引文格式：梁明,晏实江,杨根,等. 网络共享文档的雨洪灾害“求救”事件时空分析[J]. 地理空间信息, 2024, 22(6): 66-70. doi:10. 3969/j. issn. 1672-4623.2024.06.015

网络共享文档的雨洪灾害“求救”事件时空分析

梁　明[1,2]，晏实江[1,2]，杨　根[1]，黄晴晴[1]
（1. 安徽大学 资源与环境工程学院，安徽 合肥 230601；2. 自然资源部城市国土资源监测与仿真重点实验室，广东 深圳 518034）

摘　要：城市雨洪灾害频发、急需快速掌握灾害过程中“求救”事件类型、位置，并开展制图。提出了基于“网络共享文档”的文本挖掘方法。为分析不同“求救”事件在时空上分布特征，提出了核密度估计、最近邻分析等方法。以河南雨洪灾害过程中传播较为广泛的网络共享文档为例，验证了方案的可行性。结果表明：城市雨洪灾害过程中不同类型的“求救”事件在的扩散规律有较大差异；大多数求救事件在时间和空间尺度上都呈现为聚集的模式；基于网络共享文档的灾害分析和应急制图，能够在时效性、时间粒度等方面弥补传统雨洪灾害监测手段的不足。

关键词：网络共享文档；城市雨洪灾害；时空分析；核密度估计
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2024）06-0066-05

Spatio-temporal Analysis of “Call for Help” Events of Rain and Flood Based on Network Shared Document

LIANG Ming[1,2], YAN Shijiang[1,2], YANG Gen[1], HUANG Qingqing[1]
(1. School of Resources and Environmental Engineering, Anhui University, Hefei 230601, China;
2. Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, Shenzhen 518034, China)

Abstract: Urban rain and flood disasters occur frequently, and it is urgent to quickly grasp the type and location of “call for help” events in the process of disasters, and to carry out mapping. We proposed a text mining method based on network shared document. We used kernel density estimation and nearest neighbor analysis methods to analyze the spatio-temporal distribution characteristics of different “call for help” events. Taking the widely disseminated network shared documents during the rain and flood disaster in Henan for example, we verified the feasibility of the scheme. The results show that different types of “call for help” events in the process of urban rain and flood disasters have great differences in the diffusion rule. Most of the rescue events show a clustering pattern in time and space. The disaster analysis and emergency mapping based on network shared document can make up for the shortcomings of traditional rain and flood disaster monitoring methods in terms of timeliness and time granularity.

Key words: network shared document, urban rain and flood disaster, spatio-temporal analysis, kernel density estimation

全球变化的加剧进一步推动了极端气象灾害发生的频率[1]。由于城市内的人口密集、部分城市排涝能力薄弱等原因，城市内的雨洪灾害带来的人员伤亡和财产损失在诸多极端气象灾害中尤为突出[2]。城市内短时间的超强降水往往导致电力、通讯、交通等一系列与城市救援紧密相关的设施、设备的损坏。这些设施的瘫痪使得在城市雨洪灾害求援过程中，高定位精度、高时效性的灾害和求救信息匮乏，极大地延宕了救援进度。

在传统的洪灾救援中，高度依赖较大尺度的卫星遥感、无人机航拍等手段[3-4]。然而，近年来城市雨洪灾害呈现出突发性、人员密集性、短时降水量巨大和致灾速度快等特点。依赖卫星遥感的城市雨洪灾害分析和制图受到卫星影像空间分辨率较低、卫星重返周期过长等方面的限制，在时效性和空间分辨率方面已经难以满足城市雨洪救灾的需求[5]。虽然无人机具有高时效、高空间分辨率的特点，对于获取较小尺度的准实时灾害信息优势明显[6-7]，但是，由于城区内建筑密度高、地下空间多、地铁和立交桥错综复杂、监测死角多，致使无人机在城市雨洪灾害的应急救援中的

收稿日期：2022-06-21。
项目来源：安徽省重点研究与开发计划资助项目（2022l07020027）；安徽省自然科学基金资助项目（1908085QD164）；安徽省高等学校自然科学研究资助项目（2022AH050095）；衡阳师范学院创新平台开放基金资助项目（HIST20K02）。
第一作者简介：梁明（1984—），博士，讲师，主要研究方向为时空数据挖掘、自然灾害监测与管理、三维GIS，E-mail：lming@mail.ccnu.edu.cn。

应用受到了较大限制。因此，相关人员又期望通过摄像头和物联网传感器来加以弥补[8-9]。但是，在特大的城市雨洪灾害中，摄像头和传感器网络可能会完全瘫痪，无法为城市雨洪救援提供帮助（以中国郑州2021年7月22日洪灾为例）。

“网络共享文档”的类型多样、使用场景广泛。近年来，应用微博、微信、推特等社交媒体进行灾害分析或灾害制图已经受到相关研究人员的重视[10-11]。其中，文献[12]专门研究了在地震、火山等自然灾害发生后的基于网络文档的信息传播规律；文献[13]和[14]侧重于探讨基于文本的方式从海量的网络文档中提取灾害主题。此类文档的分析多数侧重于探索如何判断文档是否为“灾害”主题，而极少在更细语义尺度上关注灾害事件的发生、发展阶段和救援需求的分类[15]。此外，还有相当一部分研究，仍然着眼于网络文档信息在“事后评估”中的应用，而较少涉及实时或准实时的救灾需求[16-17]。在郑州7.22城市雨洪灾害中，“网络共享文档”发挥的作用某种程度上替代了部分预设的监测网络，为应急抢险和人员安置提供了位置、类型和数量等方面的指南。这种基于“网络共享文档”的灾害信息获取手段有2个突出的优点：一是不受供电和布线的约束；二是实时性强。此外，网络共享文档中的信息大多是具体的“求救”或“求助”需求，其指示的空间尺度较小。正因如此，此类信息较好地弥补了基于卫星、无人机等大尺度监测手段的不足，同时也利于在更细的时空粒度上探讨雨洪灾害的发生、发展、扩散和消退等时空特征，从而精准把握城市雨洪灾害的时空规律。本文以“网络共享文档”为数据源，基于文本挖掘的方式，探索灾害类型、灾害时间、灾害位置等信息的提取，并将提取结果用于分析雨洪灾害过程中不同类型“求救”事件的时空特征，并开展灾害制图。

1　“网络共享文档”的特点

与传统的微博、推特等完全的非结构化文本结构不同，“网络共享文档”通常是半结构化的。此外，“网络共享文档”在数据采集和信息共享方面的目的性、专题性更突出，而传统的微博、推特等社交文档的主题宽泛、目的性不集中。具体来看，“网络共享文档”具有以下的特点：①主题明确，信息集中。以“救灾”文档为例，文档中的文本内容都聚焦于灾害的发生、现状和求助。此特征有利于数据的搜集、处理、对比和清洗。②信息更新快，信息相互佐证，可靠性较强。在灾害应急中，信息的时效性和可信度对于科学决策至关重要。救灾类“网络共享文档”中，可以实现多用户共同编辑、实时更新，同时文档可以为广泛“共享”传播，能够快速为更广泛的人群获得和查看，时效性强。此外，相似信息或相近位置的信息通常会集中发布在一个或多个文档中，为信息的相互验证提供了可能。而在微博等社交媒体中，这类信息是相互独立分散的。要印证消息之间的相关性需要抓取大量信息，耗时较长[18-19]。③信息的半结构化。救灾类“网络共享文档”通常用类似表格的方式统计相关信息。半结构化的文本结果将人员、时间、地点、状态、需求等进行了初步的分类，为文本挖掘和文本分类提供了便利。微博等传统社交媒体中的救灾信息则大多数是“非结构化”的文本[20]。④定位信息描述不精确。微博、微信等软件通常带有定位功能，或者照片中附加有位置标签[21]。而“网络共享文档”通常用文本的方式来描述灾害发生的位置，这种描述缺乏规范化，描述方法不一、空间尺度差异大。

2　雨洪灾害中“求救”事件的文本挖掘方法

“网络共享文档”的半结构化特征，为基于文本挖掘的城市雨洪灾害的信息提取提供了便利。在此处基于文本的挖掘共包括3个不同的任务。第一，提取“网络共享文档”中每个求救事件的“求救类型”；第二，提取求救事件的位置；第三，提取求救的时间。

首先，为了更准确地分析城市雨洪灾害过程中不同类型的求救事件的时空分布特征，笔者将“求救类型”分为：“人员被困”、“人员失联”、“医疗求助”和“其他”等4个类型。如表1所示为不同“求救类型”的文本描述示例。“求救类型”的提取是一个文本分类问题，可以采用决策树、朴素贝叶斯等机器学习的分类器。然而，一方面笔者收集到的样本个数较少，难以保证训练样本的充足性、多样性；另一方面样本中的文本大多属于“短文本”，自动分类较为困难。因此，笔者直接人工分类的方式对“网络共享文档”中不同求救事件的“求救类型”进行提取。

其次，由于“求救”者的位置信息对于救援、分析和制图都至关重要，因此笔者重点探索了“位置”提取的方法。如前述分析，救灾类“网络共享文档”其位置信息通常单独存储在特定的字段中，为位置信息的提取提供了较大的便利。然而由于不同的用户对文档的理解不同，并且在约定填写位置信息的字段中，也有少部分填写不够规范，从而对位置提取造成

表1 “求救类型”的分类示例

序号	文本描述	求救类型
1	请协助不满1岁的宝宝、1位老人、一名成年人撤离危险区域！接到撤离通知，需要协助撤离到安全地方；一共3人，其中有一个不满一岁的宝宝和一位老人；信号时断时续，打不通救援电话，恳请帮助!!	被困
2	我妈刚做完手术，去医院化疗时被困在医院，两天没有吃饭了，请有能力的人帮帮她，而且她手机关机了，联系不上	失联
3	女生被铁桩挂住流血	医疗
4	小区地下室严重积水，能否支援发电机和抽水泵	其他

困扰。例如，“郑州二七区侯寨收费站被困一辆蓝色荣威SUV 车牌尾号868”中将位置和求救内容写在一起；“郑大一附院，应该是1号住院楼”中将非用户自己的判断性的助词写在地址中间，干扰了对地址的描述。此外，“二七广场附近这个避难所”这种描述，将临时的“避难所”作为地址，无法帮助程序从历史的空间数据库中提取有效的位置信息。而“已安全到家”、“已被救助”等则完全无法提供任何空间位置的有价值的参考。因此，笔者根据对救灾类“网络在线文档”的分析，采用高德地图的在线API的方式，通过输入位置描述，获得高德地图返回的经纬度信息。在输入数据之前，先对文本信息进行初步过滤，将类似“已安全到家”、“已被救助”等记录过滤掉。然后借助API发起“文本”到“经纬度”的映射请求。例如，“郑州市二七区陇海中路104号铁道陇海家园”返回坐标“113.653 502, 34.735 671”，“隆福国际小区”返回坐标“113.653 344, 34.738 688”。由于高德地图对输入的地址描述具有较高的鲁棒性和容错能力，即使不够规范的位置描述，也能够获得一个解析的坐标，例如“郑大一附院和金水河之间西边道路之间，被水流困住”返回坐标“113.644 065, 34.754 07”。这就为基于文本的位置提取提供了较大的便利。

再次，求救时间的提取也非常关键。“网络共享文档”的一个优势就是对于每一条求救记录，都有自动记录的创建时间和更新时间。这些与求救事件一一对应的时间戳为精准分析城市雨洪灾害过程中不同类型的求救事件的时空规律提供了条件。通常，在“网络共享文档”中，创建时间都在特定的字段中，并且精确到秒。因此，只需要读取特定字段即可。

3 雨洪灾害中“求救”事件的时空特征分析方法

研究不同类型“求救”事件在雨洪灾害过程中的发生、发展和蔓延的时空特征，对于“精准施救”和“科学规划”、以及“应急预案”的合理制定都具有重要的意义。例如，洪水涌入狭小的地下停车场，造成人员被困；短时强降水引致的激流将人员冲走，并且随着时间的延续，“断电”、“寻人”都类型事件的报告相较于灾害早期可能会逐渐增加。为了分析不同类型求救事件的“时空特征”，笔者分别从时间和空间维度出发对其进行分析。

从时间维度上看，不同类型求救事件的发生、扩散的规律存在差异，可以采用分时段统计的方法进行分析。从雨洪灾害开始发生的一定时间范围内，划分出若干时间区间；分别统计不同类型的“求救”事件在各个区间上的发生次数或频率；然后制作不同类型事件的“频次–时间”图。从图上观察不同类型事件随着时间的发展各事件自身发生频次的变化，对具有明显的随时间变化特征的事件进行拟合，并对拟合结果开展分析评价。

在城市雨洪灾害中，不同类型求救事件在空间上的分布也是不均衡的。分析不同类型事件在空间上的分布规律，笔者采用核密度估计（kernel density estimation，KDE）的方法。KDE以每个待计算的格网为中心进行圆形区域搜索，并以此计算每个格网的密度值，用于评估对象或事件在空间分布上的相对集中程度[22]。对于雨洪灾害过程中的不同类型的“求救”事件，估计其核密度的公式为：

$$f(x)=\sum_{i=1}^{n}\frac{1}{r^2}k\left(\frac{x-x_i}{r}\right) \tag{1}$$

式中，n 为事件 x 距离小于或等于 r 的个数；k 则为权重；r 为距离衰减的阈值。基于KDE的方法，可以分析不同类型的“求救”事件在空间上的集中趋势，以及不同事件的空间模式是“单核”、“双核”还是“多核”。

平均最近邻（average nearest neighhor，ANN）距

离通常用于度量事件或对象在空间上的聚集程度。笔者用不同类型事件在各时间段上的ANN来度量事件随着时间变化的扩散规律。ANN距离的公式如下：

$$D_{\mathrm{ANN}}=\frac{\sum_{i=0}^{n}d_i}{n} \tag{2}$$

式中，n为某种求救类型事件的个数；d_i为事件i与其空间上最邻近的事件之间的距离，通常用欧氏距离来度量。D_{ANN}为研究区内某类事件的所有最近邻距离的平均值，D的大小反映的是事件的距离程度。在本研究中，通过D_{ANN}的变化来探查在城市雨洪灾害过程中，不同类型的求救事件随时间变化的模式。

4　结果与分析

4.1　不同“求救”事件的时间特征

2021年夏中国河南省发生特大城市雨洪灾害。在整个救灾过程中多个救灾“共享文档”广泛传播。笔者共收集了其中13个传播最广泛的文档，对其进行整理、清洗和文本挖掘。从时间维度看，不同类型求救事件在时间上存在着不同的演化模式。如图1所示，“被困”和“失联”事件在灾害早期的报告最多，并且具有明显伴随着时间线性变化。笔者先对不同类型事件的报告频率进行归一化，再对“被困”和“失联”2种类型事件与事件关系进行线性拟合。结果显示，“被困”和“失联”事件与时间的R^2值分别为69.98%和71.22%。而“医疗”和“其他”事件，没有随着时间推移的显著的线性关系。

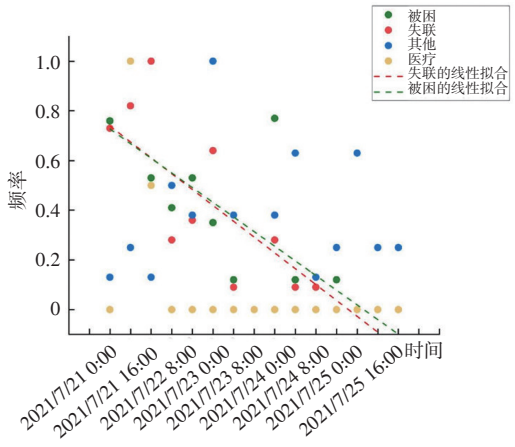


图1　不同类型事件的时间特征

4.2　不同求救事件的空间分布特征

从空间维度看，如图3所示在雨洪灾害爆发的早期，各类求救事件在空间上相对比较集中，并形成显著的高值区域（主要集中在金水区和管城区）；而随着时间的推移，“求救”事件在“网络共享文档”中报告的次数逐渐减少、并且报告位置的集中程度越来越分散。

ANN分析能够进一步量化不同事件随时间推移的空间聚集特征。ANN距离越大，表征着对应的事件在空间上的聚集程度越低；反之，如果ANN越小，则表征对应的事件在空间上的聚集程度越高。度量不同类型求救事件ANN距离随时间的变化规律，能够较为有效地揭示不同类型“求救”事件的扩散规律。不同类型的事件的ANN距离有着较大的差异。“被困”事件和“失联”事件在雨洪灾害早期的空间聚集程度明显高于其他类型事件。而且，随着时间的推移，“被困”事件的ANN距离逐渐增加，但是仍然远低于其他类型事件的ANN距离，这也反映出在此次雨洪灾害过程中，“被困”事件相对来说是高度集中的。“失联”事件则随着时间的推移快速的有集中转为分散。而且，从ANN距离的来看，“失联”事件随着时间急剧发散的速度很快。“被困”类型的事件具有较高的聚集性，从城市管理而言，应在高风险的地区加大对此类事件应急处理的设施或装备的部署。相反，“其他类型事件和“医疗”类型事件在空间的聚集程度相对低，应注意对响应此类求救事件的设备或设施的均匀部署。

4.3　基于WebGIS的“求救”事件制图

地图充分利用从“网络共享文档”提取“灾害类型”和“灾害位置”等制图要素。地图用不同符号直观的表达求救事件的类型以及它们所处的位置。点击各个位置可弹出的提示框中获取更详细的信息以及联系方式。此外，用户可以通过点击安置点的图标获取电话信息实现就近避险，并寻求进一步帮助。

5　结　论

完善城市雨洪等灾害的信息获取手段，提高信息获取的时空粒度和效率，改善灾害制图的方法势在必行。笔者选取选取雨洪灾害中广泛传播的“网络共享文档”作为数据源，探讨灾害过程中“求救”事件的文本挖掘方法。结果表明“网络共享文档”利于提取“求救”类型、“求救”时间和“求救”位置等关键信息。但是，限于“网络共享文档”中的位置信息描述能力弱的特点，建议相关机构或企业在开发的“网络共享文档”中提供类似于微博、微信中的位置获取和位置嵌入功能。

参考文献

[1] Sloggy M R， Suter J F， Rad M R，et al. Changing Climate，Changing Minds? The Effects of Natural Disasters on Public Perceptions of Climate Change[J]. Climatic Change，2021(16):25-51

[2] Dashti S， Palen L， Heris M P，et al. Supporting Disaster Reconnaissance with Social Media Data: A Design-Oriented Case Study of the 2013 Colorado Floods[C]// International Iscram Conference，2014

[3] 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

[4] 王磊，连增增. 基于Sentinel-1A的2020年鄱阳湖流域洪水灾害遥感监测[J]. 地理空间信息，2022,20(6):43 -46

[5] Havas C， Resch B . Portability of Semantic and Spatial-temporal Machine Learning Methods to Analyse Social media for Near-real-time Disaster Monitoring[J]. Natural Hazards，2021(108):2 939-2 969

[6] 郭晨，许强，董秀军，等. 无人机在重大地质灾害应急调查中的应用[J]. 测绘通报，2020(10):6-11

[7] 戴建国，张国顺，郭鹏，等. 基于无人机遥感多光谱影像的棉花倒伏信息提取[J]. 农业工程学报，2019，35(2):63-70

[8] 王星捷，刘华春，李春花. 基于多元平台洪灾报警系统设计与实现[J]. 计算机技术与发展，2017，27(4):196-199

[9] Garcia F， Retamar A E， Javier J C . A Real Time Urban Flood Monitoring System for Metro Manila[C]// TENCON 2015 - 2015 IEEE Region 10 Conference. IEEE,2016

[10] 林继贤,周志刚,陈蔚珊. 基于大数据的内涝智能指挥系统研究与应用[J]. 地理空间信息,2022,20(3):69-73

[11] 赵飞,廖永丰. 突发自然灾害事件网络舆情传播特征及影响因素研究[J]. 地球信息科学学报,2021,23(6):992-1 001

[12] Nicole J. Olynk W,Kendra R,et al. The Anatomy of Natural Disasters on Online Media:Hurricanes and Wildfires[J]. Natural Hazards,2022(10):961-998

[13] Algiriyage N,Sampath R,Prasanna R ,et al. Identifying Disaster-related Tweets:A Large-scale Detection Model Comparison[C]// ISCRAM 2021 Conference Proceedings - 18th International Conference on Information Systems for Crisis Response and Management,2021

[14] Pekar V,Binner J,Najafi H,et al. Selecting Classification Features for Detection of Mass Emergency Events on Social Media[C]//Proceedings of the International Conference on Security and Management (SAM). The Steering Committee of The World Congress in Computer Science，Computer Engineering and Applied Computing (Worldcomp),2016:192-198

[15] 邬群勇,裘钰娇. 微博数据位置信息反映台风灾情的有效性分析[J]. 测绘科学技术学报,2019,4(7):38-49

[16] Lacassin R,Devès M,Hicks S P,et al. Rapid Collaborative Knowledge Building Via Twitter After Significant Geohazard Events[J]. Geoscience Communication，2020，3(1):129-146

[17] Spruce M,Arthur R,Williams H T P. Using Social Media to Measure Impacts of Named Storm Events in the United Kingdom and Ireland[J]. Meteorological Applications，2020,27(1):1-17

[18] 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):95-103

[19] 刘淑涵,王艳东,付小康. 利用卷积神经网络提取微博中的暴雨灾害信息[J]. 地球信息科学学报,2019,21(7):1 009-1 017

[20] 王波,甄峰,孙鸿鹄. 基于社交媒体签到数据的城市居民暴雨洪涝响应时空分析[J]. 地理科学,2020,40(9):1 543-1 552

[21] Middleton S E,Middleton L,Modafferi S . Real-time Crisis Mapping of Natural Disasters Using Social Media[J]. IEEE Intelligent Systems,2014,29(2):9-17

[22] 龙飞,戴学锋,虞虎. 基于路网形态的上海住宿业空间布局特征[J]. 地理科学,2021,41(9):1 569-1 577