

面向突发事件应急管理的社交媒体 多模态信息分析研究

徐元¹, 毛进^{1,2}, 李纲^{1,2}

(1. 武汉大学信息管理学院, 武汉 430072; 2. 武汉大学信息资源研究中心, 武汉 430072)

摘要 当前, 社交媒体已经成为突发事件态势感知的一种重要信息来源, 其中包含文本、图像、视频等多模态信息, 可广泛应用于突发事件的应急管理。近年来, 作为一项具有挑战性的任务, 面向突发事件的社交媒体多模态信息分析已受到学术界和工业界的广泛关注。本文围绕突发事件中的社交媒体多模态信息分析研究展开综述, 从内容、时空和传播三个维度解析社交媒体多模态信息的多维特征, 并从信息获取、信息整合和信息挖掘三个方面归纳多模态信息分析的关键方法和技术。最后, 本文在综述的基础上, 结合“情境-应对”模式下的突发事件应急管理目标, 构建社交媒体多模态信息分析框架, 并从信息获取、描述、分析和可视化等方面展望未来的研究方向, 为面向突发事件应急管理的社交媒体多模态信息分析研究和实践提供指引, 以期通过方法创新提升突发事件的应急管理能力。

关键词 突发事件; 应急信息管理; 社交媒体; 多模态信息分析; 语义分析

Social Media Multimodal Analysis for Emergency Management

Xu Yuan¹, Mao Jin^{1,2} and Li Gang^{1,2}

(1. School of Information Management, Wuhan University, Wuhan 430072;
2. Center for Studies of Information Resources, Wuhan University, Wuhan 430072)

Abstract: Social media has become an important information source for situation awareness during emergencies. Multimodal information such as text, images, and videos viewed through social media can be widely used for emergency information management. Multimodal information analysis during emergencies is a challenging topic and has received significant attention from both academia and industry. We review studies on social media analysis in emergencies. The multi-dimensional characteristics of multimodal information in social media are analyzed using three dimensions: content, spatial-temporal, and information dissemination. The key methods and technologies of multimodal information analysis are summarized from three aspects: information acquisition, information integration, and information mining. Finally, we construct a multimodal information analysis framework for emergency management based on this review and propose future research directions on information acquisition, description, analysis, and visualization. We expect this survey to provide guidance for research on and practice of social media multimodal information analysis and to improve the ability to manage emergencies.

Key words: emergency; emergency information management; social media; multimodal information analysis; semantic analysis

收稿日期: 2021-02-05; 修回日期: 2021-06-06

基金项目: 国家自然科学基金重大项目课题“国家安全大数据综合信息集成与分析方法”(71790612); 国家自然科学基金创新群体项目“信息资源管理”(71921002)。

作者简介: 徐元, 男, 1994年生, 博士研究生, 主要研究方向为信息组织、多模态信息分析; 毛进, 男, 1988年生, 副教授, 主要研究方向为信息组织、学术大数据分析; 李纲, 男, 1966年生, 教授, 博士生导师, 主要研究方向为网络信息资源管理、竞争情报, E-mail: imiswhu@aliyun.com。

1 引言

随着移动互联网的迅速发展,微博、微信、Twitter等社交媒体平台作为桥接物理和虚拟世界的信息通信工具之一,逐渐成为突发事件应急管理中重要的实时信息获取渠道^[1]。在突发事件期间,大量与人员伤亡、设施损毁、紧急求助等内容相关的文本、图像、视频被用户实时分享到社交媒体上,这些多模态信息对于突发事件的应急管理具有重要作用^[2-3]。在大数据环境下,突发事件应急管理的思维模式也发生了转变,传统基于专家经验的管理决策模式转变为“情景-应对”型应急管理的新模式,即从大数据中挖掘有价值的知识来深入认识突发事件的时空格局、活动模式、演化规律与内在运行机制,以实现对突发事件态势及演化趋势的全面理解^[4]。社交媒体为“情景-应对”模式中的突发事件应急管理提供了新的数据来源,不同模态的信息在内容上相互关联和呼应,所包含的内容也存在差异,如文本中包含事件的细节内容,而图像和视频则更为直观地展示事件的场景,文本与图像、视频之间能够相互补充和印证,共同反映物理世界中事件的真实发展状态。从不同层次、不同侧面对复杂的社交媒体多模态信息进行抽象、概括和融合分析,总结并提炼出比单一模态更为准确和全面的综合情报,能够帮助用户全方位、深层次地理解突发事件的实时态势,从而减少突发事件应急管理中的不确定性。

社交媒体多模态数据呈现规模大、增长速度快以及来源、种类、模态多样等特征,如何将海量、分散、无序、动态变化的多模态信息转化为能够有效支撑应急管理的情报是突发事件下社交媒体多模态信息分析的核心问题,其关键在于正确认知数据背后的特征、逻辑和科学内涵,从而快速准确认

识、判别和分析突发事件在不同阶段的态势变化,并以此为依据辅助应急管理人员制定科学有效的应对决策^[4]。近年来,相关研究和应用实践已较为丰富,且仍在逐步增多。因此,本文采用文献综述方法,通过梳理国内外有关突发事件下社交媒体多模态信息的研究成果,从信息资源特征和信息分析方法两个方面解析社交媒体多模态信息的多维特征,归纳特定事件场景下的多模态信息分析方法体系,并以此为基础构建面向突发事件应急管理的社交媒体多模态信息分析框架,指出现有研究中存在的不足和挑战,为深入研究面向突发事件应急管理的社交媒体多模态信息分析方法提供新的线索与方向。

2 突发事件中的社交媒体研究概况

本文选择以 Web of Science 和中国知网中的期刊论文和会议论文为文献来源,进行基于标题的检索来获取国内外有关突发事件中的社交媒体研究的文献,时间范围是2020年及其之前的年份。在关键词的选择上,主要从两个方面考虑:一是与突发事件相关的词,包括“突发事件”“灾害”“洪水”“台风”“地震”“火灾”等;二是与社交媒体相关的词,包括“社交媒体”“微博”“Twitter”“Facebook”等。通过审查标题和摘要去除不相关的文献,筛选后得到801篇文献,包括625篇英文文献和176篇中文文献。文献的年份分布如图1所示,可以发现突发事件中的社交媒体研究在2011年后开始逐渐增加,这与社交媒体平台和智能手机的发展过程相吻合。随着社交媒体用户数量的增加,突发事件中可用于分析的社交媒体数据也在逐渐增加,使得社交媒体成为突发事件应急管理的重要数据来源成为可能,也促进了突发事件中社交媒体的研究。

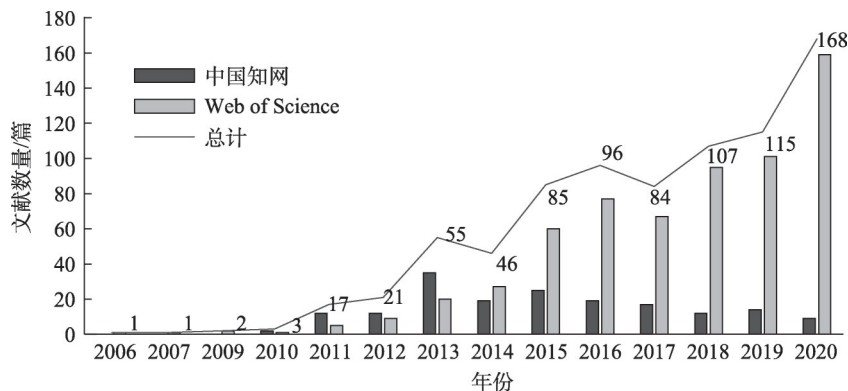


图1 国内外有关突发事件中社交媒体研究文献的年份分布图

通过分析文献所属学科领域可发现,当前有关突发事件中社交媒体信息分析的研究主要集中在计算机科学、传播学、管理学和信息科学等研究领域,存在大量的跨学科研究。不同学科的研究侧重点也存在差异,例如,从计算机科学的角度,主要研究突发事件下社交媒体数据智能化处理技术;从传播学的角度,主要关注突发事件下社交媒体信息的传播与扩散规律;从管理学的角度,主要关注社交媒体信息在突发事件应急管理中的应用。在突发事件的类型上,主要关注地震、飓风、洪水、袭击、传染病等危害人类生命和财产损失的重大自然灾害或公共安全事件。本文研究数据的来源主要包括 Twitter、Facebook 和微博等主流社交媒体平台, Twitter 由于其用户量大、数据获取便捷等特点,成为最主要的国外社交媒体数据来源,国内主要利用微博数据进行研究。

在检索到的文献中,标题或摘要中直接出现“多模态”表述的文献只占总文献数的2%,且均为近5年发表,但是将多种模态的社交媒体信息作为分析对象或能够体现多模态融合思想的文献有216篇,占总文献数的27%,对于文本和图像两种模态的信息分析最为常见。随着人工智能和大数据等技术取得突破性进展,基于机器学习和深度学习的自然语言处理和计算机视觉技术在多模态数据处理任务中的性能和效果得到了明显提升,并已经成为社交媒体多模态信息分析的主要手段。这表明突发事件中社交媒体多模态信息分析研究是一个较新的研究领域,正逐渐被更多研究者所关注,并且形成了一定的理论和方法技术体系,所以有必要从整体上了解突发事件中社交媒体多模态信息的研究现状,

为面向突发事件应急管理的社交媒体多模态信息分析的研究提供理论和方法支撑。

本文将包含多种模态的社交媒体信息分析或能够体现多模态融合思想的文献作为主要的综述文献,重点关注如何利用多模态信息分析方法将原始的社交媒体信息转化为能够有效支撑应急管理决策的情报。社交媒体中包含多种模态的内容,用户以文本、符号等形式发布内容的同时,上传相关的照片、音频、视频等附件,并标记当前所处地理位置和时间戳,此外,用户之间通过浏览、点赞、评论和转发等操作进行交流和互动,这些用户之间的互动行为可以抽象为复杂的网络结构,上述内容共同形成了社交媒体多模态信息在内容、时空和网络三个层面的特征。而有关突发事件中社交媒体多模态信息分析可以概括为信息获取、信息整合和信息挖掘三个层次。从社交媒体中的实时数据中可发现和获取与突发事件相关的多模态信息;以主题或事件为基本单元对多模态信息进行语义描述和关联;最后面向特定的应用场景,从这些多模态信息中挖掘出对应决策有价值的隐性知识并应用到应急管理中。

综上所述,本文将突发事件下社交媒体的研究归纳为资源和方法两个方面:一是社交媒体信息本身在突发事件中所具有的多维特征;二是突发事件中面向特定应用场景的社交媒体信息分析方法,如图2所示。充分掌握社交媒体多模态信息特征为信息分析方法的研究提供了指引,而借助高效的信息方法能够深层次挖掘出社交媒体多模态信息中的隐含内容和规律,二者联系紧密,共同为突发事件的应急管理提供支撑,也能较好地涵盖不同学科视角下的研究成果。

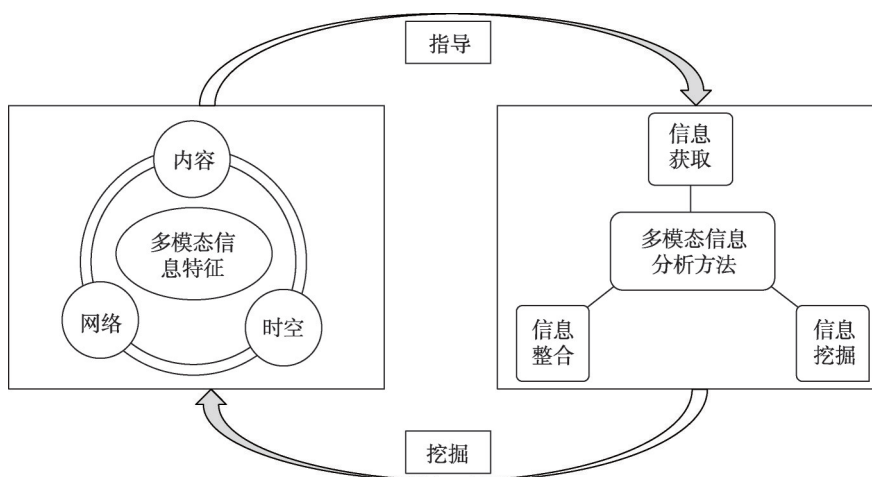


图2 突发事件社交媒体多模态信息的主要研究内容

3 突发事件中社交媒体多模态信息的多维特征

3.1 内容特征

社交媒体本身的信息内容是突发事件实时态势最直观的反映,蕴含了人物、地点、状态等各类突发事件的构成要素,从总体上感知和把握社交媒体信息的内容特征能够发现用户普遍关注的问题和应急管理中有价值的信息。不同类型的突发事件中,研究人员所关注的信息内容也存在一定的差异,形成了不同的分类标准和体系。例如,Alam等^[5]将哈维等飓风中的社交媒体信息分为受影响的人群、损坏的基础设施、警告和建议、捐赠和志愿服务等信息;Cho等^[6]将2011年日本地震中的社交媒体信息分为灾害态势、个人经历、建议、技术、情感和行动;Vieweg^[7]根据事件中的受影响对象,将社交媒体信息分为社会环境信息(建议、警告、疏散、死亡、伤害、医疗、失踪和帮助等与人相关的内容)和物理环境信息(建筑物、道路等与物相关的内容);Rudra等^[8]根据信息是否反应事件实时状态,将社交媒体信息分为情境信息(受影响的人群、设施以及对应措施等)和非情境信息(事件原因、捐赠和事后分析等)。这种粗粒度的分类能够帮助用户尽早地获取其所需要的信息,而细粒度的社交媒体信息分类体系能够帮助用户更加深入地掌握事件的细节,如Huang等^[9]根据突发事件的预警、准备、响应和恢复四个阶段,定义不同的阶段下社交媒体信息的子类别,该细粒度分类体系及其关键词能够反映整个事件期间的社交媒体内容。

不同类型的社交媒体用户在发布的内容和语言风格上也存在差异。个体用户发布的推文通常是与事件相关的第一手信息,表达个人的主观态度、情感、需求和经历^[10],语言较为随意,信息的真实性有待进一步考证,在分析过程中需要结合其他用户推文和实际情况来合理使用;政府部门、主流新闻媒体等官方用户发布的内容包括官方预警、事件态势更新、建议对策、救援行动等^[11],在语言上更为严谨,具有更高的可信度和权威性,能够被直接转发和使用。这种差异体现了在突发事件期间,不同社会角色关注的问题也不同,可能导致不同用户发布内容在传播效率和利用效率上存在差异。

除了文本信息,社交媒体中的图像、视频等多媒体信息也提供了与突发事件相关的内容,能够更加直观和全面地展现事件的态势详情,对于应急管

理同样具有重要作用^[12]。一般情况下,在社交媒体内容中,文字、图像等不同模态的信息同时存在,单一模态的信息通常缺乏部分事实性内容,例如,图像能够直观展示受到损坏的设施状态,而文本中则包含更精确的受影响人员信息^[5]。研究结果表明,大部分社交媒体的内容中图像和文本是对象相关和场景相关,不同模态和载体的信息能够相互补充、相互印证,对其进行融合分析能够为态势感知提供比单一模态更为准确和全面的内容,在事件的识别以及事件带来影响的评估上具有更大价值^[13]。

现有的研究主要是针对特定的突发事件、数据集和研究任务来构建相应的分类和描述体系,而较少有研究考虑图像、视频等其他模态的信息。不同研究中,对于社交媒体多模态信息的分类没有形成一个被普遍接受的统一标准,不同信息分类体系也缺乏语义上的关联,导致不同研究的成果之间不能进行交叉比较和验证,这将会导致对于突发事件中社交媒体内容的一般性模式和规律揭示存在阻碍。

3.2 时空特征

社交媒体提供的时间戳和地理位置数据^[14],将时间和空间维度嵌入到语义内容中,通过联合分析社交媒体信息的时空分布和语义内容,有助于进一步了解突发事件期间不同位置的用户在不同阶段的关注热点和看法。

在时间维度上,不同类型突发事件中的社交媒体信息数量和内容随时间的演变模式存在差异。地震经常导致相关推文的数量激增(推文数量在一小时内达到当地峰值)^[15],与情感有关的推文比例随着时间推移逐渐下降^[6];而在洪水和飓风等事件中的推文数量增加较为缓和,与情感相关的推文比例随着时间的推移而持续增加^[16]。这可能是由于该类型的事件能够被有效预测和预警,并且发生较为频繁,而地震事发突然且造成的影响巨大,具有极高的用户关注度。

在空间维度上,与未受影响的用户相比,受灾地区的用户倾向于在社交媒体上发布更多与灾情相关的信息^[17],如在火灾事件中,发布位置与火灾位置之间的距离与推文发布数量之间存在显著的负相关^[18]。此外,事件受影响地区的用户和未受影响地区用户发布推文的内容也存在差异。例如,在“海燕”台风中,居住在受灾最严重的菲律宾的用户发布了更多有关协调救济的推文,而菲律宾以外的用

户则发布了更多有关悼念活动的信息^[19]。在“哈维”飓风中,灾害爆发时,受灾用户主要发布物理环境类话题,而非受灾用户倾向于情感的表达;灾害过后,受灾用户通常发布情感类话题的推文,而非受灾用户则更倾向于发布建设环境和物理环境类话题^[20]。

突发事件中社交媒体信息的时空特征会受到一些外部因素的干扰,如地区人口密度和信息基础设施,可能导致社交媒体信息的空间数量分布不平衡,当突发事件发生在人口密度较大且网络基础设施较好的地区时,区域内信息的总数也会较多,而在人口较少或信息基础设施较为薄弱的地区,区域内信息的数量较少。如何通过获取其他社会感知数据源来补全缺失的社交媒体信息,以及纠正这些地区社交媒体数据不足所导致的社交媒体时空分析偏差,也是在数据获取中需要考虑的问题。

3.3 网络特征

在社交媒体中,每个用户是信息发布和转发的节点,形成了一种可参与并具有自组织特性的网络结构来交流和传播信息,信息传播的速度和广度远远超过了传统新闻媒介。突发事件下的社交媒体网络分析能够发现用户如何通过社交媒体来传递相关信息,从而深刻地理解突发事件在事前、事中和事后的信息在社交媒体中的传播规律。

社交媒体中的信息转发行为是信息传播的主要渠道,通过社交媒体中的用户网络结构和连接交互模式能够了解不同信息流以及不同用户对于有效信息传播的贡献^[21],用户的关注数和推文的转发数在社交网络中大致符合幂律分布的特点^[22],即中心用户(拥有大量粉丝的用户)和关键推文(大量被转发的推文)主导着在线社交网络中的信息传播过程,突发事件的信息传播主要由关键节点推动^[23]。因此,拥有更多关注者的个体账号所发布的推文通常具有更高的曝光度,在传播紧急信息中往往起着更加关键的作用,而官方账号发布的权威信息也更容易被转发^[22],在连接不同个人和网络之间尤为重要^[24]。

从传播扩散的角度来说,通过社交媒体网络进行内部传播的速度要快于这些网络中来自外部来源的信息,在灾难期间较早发布的信息比在灾难更重要的阶段较晚发布的信息具有更高的传播速度^[25]。此外,推文的内容、形式和用户特征也会影响信息的转发率,如包含更长文本、主题标签和链接^[26],

并与预警建议、损失情况和危险位置等事件关键信息的原创推文更容易被转发^[27],包含请求和情感的推文也更容易被传播和扩散^[28]。在综合考虑用户行为特征、网络全局信息和影响力衰退机制的基础上,动态识别不同阶段的关键节点动态^[29],并制定相应的内容编辑策略,有助于把相关信息快速、准确地传播给有需要的人群。

社交媒体的用户网络、转发网络、链接网络,在一定程度上反映了与突发事件相关的信息在社交媒体中的传播模式,然而,当前只是停留在单一网络特征的揭示,对于网络动态演变过程及其因果关系的探索不足,对于加入不同类型实体和时空关联的异质网络特征以及突发事件信息在社交媒体中的传播机理仍有待进一步探索。

4 突发事件中社交媒体多模态信息分析关键方法

4.1 信息获取方法

社交媒体中包含大量与突发事件无关、冗余或虚假的噪声数据,有用信息分散在不同的社交媒体信息中,存在严重的信息过载问题。需要通过分类的方式从大规模的社交媒体数据中自动识别与突发事件相关的信息,并进一步抽取出其中的细粒度信息。

4.1.1 多模态信息分类

在突发事件中,社交媒体多模态信息分类任务主要有两种:一是从大规模社交媒体信息中识别与突发事件相关并包含丰富信息的内容,这是一个二分类问题;二是根据不同场景将信息分配到更为细粒度的类目中,以满足不同的信息需求,这是一个多分类问题。社交媒体多模态信息分类主要有基于关键词、基于特征工程和基于深度学习这三种方法。

基于关键词的方法通过目标关键词检索和精确匹配以获取相关社交媒体信息,为了适应事件的动态演变,关键词需要动态生成和实时更新^[30],但是在短时间内选取合适的关键词进行精确检索和过滤较为困难。基于特征工程的方式,通过提取词汇、句法、视觉等外部特征和内容特征,将社交媒体信息转化为可用于分类的指标,再利用逻辑回归、决策树、支持向量机等数学模型构建有监督的机器学习分类器,实现自动化的信息分类^[31]。基于深度学

习的方法是利用预训练模型来整合和表示社交媒体内容,并结合灾害事件语料库进行微调,最后通过迁移学习来完成下游的社交媒体内容分类任务,在一定程度上考虑了信息上下文之间的语义关联,通常能够展现出比传统机器学习更好的性能^[32]。除了单一模态的信息分类,通过多模态融合的方式构建多模态信息的联合表示空间,能够有效提高信息分类的性能^[33]。例如,根据预定义的层次结构为每个模态特征分配不同的权重,再构建深度学习模型,将推文分配到不同的灾害类别及其子类别中^[34]。

此外,社交媒体中存在大量被重复转发的内容,从信息内容的角度,这些社交媒体信息中并没有提供除原文之外的其他有用信息,需要在数据处理的过程中删除这些冗余信息,以保证在有效利用信息的情况下减少信息分析的工作量^[35],可以采用余弦相似度、欧氏距离、杰卡德相似、感知哈希等算法计算文本相似度和图像相似度,找出数据集中相同或相似的信息,并制定一定的规则来删除这些冗余信息。

4.1.2 多模态信息抽取

与正式的官方消息和新闻相比,突发事件情境下的社交媒体中存在非正式语言、缩写、拼写错误、用词模糊、语法随意等问题,原始信息的有效性和真实性有待进一步验证。这就需要利用命名实体识别等方法抽取受影响的人员、时间、地点及其属性和关系,进行指代消解和实体消歧,并将实体链接到现实世界中,从原始信息中获得更为丰富和准确的语义内容^[36]。Rexiline Ragini 等^[37]通过分析用户在紧急情况下的用词风格,利用机器学习和预定义规则相结合的方法识别事件中处于危险状态的人群。此外,也可以引入外部语料库和规范化词表,利用预定义规则来扩展词间语义关系和词法搭配关系,构建损失知识库,以此来提高信息抽取的精度^[38]。社交媒体信息中的图像和视频也提供了有关设施损坏和人员伤亡等的重要信息,文本上下文可能包含损坏对象、位置、原因等详细描述信息,通过构建多模态融合的深度学习模型,分别从图像和文本中提取语义特征,从而更加全面地识别细粒度信息。

精确的地理位置也是突发事件应急管理中的重要要素,将社交媒体内容映射到事件空间地图中,能够增强突发事件的地理态势感知能力。由于互联网中的隐私保护政策,部分用户在发帖时选择隐藏地理位置,大量与事件相关的内容中缺少位置信

息,需要通过图像、视频和用户等特征来预测和解析社交媒体信息中描述的事件所在位置。社交媒体中地理位置解析方法主要有三种:一是构建外部知识库或地理词典,通过间接相关的实体与外部知识库或现实世界空间对象进行匹配^[39-40],以此来预测用户发布内容时所在的位置;二是根据用户的社交个人资料来获取社交媒体用户或内容的位置^[41];三是通过地理空间区域的视觉元素和属性来识别地理位置,通过领域自适应技术和图像转换方法可以将损坏的场景与未损坏的场景进行匹配^[42],实现图像的视觉内容向地理坐标的映射^[43]。

基于有监督机器学习是多模态信息分类和信息抽取的主要方式,为机器学习提供丰富的训练数据来源,是有效提高和评价多模态信息分类和抽取模型性能的重要方式之一。目前这方面的可用数据集较少,其规模和领域覆盖范围也难以满足模型训练的需求,需要构建面向多种突发事件应用场景,涵盖多种类型、多种模态信息的标准数据集。

4.2 信息整合方法

社交媒体中与突发事件相关的信息以独立条目的形式存在,存在实际内容冗余、语义不一致等问题,其中的信息量超过人脑的信息处理能力,并且还要进一步将不同模态和类型的信息纳入统一的高层语义表示空间中;事件作为一种信息的统一表示和组织方法,能够准确清晰地表示发生的各种事实型信息,反映特定实体在特定时间和地点相互作用的客观事实,是整合社交媒体信息的主要方式,对于突发事件的深度理解具有重要的作用。

4.2.1 事件检测

在突发事件发生后,社交媒体在一段时间和一定地区内的与某一主题相关的信息数量会突然增加,在一定程度上反映了现实世界中事件的真实状态。通过汇集大规模实时社交媒体流数据,监测一定时空范围内的主题内容,能够在第一时间检测到正在发生的异常事件或突发事件。事件检测分为重大突发事件检测和子事件检测^[44]。前者是从大规模的社交媒体数据流中第一时间检测到实时发生的重大事件,以支持突发事件的快速响应;后者是检测伴随着突发事件发生而产生的一些次生或衍生事件,以追踪突发事件的演化和发展过程。

1) 重大突发事件检测

对于特定灾害事件的检测,通过统计和比较当前时刻与前一时刻或历史时刻信息中的主题标签频

率来发现周期性或异常事件。例如,在地震检测中,随着信息发布位置与震中之间距离的增加,与地震相关的信息数量会明显减少,单位时间内的信息发布频率也会明显下降。Poblete等^[45]和Sakaki等^[46]基于此规律构建了社交媒体信息中与“地震”相关的关键词频率的时序分布模型,用于判断地震事件的发生,并进一步通过地理标签过滤的方式确定地震发生的位置。对于未知事件的检测,在考虑时间和地理特征的基础上,Zhou等^[47]利用聚类、主题模型等无监督的语义分析技术,构建基于概率的主题时空模型,挖掘原始推文集合的潜在主题或事件,并计算事件发生概率。社交媒体中的图像和视频等多模态数据能够更加直观地反映事件状态,对于突发事件的检测更具价值^[48]。Daly等^[49]利用图像识别的机器学习模型识别社交媒体中实时上传的火灾信息,再结合时间戳和地理标签确定发生火灾的时间和具体位置,实现区域内火灾事件的实时监测。

2) 子事件检测

重大突发事件的发生,通常伴随着衍生事件或子事件的发生,如“洪水导致人员被困”“地震导致道路损坏”等。所有子事件共同构成了突发事件多侧面的热点问题,反映整个事件的全貌和细节。子事件或衍生事件的检测方法与突发事件的检测方法相类似,通过主事件关键词来跟踪相关信息,但更注重信息之间的时序特征。其核心步骤是对相似的推文进行二次聚类或主题建模,得到突发事件中的子主题,再对子主题进行加权、排序和合并,并结合时间和地理特征,对子事件进行增量更新^[50-51]。此外,还可以借助图模型来检测子事件,Meladianos等^[52]以关键词作为节点,将较短时间间隔内连续的推文表示为一个加权的词共现图,再使用凸优化算法来检测图边缘权重的变化,找出与事件最相关的推文或关键词来表示不同子事件。

4.2.2 事件摘要

事件摘要是指将检测到的突发事件及其子事件片段进行语义编码和整合,形成全面而简短的事件描述的过程,分为抽取式摘要和生成式摘要。抽取式摘要是从原始社交媒体信息中识别出可能表示事件的关键词语和句子,通过聚类、图模型或规则将这些词和句子进行重新排列和组合,得到有关事件的描述^[53];生成式摘要是在语义理解的基础上,凝练核心内容,提取社交媒体信息中的主要实体、关系和动作短语,实现序列化的语言重构。事件摘要

的表示通常遵循“5W”模式(what, where, when, who, why),即包含事件的五要素:内容、地点、时间、涉及人物和发生原因^[54]。

在社交媒体中的图像对于突发事件的展现更为直观,Kuang等^[55]和Xu等^[56]将图像特征融入事件摘要中,利用关键字、地理位置、代表性推文和代表性图像来表示每个子事件,生成关于突发事件的多模态连贯性摘要报告。为了获取每个子事件下最相关的图像,李显鑫利用种子查询词不断扩展子事件下的查询文本,构建融合图文多模态信息的图片事件相关性判定模型,同时考虑对各模态内特征和跨模态间关联特征,以实现事件相关图像的采集^[57]。在基于主题模型的多模态事件摘要中,Bian等^[58]和Schinas等^[59]通过基于图模型的排序算法,从文本、图像、时间和社交关系等方面计算图像与主题的相关性,构建跨模态的概率模型,通过不同模态信息与事件之间的相关性来发现事件中的子主题,并进一步汇总形成可视化的事件摘要。除了主题建模的方式,Qian等^[60]通过文本-图像-用户聚类的方式寻找事件的子节点,从视觉相似性、重要性和多样性等维度为每个子节点分配最具相关性的文本和图像,形成融合文本和图像的事件摘要。

4.2.3 事件演化分析

事件摘要把突发事件当作一个整体来看待,忽略了事件的动态演化过程以及子事件之间的关联关系。从突发事件的生命周期出发,建立事件从起始期、爆发期、蔓延期和衰退期的全生命周期事件主题演化模型和事件脉络模型,构建融合多源、多模态数据的时空语义模型,分析突发事件中更为复杂的时间-空间-内容的协同演化过程,可实现更细粒度、更深层次、更全角度以及更多侧面的突发事件描述。

1) 事件主题演化建模

在基于时序关系的事件演化分析中,利用社交媒体捕获事件演化各个阶段的讨论数量、热度和内容等特征,建立基于时间序列的主题动态演化模型^[61-62],能够有效呈现不同时间段内事件主题的变化情况。此外,等结合维基百科、遥感和地理信息等外部数据源,利用时空分析和文本分析相结合方法,构建突发事件的时空演化语义模型,分析突发事件中更为复杂的时间-空间-内容的协同演化过程,实时跟踪事件的时空演化态势和特征^[63-64],以更简洁的方式理解整个事件的发展趋势。事件画像的概念和方法也被应用于事件演化分析中,构建由

基于事件演化的一系列话题主题构成的“主线信息链”以及基于各个话题下用户评论信息和事件传播特征的“副线信息链”，来表征事件多个维度的发展状态^[65]。

2) 事件脉络生成

为了更加清晰地理解和掌握社交媒体中有关突发事件的构成情况，需要从更细的粒度揭示事件之间、事件要素之间的潜在逻辑和时空关联，引入事件链、事件图谱等细粒度的信息组织方法，进行突发事件的数据建模。从叙事的角度，进一步识别事件中更细粒度的实体、属性和关系，将无序、杂乱的事件组合成有序、有规律的故事情节，构建事件图谱，并以可视化的形式展现事件发展脉络^[66]，形成一个连贯且自成体系的叙事结构。其核心任务是事件之间的关系刻画，通过图模型或机器学习的方法识别事件之间在时间和逻辑上的关联，以及事件中地理位置、参与者、核心词之间的关联，形成事件下各个主题的分支结构，将这些分支利用生成树生成事件的脉络结构^[67]。图像也被应用到事件演化的建模中，理解并抽取图像的语义特征，将其分配到事件的文本时间线中，生成融合图像-文本摘要的全局事件时间线^[68]。

当前，对于社交媒体多模态信息整合的研究仍主要是以单一模态的信息或简单的图像和文本融合作为整合的对象，尚未较好地建立不同模态信息之间的关联，对于多模态融合的研究有待进一步深入探索。来自不同模式的信息通常带来关于概念、对象、事件等的互补和印证，如何将这些不同模态信息进行深度融合得到更为稳健的事件态势演化推理结果是多模态信息融合中最主要的研究问题。此外，越来越多的用户开始倾向于在社交媒体中发布视频，其中也包含丰富的图像、文本、声音等多模态信息；但是目前对于视频分析的手段有限，只是简单地将视频中的关键帧转换为图片进行分析，对于视频内容的解析和融合也将是有待解决的问题之一。

4.3 面向应用场景的信息挖掘方法

社交媒体数据中除了与突发事件有关的显性信息，还包含大量有待进一步挖掘的隐性知识。根据特定的应用场景，利用特征关联和语义挖掘等手段发现社交媒体中的有关用户情感和事件态势等内容，对于事件的态势评估和走势预测均具有重要作用。

4.3.1 观点挖掘

在突发事件期间，每个用户都可以在社交媒体上自由地表达自己的观点、反应、情感、态度和看法，有助于深入了解公众面对突发事件和应急响应的真实反应^[69]，对于用户情感识别、舆情演化分析、需求挖掘以及敏感人群监测都具有重要作用。

1) 用户情感分析

情感分析的目的是识别社交媒体信息中隐藏的用户对于突发事件的主观情感倾向，采用机器学习或深度学习等方式，建立语法、语义和表情符号等特征与情感之间的潜在关系，将用户发布内容分为正向、负向或中立情感，或者通过开发情感词典，使用更为准确和细粒度的情感词来标记信息。社交媒体信息中的图像同样是通过视觉、对象或场景传递情感的，可以是图像中人物本身所表现出来的情感，也可以是用户看到图像所产生的情感，但是图像中的视觉特征和主观情感之间存在语义鸿沟，这就导致了每个人对于相同图像所产生的情感可能会产生差异。图像与文本的融合方法也受到了研究者更多的关注，构建基于注意力机制的网络，利用视觉和文本信息之间的相关性进行情感分析，能够有效提升情感识别的性能^[70-71]。

2) 舆情演化分析

分析社交媒体中不同阶段突发事件的主题与情感演化特征和规律，帮助应急管理部门了解突发事件中公众对事件的看法和情感倾向，预测未来的舆论趋势走向，并妥善应对民众关注热点^[72]。从用户的角度，通过分析突发事件中不同利益相关者的情感状态和分布，探寻利益相关者之间的情感传播路径，结合舆情话题综合分析利益相关者的情感演化态势^[73]。一般情况下，意见领袖对网络舆情的传播具有决定性作用，利用主题一致性和情感支持识别网络评论意见领袖，从而引导正面评论意见领袖积极发声，减少消极情绪，合理回应存在的问题，避免情绪的极端化和舆情态势的恶化^[74]。

3) 需求挖掘

社交媒体用户也会发布内容来描述自身或其他人的困难并寻求帮助，需求挖掘是指从社交媒体信息中抽取受影响人员、资源需求、损失情况等更细粒度的事件要素信息，从而指导应急管理部门以此为依据，优化物资和人员的调度策略，提供精准的救助服务。利用分类模型识别出包含用户需求的社交媒体内容和需求类型（如食物、住所、医疗、救援等），进一步抽取需求主体、需求资源、需求

时间和地点等结构化信息,并计算紧急状况下受影响人群的各种需求级别^[75]。除了单一的需求挖掘,Purohit等^[76]从推文中识别突发事件中与捐助、救援等相关的资源供应信息,利用需求和供应之间的相似度计算或属性配对进行自动供需匹配,辅助资源的调度,从而提高突发事件应急响应的速度。

4) 敏感人群监测

突发事件的破坏性和不确定性加剧了公众的焦虑和恐慌等负面情绪,事件中受影响的人群更容易出现心理方面的问题,如创伤后应激障碍。通过社交媒体信息的情感挖掘,发现不同用户在事件发生前后一系列(包括恐慌、焦虑或悲伤等)负面情绪^[77],再结合需求动机模型编制用户的需求-动机映射表,最终得到不同类型用户的深层次情感需求特征^[78]。而结合时空数据,将公众情感映射到时空地图中,追踪公众情感变化情况,可以及时发现心理状况存在异常的人群,并及时提供一定的心理援助和情绪引导,从而减少与突发事件相关的心理问题发生^[79]。

4.3.2 态势分析

在紧急情况下,传统的依靠专家来评估事件发展态势是一种耗时费力的方式,而随着与突发事件相关的社交媒体多模态数据的不断汇集,利用社交媒体中多模态信息的深层次关联挖掘,能够有效估计突发事件的严重程度、宏观态势等事件的隐藏信息,为应急管理提供更加快速可靠的决策支持情报。

1) 事件严重性评估

社交媒体信息数量与灾害破坏和财产损失之间存在一定的相关性,可以利用社交媒体信息的数量分布和内容快速推测和评估突发事件的影响和潜在风险^[80]。因此,通过计算突发事件后的社交媒体推文密度、转发量、图文数量等指标,可以估计和预测事件的影响和破坏程度^[81],如Kropivnitskaya等^[82]利用地震发生后短时间内的社交媒体中与地震相关的推文数量来估计地震烈度,而Nguyen等和Li等借助深度学习方法将事件根据不同推文特征分为不同的严重等级^[83]或量化为连续型的损害指数值^[84]。

2) 态势估计

从社交媒体中挖掘得到的信息本身并不能直接揭示当前的事件态势,还需要揭示社交媒体信息与事件态势的关联关系及映射机制,通过融合社交媒体和其他来源数据(如传感数据、水文数据和数学模型)来实现事件总体态势的估计。de Albuquerque

等^[85]和Arthur等^[86]将与洪水相关的推文通过地理编码映射到GIS系统中,通过推文中图像和文本描述来估计特定区域中洪水的淹没和破坏情况,并形成洪水地图。为了进一步精确估计和预测洪水深度和淹没区域,将推文信息转换为降雨量的值,Brouwer等^[87]和Li等^[88]构建地理高程模型来计算排水渠上方的水位高度,再从已知淹没位置与目标位置之间的距离和高程差得到水位的概率指数,能够有效降低水位估计和预测的不确定性。

当前针对特定的场景和信息需求的社交媒体多模态信息挖掘研究中,分析数据和目标都具有较强的领域性,使得其信息挖掘的方法和结果缺乏多领域的适用性,对于挖掘方法的评价上也存在一定的困难。如何针对突发事件中典型场景和任务,探索具有一定普适性的信息挖掘方法,是值得进一步研究的问题。此外,目前只有少量研究涉及分析结果可视化的问题,简单地将事件要素以统计图表或时空地图的形式呈现出来,缺乏有针对性的对突发事件态势可视化理论和方法的研究。

5 总结与展望

尽管社交媒体信息已经被广泛用于突发事件的应急管理中,但是其研究成果和应用效果具有一定的局限性:在当前的研究中,通常是利用单一的信息分析方法,对单一模态的社交媒体信息进行分析,来满足应急管理中某一环节或阶段中的信息需求,对于突发事件中社交媒体多模态信息分析的一般流程、方法技术和应用场景缺乏系统全面的探讨,不能满足“情景-应对”模式中的突发事件应急管理模式对于社交媒体多模态信息处理的要求。

面向突发事件应急管理的社交媒体多模态信息分析目标,贯穿于突发事件应急管理中的事前预警、事中响应和事后评估三个阶段。在事前,通过对社交媒体数据的动态监测,预测将会发生的事件和正在发生的事件,以实现危机的预警;在事中,通过对与突发事件相关社交媒体多模态信息的整合和挖掘,进行事件的全景式建模和事件推演,以实现事件的态势感知;在事后,通过实际情况与分析结果的对比,对整个信息处理过程进行总体评估,优化信息处理的流程和分析方法。

为了实现上述目标,本文构建了面向突发事件应急管理的社交媒体多模态信息分析框架,如图3所示,从信息处理流程出发,以大规模社交媒体多

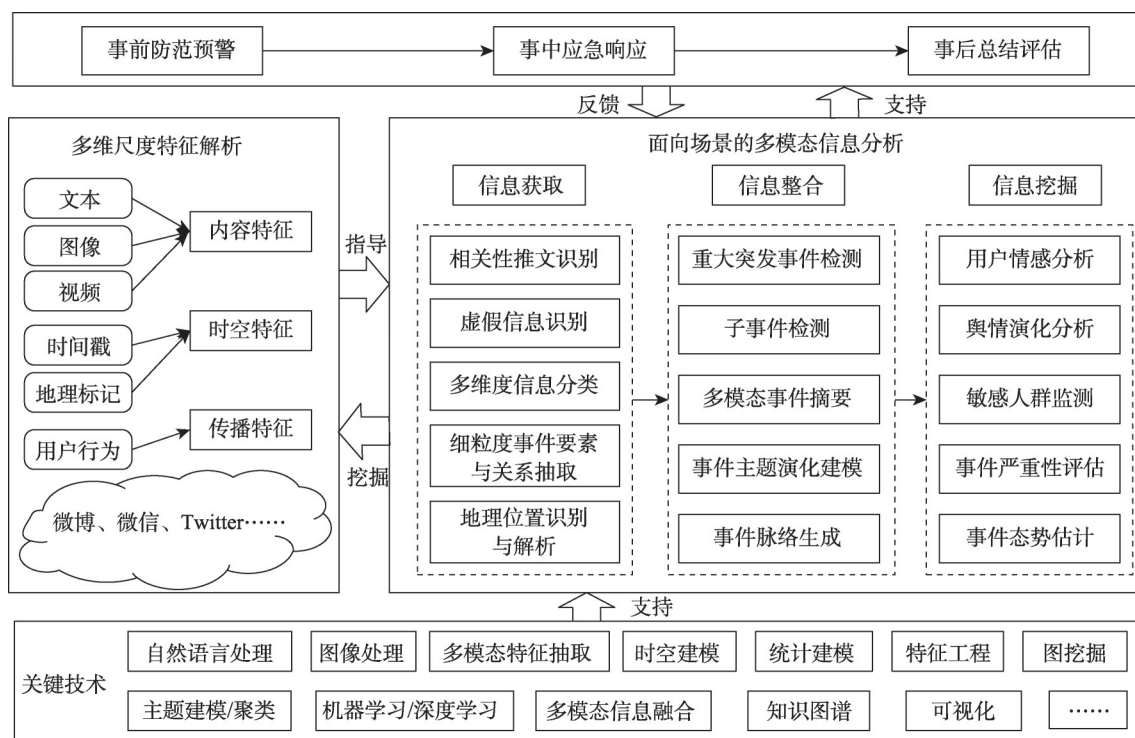


图3 面向突发事件应急管理的社交媒体多模态信息分析框架

模态信息获取为前提，以信息整合和挖掘为手段，为不同场景和需求下的应急管理提供实时可靠的信息支持，最终实现虚拟社交网络中的事件与物理世界中的实际事件态势的同步呈现。

然而，通过梳理可发现，现有的研究存在信息获取的边界不清晰、信息整合的标准多样、信息挖掘方法的多场景适用性不强、多模态信息融合的实际应用较少、信息可视化手段较为单一等问题，并且在数据来源的多样性、分析方法的适用性、分析结果的可信度以及应用场景的多样性等方面存在一定的不足，未能较好地支撑本文提出的面向突发事件应急管理的社交媒体多模态信息分析框架。在未来的研究中，可从以下几个方面做进一步探索。

(1) 确立实时高效的信息获取机制。社交媒体信息通常具有传播速度快、扩散广等特点，数据的使用价值随着时间流逝而急剧衰减，突发事件的应急决策对于数据的时效性和全面性有较高要求。如果只关注事件发生前后某一时间段或者地域范围内的截面数据，那么数据获取的边界不明确，实时的全量数据获取困难，导致分析结果容易出现偏差。需要在信息过滤和相关性识别的基础上确定合理的数据获取边界和实时信息获取策略，提高突发事件相关数据感知的效率、准确性和时效性，既要获取全面的信息来满足用户的信息需求，又要尽量减少

无关信息，降低信息处理的难度。

(2) 构建统一的多模态信息描述和组织体系。由于突发事件、信息需求和信息模态本身具有多样性，建立统一的突发事件信息分类体系和描述标准存在巨大的困难。随着语义网和知识图谱技术的发展，针对代表性的事件类型建立分类框架和本体模型更具有可能性和可操作性，利用语义映射方法将公共安全大数据资源映射到高层语义空间中，形成统一的内容理解和语义关联。这种方式既能提供不同模态信息的整合框架来集成突发事件的相关信息，也能支持跨事件、跨模态的分析以及不同机构和研究中的信息交换。

(3) 探索可复用、可解释的多模态信息分析方法。大规模机器学习和深度学习方法的被应用，在提升了分析效率的同时，其分析结果在解释性、可靠性和普适性方面有待进一步验证。需要研究可泛化、可扩展、可移植的信息分析方法，对于一般性的场景，研究共性化的分析方法和工具，再通过迁移学习和自适应的方式满足个性化的应用场景需求。同时，研究数据驱动与知识驱动相结合的协同分析方法，利用先验知识和相关理论来指导知识发现的过程，将分析问题简化为逻辑计算和算法规则的过程，通过数据驱动的方式来揭示隐藏的和以前未知的模式、趋势和变化，从而提升深度学习的可

解释性和可靠性。

(4) 优化多维度、多粒度的信息可视化方法。可视化界面能够向决策者更为直观地展示事件的主要结构以及发展变化情况, 如何将突发事件复杂变化的态势以简洁的方式展现对于应急管理具有重要的意义。需要将复杂、模糊、多样的一系列突发事件纳入到时空参考框架中进行可视化研究, 建立符合用户常识性认知的可视化表达方式。将事件要素间的动态交互关系客观地投影到时间和空间轴上, 在可自定义的图表中显示相关指标, 在地图上显示地理位置以及在特定视图中显示重要的数据、图像或视频, 并实时更新突发事件的动态演化过程。

6 结 语

突发事件的发生和演化是一个动态的过程, 所面临的情境复杂多变, 具有高度的不确定性, 又涉及安全、社会稳定等敏感问题, 需要以及时可靠的情报作为依据来减少管理决策的不确定性。快速高效地获取、整合、分析和挖掘社交媒体中不同模态的碎片化信息, 提取出与突发事件预警警报、实时态势和发展趋势等相关的有用情报, 有助于提升应急管理部门对突发事件的实时预警、监测和态势感知能力。本文通过回顾突发事件中社交媒体多模态信息分析的相关文献, 梳理了突发事件中社交媒体信息多维特征和多模态分析方法, 构建了面向突发事件应急管理的社交媒体多模态信息分析框架, 为突发事件中社交媒体多模态信息的研究和实践提供指引, 以期通过创新社交媒体多模态信息分析方法促进突发事件应急管理模式的创新和突破。

参 考 文 献

- [1] Reuter C, Ludwig T, Kaufhold M A, et al. Emergency services' attitudes towards social media: a quantitative and qualitative survey across Europe[J]. *International Journal of Human-Computer Studies*, 2016, 95: 96-111.
- [2] Zhang C, Fan C, Yao W L, et al. Social media for intelligent public information and warning in disasters: an interdisciplinary review[J]. *International Journal of Information Management*, 2019, 49: 190-207.
- [3] Steiger E, de Albuquerque J P, Zipf A. An advanced systematic literature review on spatiotemporal analyses of Twitter data[J]. *Transactions in GIS*, 2015, 19(6): 809-834.
- [4] 曾大军, 曹志冬. 突发事件态势感知与决策支持的大数据解决方案[J]. *中国应急管理*, 2013(11): 15-23.
- [5] Alam F, Ofli F, Imran M. Descriptive and visual summaries of disaster events using artificial intelligence techniques: case studies of Hurricanes Harvey, Irma, and Maria[J]. *Behaviour & Information Technology*, 2020, 39(3): 288-318.
- [6] Cho S E, Jung K, Park H W. Social media use during Japan's 2011 earthquake: how Twitter transforms the locus of crisis communication[J]. *Media International Australia*, 2013, 149(1): 28-40.
- [7] Vieweg S E. Situational awareness in mass emergency: a behavioral and linguistic analysis of Microblogged communications [D]. Ann Arbor: University of Colorado at Boulder, 2012.
- [8] Rudra K, Ghosh S, Ganguly N, et al. Extracting situational information from Microblogs during disaster events: a classification-summarization approach[C]// *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. New York: ACM Press, 2015: 583-592.
- [9] Huang Q Y, Xiao Y. Geographic situational awareness: mining tweets for disaster preparedness, emergency response, impact, and recovery[J]. *ISPRS International Journal of Geo-Information*, 2015, 4(3): 1549-1568.
- [10] David C C, Ong J C, Legara E F T. Tweeting supertyphoon Haiyan: evolving functions of Twitter during and after a disaster event[J]. *PLoS one*, 2016, 11(3): e0150190.
- [11] Kim J, Bae J, Hastak M. Emergency information diffusion on online social media during storm Cindy in US[J]. *International Journal of Information Management*, 2018, 40: 153-165.
- [12] Peters R, Albuquerque J P D. Investigating images as indicators for relevant social media messages in disaster management[C]// *Proceedings of the ISCRAM 2015 Conference*, 2015: 24-27.
- [13] 李纲, 张霁, 毛进, 等. 灾害事件下社交媒体图文相关性研究[J]. *情报学报*, 2020, 39(11): 1223-1231.
- [14] Chae J, Thom D, Jang Y, et al. Public behavior response analysis in disaster events utilizing visual analytics of microblog data[J]. *Computers & Graphics*, 2014, 38: 51-60.
- [15] Avvenuti M, Cresci S, La Polla M N, et al. Nowcasting of earthquake consequences using big social data[J]. *IEEE Internet Computing*, 2017, 21(6): 37-45.
- [16] Spence P R, Lachlan K A, Lin X L, et al. Variability in Twitter content across the stages of a natural disaster: implications for crisis communication[J]. *Communication Quarterly*, 2015, 63(2): 171-186.
- [17] Crooks A, Croitoru A, Stefanidis A, et al. Earthquake: Twitter as a distributed sensor system[J]. *Transactions in GIS*, 2013, 17(1): 124-147.
- [18] Kent J D, Capello H T. Spatial patterns and demographic indicators of effective social media content during the Horseshoe Canyon fire of 2012[J]. *Cartography and Geographic Information Science*, 2013, 40(2): 78-89.
- [19] Takahashi B, Tandoc E C, Carmichael C. Communicating on

- Twitter during a disaster: an analysis of tweets during Typhoon Haiyan in the Philippines[J]. *Computers in Human Behavior*, 2015, 50: 392-398.
- [20] 李纲, 陈思菁, 毛进, 等. 自然灾害事件微博热点话题的时空对比分析[J]. *数据分析与知识发现*, 2019, 3(11): 1-15.
- [21] Hagen L, Keller T, Neely S, et al. Crisis communications in the age of social media: a network analysis of Zika-related tweets[J]. *Social Science Computer Review*, 2018, 36(5): 523-541.
- [22] Dong R S, Li L B, Zhang Q P, et al. Information diffusion on social media during natural disasters[J]. *IEEE Transactions on Computational Social Systems*, 2018, 5(1): 265-276.
- [23] 曹学艳, 段飞飞, 方宽, 等. 网络论坛视角下突发事件舆情的关键节点识别及分类研究[J]. *图书情报工作*, 2014, 58(4): 65-70.
- [24] Kim J, Hastak M. Social network analysis: characteristics of on-line social networks after a disaster[J]. *International Journal of Information Management*, 2018, 38(1): 86-96.
- [25] Yoo E, Rand W, Eftekhari M, et al. Evaluating information diffusion speed and its determinants in social media networks during humanitarian crises[J]. *Journal of Operations Management*, 2016, 45: 123-133.
- [26] Son J, Lee H K, Jin S, et al. Content features of tweets for effective communication during disasters: a media synchronicity theory perspective[J]. *International Journal of Information Management*, 2019, 45: 56-68.
- [27] Sutton J, Spiro E S, Johnson B, et al. Warning tweets: serial transmission of messages during the warning phase of a disaster event [J]. *Information, Communication & Society*, 2014, 17(6): 765-787.
- [28] Li L F, Zhang Q P, Wang X, et al. Characterizing the propagation of situational information in social media during COVID-19 epidemic: a case study on Weibo[J]. *IEEE Transactions on Computational Social Systems*, 2020, 7(2): 556-562.
- [29] 陈思菁, 李纲, 毛进, 等. 突发事件信息传播网络中的关键节点动态识别研究[J]. *情报学报*, 2019, 38(2): 178-190.
- [30] Zheng X, Sun A X. Collecting event-related tweets from twitter stream[J]. *Journal of the Association for Information Science and Technology*, 2019, 70(2): 176-186.
- [31] 赵洁, 何文娟, 鲁晓阳. 一种基于新闻学的微博事件特征选择方法[J]. *情报学报*, 2015, 34(4): 424-431.
- [32] Pogrebnikov N, Maldonado E. Didn't realize that: social media message complexity and situational awareness of emergency responders[J]. *International Journal of Information Management*, 2018, 40: 166-174.
- [33] 马超, 李纲, 陈思菁, 等. 基于多模态数据语义融合的旅游在线评论有用性识别研究[J]. *情报学报*, 2020, 39(2): 199-207.
- [34] Yang Y M, Ha H Y, Fleites F, et al. Hierarchical disaster image classification for situation report enhancement[C]// *Proceedings of the 2011 IEEE International Conference on Information Reuse & Integration*. IEEE, 2011: 181-186.
- [35] Alam F, Ofli F, Imran M. Processing social media images by combining human and machine computing during crises[J]. *International Journal of Human-Computer Interaction*, 2018, 34(4): 311-327.
- [36] 杨腾飞, 解吉波, 李振宇, 等. 微博中蕴含台风灾害损失信息识别和分类方法[J]. *地球信息科学学报*, 2018, 20(7): 906-917.
- [37] Rexiline Ragini J, Rubesh Anand P M, Bhaskar V. Mining crisis information: a strategic approach for detection of people at risk through social media analysis[J]. *International Journal of Disaster Risk Reduction*, 2018, 27: 556-566.
- [38] 邱奇志, 周三三, 刘长发, 等. 基于文体和词表的突发事件信息抽取研究[J]. *中文信息学报*, 2018, 32(9): 56-65, 74.
- [39] Cheng Z Y, Caverlee J, Lee K. You are where you tweet: a content-based approach to Geo-locating Twitter users[C]// *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*. New York: ACM Press, 2010: 759-768.
- [40] Dalvi N, Kumar R, Pang B. Object matching in tweets with spatial models[C]// *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*. New York: ACM Press, 2012: 43.
- [41] Lee B, Hwang B Y. A study of the correlation between the spatial attributes on Twitter[C]// *Proceedings of the 2012 IEEE 28th International Conference on Data Engineering Workshops*. IEEE, 2012: 337-340.
- [42] Doersch C, Singh S, Gupta A, et al. What makes Paris look like Paris?[J]. *ACM Transactions on Graphics*, 2012, 31(4): 1-9.
- [43] Vo N, Jacobs N, Hays J. Revisiting IM2GPS in the deep learning era[C]// *Proceedings of the 2017 IEEE International Conference on Computer Vision*. IEEE, 2017: 2640-2649.
- [44] Imran M, Ofli F, Caragea D, et al. Using AI and social media multimodal content for disaster response and management: opportunities, challenges, and future directions[J]. *Information Processing & Management*, 2020, 57(5): 102261.
- [45] Poblete B, Guzmán J, Maldonado J, et al. Robust detection of extreme events using Twitter: worldwide earthquake monitoring[J]. *IEEE Transactions on Multimedia*, 2018, 20(10): 2551-2561.
- [46] Sakaki T, Okazaki M, Matsuo Y. Tweet analysis for real-time event detection and earthquake reporting system development[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2013, 25(4): 919-931.
- [47] Zhou X M, Chen L. Event detection over twitter social media streams[J]. *The VLDB Journal*, 2014, 23(3): 381-400.
- [48] Yilmaz Y, Hero A O. Multimodal event detection in Twitter hashtag networks[J]. *Journal of Signal Processing Systems*, 2018, 90(2): 185-200.
- [49] Daly S, Thom J A. Mining and classifying image posts on social

- media to analyse fires[C]// Proceedings of the International Conference on Information Systems, 2016: 1-14.
- [50] Pohl D, Bouchachia A, Hellwagner H. Online indexing and clustering of social media data for emergency management[J]. *Neurocomputing*, 2016, 172: 168-179.
- [51] Wu Q H, Ma S L, Liu Y Z. Sub-event discovery and retrieval during natural hazards on social media data[J]. *World Wide Web*, 2016, 19(2): 277-297.
- [52] Meladianos P, Xypolopoulos C, Nikolentzos G, et al. An optimization approach for sub-event detection and summarization in Twitter[C]// Proceedings of the European Conference on Information Retrieval. Cham: Springer, 2018: 481-493.
- [53] Rudra K, Banerjee S, Ganguly N, et al. Summarizing situational tweets in crisis scenario[C]// Proceedings of the 27th ACM Conference on Hypertext and Social Media. New York: ACM Press, 2016: 137-147.
- [54] Xu Z, Liu Y H, Yen N Y, et al. Crowdsourcing based description of urban emergency events using social media big data[J]. *IEEE Transactions on Cloud Computing*, 2020, 8(2): 387-397.
- [55] Kuang C, Tang J Y, Liu Z Y, et al. ImgWordle: image and text visualization for events in microblogging services[C]// Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces. New York: ACM Press, 2014: 371-372.
- [56] Xu Z, Liu Y H, Zhang H, et al. Building the multi-modal storytelling of urban emergency events based on crowdsensing of social media analytics[J]. *Mobile Networks and Applications*, 2017, 22(2): 218-227.
- [57] 李显鑫. 基于图文相关的突发事件图片采集研究[D]. 武汉: 武汉大学, 2020.
- [58] Bian J W, Yang Y, Zhang H W, et al. Multimedia summarization for social events in Microblog stream[J]. *IEEE Transactions on Multimedia*, 2015, 17(2): 216-228.
- [59] Schinas M, Papadopoulos S, Kompatsiaris Y, et al. Visual event summarization on social media using topic modelling and graph-based ranking algorithms[C]// Proceedings of the 5th ACM on International Conference on Multimedia Retrieval. New York: ACM Press, 2015: 203-210.
- [60] Qian X M, Li M D, Ren Y Y, et al. Social media based event summarization by user-text-image co-clustering[J]. *Knowledge-Based Systems*, 2019, 164: 107-121.
- [61] 刘雅姝, 张海涛, 徐海玲, 等. 多维特征融合的网络舆情突发事件演化话题图谱研究[J]. *情报学报*, 2019, 38(8): 798-806.
- [62] 安璐, 杜廷尧, 余传明, 等. 突发公共卫生事件的微博主题演化模式和时序趋势——以 Twitter 和 Weibo 的埃博拉微博为例[J]. *情报资料工作*, 2016(5): 44-52.
- [63] Huang Q Y, Cervone G, Zhang G M. A cloud-enabled automatic disaster analysis system of multi-sourced data streams: an example synthesizing social media, remote sensing and Wikipedia data[J]. *Computers, Environment and Urban Systems*, 2017, 66: 23-37.
- [64] 赵又霖, 庞烁, 吴宗大. 社会感知数据驱动下突发事件应急管理时空语义模型构建研究[J]. *情报科学*, 2021, 39(2): 44-53.
- [65] 吴林, 安璐, 孙冉. 面向企业舆情监测的事件画像与高危人群预测研究[J]. *信息资源管理学报*, 2020, 10(1): 15-28.
- [66] 夏立新, 陈健瑶, 余华娟. 基于事理图谱的多维特征网络舆情事件可视化摘要生成研究[J]. *情报理论与实践*, 2020, 43(10): 157-164.
- [67] 张晨昕, 饶元, 樊笑冰, 等. 基于社交媒体的事件脉络挖掘研究进展[J]. *中文信息学报*, 2019, 33(11): 15-30.
- [68] Xu S Z, Kong L, Zhang Y. A picture paints a thousand words: a method of generating image-text timelines[C]// Proceedings of the 21st ACM International Conference on Information and Knowledge Management. New York: ACM Press, 2012: 2511-2514.
- [69] Utz S, Schultz F, Glocka S. Crisis communication online: how medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster[J]. *Public Relations Review*, 2013, 39(1): 40-46.
- [70] Kumar A, Garg G. Sentiment analysis of multimodal twitter data[J]. *Multimedia Tools and Applications*, 2019, 78(17): 24103-24119.
- [71] Huang F R, Wei K M, Weng J, et al. Attention-based modality-gated networks for image-text sentiment analysis[J]. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2020, 16(3): 1-19.
- [72] 曹树金, 岳文玉. 突发公共卫生事件微博舆情主题挖掘与演化分析[J]. *信息资源管理学报*, 2020, 10(6): 28-37.
- [73] 安璐, 欧孟花. 突发公共卫生事件利益相关者的社会网络情感图谱研究[J]. *图书情报工作*, 2017, 61(20): 120-130.
- [74] 安璐, 胡俊阳, 李纲. 基于主题一致性和情感支持的评论意见领袖识别方法研究[J]. *管理科学*, 2019, 32(1): 3-13.
- [75] Ragini J R, Anand P M R, Bhaskar V. Big data analytics for disaster response and recovery through sentiment analysis[J]. *International Journal of Information Management*, 2018, 42: 13-24.
- [76] Purohit H, Castillo C, Diaz F, et al. Emergency-relief coordination on social media: automatically matching resource requests and offers[J/OL]. *First Monday*, 2014, 19(1). (2013-12-28). <https://doi.org/10.5210/fm.v19i1.4848>.
- [77] Harb J G D, Ebeling R, Becker K. A framework to analyze the emotional reactions to mass violent events on Twitter and influential factors[J]. *Information Processing & Management*, 2020, 57(6): 102372.
- [78] 徐永顺, 周宇, 刘渊, 等. 重大突发公共卫生事件中的居民情感需求挖掘[J]. *图书馆论坛*, 2021, 41(9): 76-86.
- [79] Yang T F, Xie J B, Li G Q, et al. Social media big data mining and spatio-temporal analysis on public emotions for disaster miti-

- gation[J]. ISPRS International Journal of Geo-Information, 2019, 8(1): 29.
- [80] Kryvasheyeu Y, Chen H, Obradovich N, et al. Rapid assessment of disaster damage using social media activity[J]. Science Advances, 2016, 2(3): e1500779.
- [81] Liang Y, Caverlee J, Mander J. Text vs. images: on the viability of social media to assess earthquake damage[C]// Proceedings of the 22nd International Conference on World Wide Web. New York: ACM Press, 2013: 1003-1006.
- [82] Kropivnitskaya Y, Tiampo K F, Qin J H, et al. Real-time earthquake intensity estimation using streaming data analysis of social and physical sensors[J]. Pure and Applied Geophysics, 2017, 174(6): 2331-2349.
- [83] Nguyen D T, Ofli F, Imran M, et al. Damage assessment from social media imagery data during disasters[C]// Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017. New York: ACM Press, 2017: 569-576.
- [84] Li X K, Caragea D, Zhang H Y, et al. Localizing and quantifying damage in social media images[C]// Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. IEEE, 2018: 194-201.
- [85] de Albuquerque J P, Herfort B, Brenning A, et al. A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management[J]. International Journal of Geographical Information Science, 2015, 29(4): 667-689.
- [86] Arthur R, Boulton C A, Shotton H, et al. Social sensing of floods in the UK[J]. PLoS One, 2018, 13(1): e0189327.
- [87] Brouwer T, Eilander D, van Loenen A, et al. Probabilistic flood extent estimates from social media flood observations[J]. Natural Hazards and Earth System Sciences, 2017, 17(5): 735-747.
- [88] Li Z L, Wang C Z, Emrich C T, et al. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods[J]. Cartography and Geographic Information Science, 2018, 45(2): 97-110.

(责任编辑 魏瑞斌)