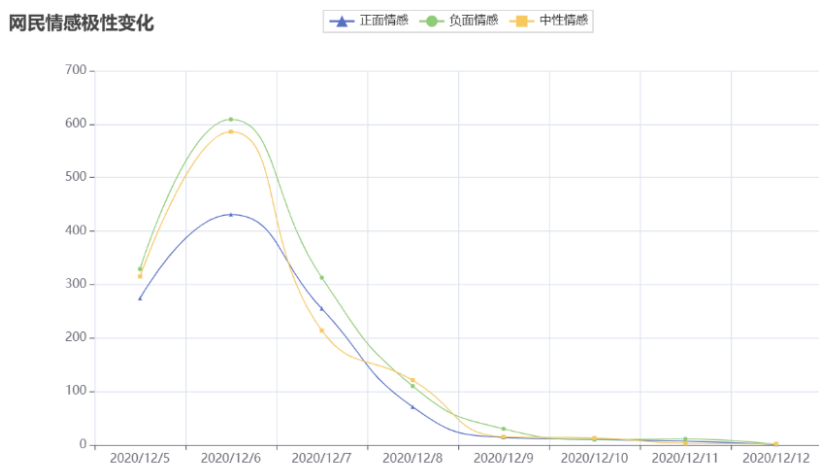


图4 网民情感极性

3.2 舆情主题发现

本文通过困惑度和主题一致性综合考虑选取最佳主题数。第一阶段的困惑度及一致性随

主题数目变化情况如图5、图6所示。困惑度呈现递增状态，而主题数目为3时一致性最大，且此时的困惑度比大部分主题数的困惑度小，所以，该阶段的主题数目取3。

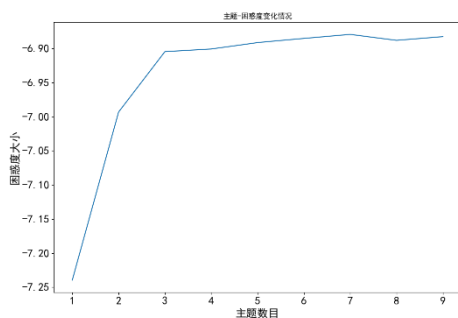


图5 第一阶段困惑度变化图

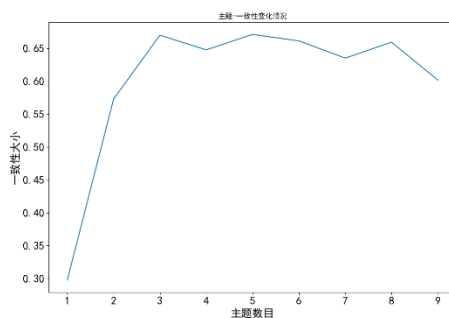


图6 第一阶段一致性变化图

使用 TF-IDF 对各个主题的特征词进行扩充，如表4所示。

在第一阶段，主题A主要话题是救援溺水女孩的警察不会游泳；主题B是警察注视安徽女孩跳河轻生；主题C主要讨论的是望江溺亡

女孩事件中的警察被停职。

表4 主题特征词

主题	特征词
主题A	救援, 女孩, 溺水, 警察, 游泳, 不会
主题B	警察, 女孩, 安徽, 注视, 轻生, 跳河
主题C	警察, 溺亡, 女孩, 停职, 望江县

第二阶段的困惑度及一致性随主题数目变化情况如图 7、图 8 所示。困惑度呈现递增状态，而主题数目为 3 和 4 时一致性最大，

主题数为 3 时的困惑度比主题数为 4 时的困惑度小得多，所以，该阶段的主题数目取 3。

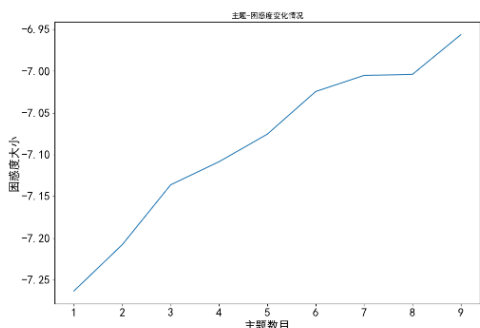


图 7 第二阶段困惑度变化图

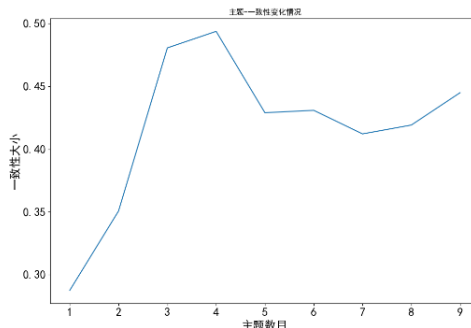


图 8 第二阶段一致性变化图

使用 TF-IDF 对各个主题的特征词进行扩充，如表 5 所示。

第一阶段的 13%。网民的关注点从警方的救援行动及责任问题，聚焦到了基层治理及应急救援上。

表 5 主题特征词

主题	特征词
主题 D	警察，望江，女孩，视频，视角
主题 E	警察，女孩，安徽，注视，身亡，跳河
主题 F	警察，女孩，溺水，救人，救生衣

3.3 各主题下情感演化

在第二阶段，主题 D 主要话题是关于望江女孩的其他视角的视频曝光；主题 E 主要讨论的是警察注视安徽女孩跳河身亡；主题 F 讨论了女孩溺水有警察送来救生衣。

两个阶段的舆情主题归纳完毕后，对各主题下的评论数据进行情感分析，并绘制折线图，如图 9 所示。舆情处于热度增长阶段时，在“救援溺水女孩的警察不会游泳”这一主题下，网民的情感以中性为主，在“警察注视安徽女孩跳河身亡”主题下，网民情感则以负面为主，而“涉事警察被停职”主题下的负面情感占比逐渐增加，是因为有网民认为民警的救援存在问题造成女孩溺亡的主要原因之一，应当追责。在这一阶段，网民对于该事件中警察的救援行为感到不满而发布了大量负面评论，警方也及时进行了回应并对涉事警察作出了处罚。

综上所述，当涉警舆情事件发生时，舆情信息传播速度极快。在不同的发展阶段，网民对于热点事件会密切关注，并且关注的热点话题会受到事件最新信息的影响而产生变化。每个阶段的热点话题之间阅读量和讨论量都相差不大，随着时间的推移和事件全貌的还原，网民的关注度慢慢下降，负面情绪占比及讨论量也随之下降，第二阶段的总体讨论量只占到了

舆情处于热度消退阶段时，在“关于望江女孩的其他视角的视频曝光”这一主题下，网民的情感以中性为主。在“警察注视安徽女孩跳河轻生”主题下，网民的负面情感逐渐减少。

“女孩溺水有警察送救生衣”主题下网民的情感以负面为主。随着更多事件细节披露出来,

网民逐渐了解事件真相,部分网民呼吁不应过分苛责警方,负面情感也随之下降。

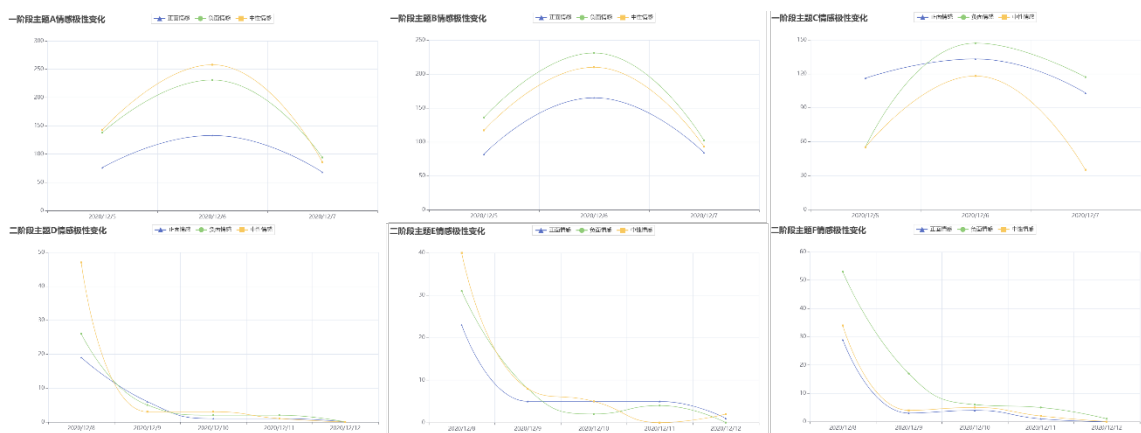


图9 各主题下网民情感变化

4 意见建议

警察在执法执勤过程中,任何一个过程都有可能引起群众的关注,一旦有人将事件发布到网络上,就会迅速形成舆情,舆情有正面也有负面。现如今,随着网民群体的扩大,部分网民很容易受到不良言论的影响,在未了解事件全貌的情况下就发表负面评论,抹黑工作人员甚至是政府的形象,造成了不良的社会影响。通过 3.1 至 3.3 的实证研究,本文向公众、社交平台、媒体以及警方提出如下意见建议。

4.1 公众应理性看待舆情

公众应该理性看待涉警网络舆情,不盲目相信不实信息,遵守法律法规,不发表和散布谣言。在平台上合法合理地表达自己的意见和看法。

4.2 社交平台应加强内容审核

社交平台作为网络舆情传播的重要媒介,有责任对网络舆情进行监控和管理,防止网络

舆情造成恶劣影响。加强内容审核,及时删除虚假、不良信息。对发布不良和过激言论的相关人员进行禁言等处理。平台及时推荐官方声明及新闻报道,向用户提供更全面和正确的信息,进行舆情引导。

4.3 媒体应遵循客观公正原则

媒体应该保持客观公正的态度,以确保新闻报道的公正性。确保所报道的内容是真实、准确的,避免发布错误虚假的信息,全面客观地报道事实,避免误导公众。媒体在报道时,应该遵守法律法规,不发表违法言论或者侵犯他人的合法权益,应该尊重警察形象,不进行无端攻击或者诋毁,引导舆论正能量,促进社会和谐稳定。

4.4 警方应加强舆情相关建设

警方应该加强网络舆情的监测,通过实时监测网络舆情,及时发现负面涉警网络舆情。当负面舆情发生时,警方应及时介入,对网络舆情进行监测和评估,了解舆情动态,研究应

对措施,合理配置资源进行应对。与公众保持积极的交流,及时回应群众关心的问题,提供真实、准确、权威的信息,保证舆情的公开透明,让群众了解真相,并以积极主动的态度引导舆情,防止负面信息的扩散。对发布不当言论的相关人员依法进行处罚,重视意见领袖的影响力,发挥网络大V的作用,与其他相关组织和部门进行协调合作。通过科学有效的网络舆情调控,避免网络舆情恶化,促进舆论氛围的稳定。公安机关应当提升各个部门的政治素养和业务素质,提升公安机关的公信力。公安机关应重视涉警网络舆情,在各大社交平台上树立良好的形象,构建良好、坚固的警民关系。

5 结论

本文使用“安徽一女生疑在警察注视下自杀溺亡引争议”这一涉警舆情事件进行实证研究,通过LDA主题模型及TF-IDF在事件不同时期归纳得到不同的主题,并分析了该话题下网民的情感极性以及讨论的热门话题。实验结果表明,本文提出的涉警网络舆情主题发现和情感分析研究方法,可以有效地展示舆情的发展情况,可帮助有关部门挖掘研究涉警网络舆情的相关规律,同时,为相关部门进行涉警网络舆情风险研判提供参考。

本文还存在局限和不足之处,一是实验数据只涉及文本类型,部分社交平台未能包括在内,未来将对图片、视频等数据进行研究,平台研究数据也会进一步增加;二是构建领域情感词典所添加的新词语只包含了两个涉警网络舆情的数据,后续仍需进行扩充;三是情感词

典不能及时更新且无法识别反讽特殊语句,因此,如何更为精确地判断网民情感极性还有待进一步深入。

参考文献

- [1] 中国互联网络信息中心. 第51次中国互联网络发展状况统计报告 [EB/OL]. [2023-03-02]. <https://cnmic.cn/n4/2023/0302/c199-10755.html>.
- [2] 宋全喜,寻其锋,宁祎娜. 网络舆情的研究与发展——以涉警网络舆情为例 [J]. 运筹与管理, 2018, 27(4): 93-104.
- [3] 王欢. 涉警舆情态势分析:根源、现状与趋势 [J]. 山东警察学院学报, 2017, 29(3): 142-154.
- [4] 孙锦露,李玫瑾. 涉警事件的网络舆情形成机制与应对策略 [J]. 中国人民公安大学学报(社会科学版), 2018, 34(1): 112-121.
- [5] 孙建,向学亮. 基于情景危机传播理论的涉警舆情应对策略研究 [J]. 信息安全, 2021(S1): 74-77.
- [6] 李若愚. 短视频网络涉警舆情的演化与对策研究 [J]. 网络安全技术与应用, 2022(11): 159-162.
- [7] 张鹏,刘静,兰月新,等. 新媒体涉警舆情的动力学仿真与治理对策研究 [J/OL]. 系统科学学报: 1-7[2023-07-30]. <http://kns.cnki.net/kcms/detail/14.1333.N.20221214.1350.001.html>.
- [8] 张露露. 从美国警察枪杀黑人案看中国涉警负面舆情应对 [J]. 中国人民公安大学学报(社会科学版), 2016, 32(3): 133-137.
- [9] 翟宇翔. 负面涉警舆情对社会稳定的影响及应对 [J]. 领导科学论坛, 2021(6): 79-83.
- [10] 孙建. 涉警网络舆情对网民安全感、满意度的影响 [J]. 河北公安警察职业学院学报, 2019, 19(4): 65-68.
- [11] 徐梓童. 涉警网络舆情处置与公安机关公信力提升的实证研究 [J]. 网络安全技术与应用, 2022(7): 152-154.
- [12] 邵琦,牟冬梅,王萍,等. 基于语义的突发公共卫生事件网络舆情主题发现研究 [J]. 数据分析与知识发现, 2020, 4(9): 68-80.

- [13] 王曰芬, 王一山, 杨洁. 基于社区发现和关键节点识别的网络舆情主题发现与实证分析 [J]. 图书与情报, 2020(5): 48-58.
- [14] 邱均平, 沈超. 基于 LDA 模型的国内大数据研究热点主题分析 [J]. 现代情报, 2021, 41(9): 22-31.
- [15] 徐红, 张斯婷, 李凌方. 基于 LDA 模型与共词分析法的农村阅读推广主题发现与热点分析 [J]. 情报科学, 2022, 40(10): 67-73.
- [16] 王春东, 张卉, 莫秀良, 等. 微博情感分析综述 [J]. 计算机工程与科学, 2022, 44(1): 165-175.
- [17] 崔彦琛, 张鹏, 兰月新, 等. 消防突发事件网络舆情情感词典构建研究 [J]. 情报杂志, 2018, 37(10): 154-160.
- [18] 周知, 王春迎, 朱佳丽. 基于超短评论的图书领域情感词典构建研究 [J]. 情报理论与实践, 2021, 44(9): 183-189.
- [19] BRAVO-Marquez F, KHANCHANDANI A, PFAHRINGER B. Incremental word vectors for time-evolving sentiment lexicon induction[J]. Cognitive Computation, 2021(5): 1-17.
- [20] ARIBOWO A S, KHOMSAH S. Implementation of text mining for emotion detection using the lexicon method (case study: tweets about Covid-19)[J]. Telematika, 2021(1): 49-60.
- [21] 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.
- [22] 李秀茹, 王晓, 李朋朋, 等. Word2vec 和支持向量机的 POI 自动分类方法 [J]. 测绘科学, 2022, 47(6): 195-203.
- [23] BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation[J]. Journal of Machine Learning Research, 2003, 3(1): 993-1022.
- [24] 韩亚楠, 刘建伟, 罗雄麟. 概率主题模型综述 [J]. 计算机学报, 2021, 44(6): 1095-1139.
- [25] 陈果, 吴微. 细分领域 LDA 主题分析中选词方案的效果对比研究 [J]. 情报理论与实践, 2019, 42(6): 138-143.
- [26] 石进, 韩进, 赵小柯, 等. 基于语境概念核心词提取算法研究 [J]. 情报学报, 2019, 38(11): 1177-1186.