

# 新冠肺炎疫情期间社交媒体中情境信息传播的特征:一个案例研究 在微博上

李丽芳、张庆鹏、IEEE 成员、小王、IEEE 高级成员、IEEE、田、蔡锦辉、IEEE 研究员

**摘要**——在冠状病毒疾病(新冠肺炎)持续爆发期间,人们使用社交媒体以前所未有的规模获取和交流各种类型的信息。只有情况信息对于公众和当局应对疫情才是有价值的。因此,重要的是识别这种情况信息,并了解它是在社交媒体上传播的,以便可以为新冠肺炎疫情提供适当的信息发布策略。本文试图通过利用微博数据和自然语言处理技术将新冠肺炎相关信息分类为七种情景信息来填补这一空白。我们在预测每类信息的转发量时发现了具体的特征。结果提供了对信息需求和公众关注的驱动力的见解。

**索引术语**——新冠肺炎、危机信息共享、传染病、信息传播、社交媒体、社交网络分析。

## I. 介绍

URST 出在中国武汉,持续爆发的冠状病毒病(新冠肺炎)已造成区域性和全球公共健康危机[1]。在像新冠肺炎疫情这样的危机中,公众倾向于社交媒体

2020年2月19日收到稿件;2020年3月2日和3月4日修订,2020;2020年3月5日受理。出版日期2020年3月20日;当前版本日期2020年4月3日。本工作得到了国家自然科学基金资助(71672163, 71972164, 61702519, 71701206),卫生与医学研究基金(HMRF)资助(16171991),中央大学项目(C2191110)和国家社会科学重大专项基金(17VZL017)的资助。(通讯作者:张庆鹏。)

李丽芳在华南理工大学公共管理学院工作,中国广州510641(电子邮件:llfang@scut.edu.cn)。张庆鹏就职于香港城市大学数据科学学院香港(电邮:qingpeng.zhang@cityu.edu.hk)。

小王和王飞跃在中国科学院自动化研究所,北京100864,中国(电子邮件:xiao.wang@ia.ac.cn;岳飞@ieee.org)。

和田在武汉大学电气与自动化学院,武汉430072,中国(电子邮件:ianlu.gao@whu.edu.cn)。

国防科技大学系统工程学院,王涛和段威,中国长沙410073(电子邮件:wangtao@nudt.edu.cn;段威@nudt.edu.cn)。

蔡锦辉(Kelvin Kam-fai Tsoi)就职于香港中文大学公共卫生与初级保健学院(电邮:kelvintsoi@cuhk.edu.hk)。

数字对象标识符 10.1109/tcss.25866.188686868616

获取所需信息和交流观点的平台[2],[3]。社交媒体平台上有许多不同类型的信息,情境信息,即帮助有关当局或个人了解紧急情况下的情况的信息(包括可操作的信息,如求助、受影响人数)[4],对公众和当局指导他们的应对措施很有用[5]、[6]。识别这些类型的信息并预测其传播规模将有助于有关当局感知公众的情绪、当局与公众之间的信息差距以及公众的信息需求。这将有助于当局制定适当的应急反应策略[6]。

现有的研究尚未就情境信息的定义达成一致。一些人将求助和捐赠归类为情境信息,但忽略了其他类型的情境信息,如来自公众的批评,这揭示了公众的关注,以及情感支持,这揭示了他人对受害者的同情[7]。然而,有必要以全面的方式识别情境信息。例如,识别与批评相关的信息可以帮助当局了解公众关注的主要问题,并提出适当的应对措施。识别情感支持信息可以帮助当局了解社交媒体用户的社会支持模式,并更好地利用这些自愿资源。

此外,确定能够预测情境信息传播规模的关键特征至关重要,以确保当局能够根据公众的需求发布不同类型的情境信息。在像新冠肺炎这样的突发流行病期间,满足这种信息需求至关重要。

为了填补这些研究空白,我们利用中国主要微博网站新浪微博(相当于中国的Twitter)上与新冠肺炎相关的讨论来回答以下研究问题:

RQ1. 如何识别和分类社交媒体中的情境信息?

RQ2. 不同类型情境信息传播尺度特征的各种可预测性是什么?

## II. 社交媒体中的情境信息

人们广泛使用社交媒体平台在不同情况下分享信息[8]–[12]。在危机期间，社交媒体用户会产生丰富的情境信息[4]、[7]、[13]。

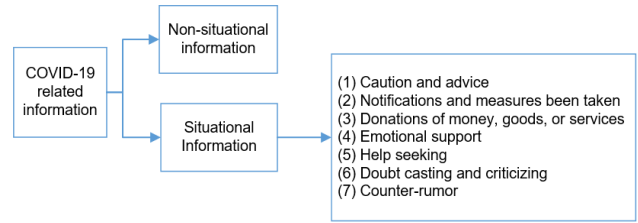
然而，不同的研究者将情境信息分为不同的类型[7]。例如，Rudra 等人[7]将情境信息定义为伤亡或受伤/被困人员或帮助救援行动的通知，并将同情受害者、赞扬或批评救援行动、对危机发生原因的事后分析以及与捐赠相关的信息归类为非情境信息。而在 Vieweg (2012) [14] 的研究中，Rudra 等人[7]定义的非情景信息被归类为情景信息。具体来说，她将情境信息分为社会环境信息、建筑环境信息和物理环境信息。社会环境信息包含建议、警告、疏散、死亡、伤害、医疗护理、人员失踪和提供帮助。建筑环境信息包含由危机和基础设施状态引起的损害。物理环境信息包括环境影响、一般区域信息(危险状态)和一般危险信息(如天气报告)[14]。

此外，Imran 等人[15]进一步将情况信息分为七种类型，如警告和建议、人员伤亡和损失、捐款、货物或服务、失踪、找到或看到的人员，以及基于 Vieweg 工作的信息来源。Mukkamala 和 Beck [4]对 10 种情境信息进行了分类，包括关于灾难的初始信息、情境更新、对注意力不足的批评、道德支持、准备工作、批评和控制谣言、请求帮助、提供帮助、自组织支持和主动志愿服务。

根据 Vieweg [14]的说法，情境信息是提供“战术性、可操作性信息的帖子，这些信息可以帮助人们做出决策，建议他人如何从各种来源获得特定信息，或者为受大规模紧急事件影响的人提供即时的影响后帮助。”分享这些信息有助于有关当局和个人理解危机并指导他们的行为[7]。基于这一定义，本研究将七种与新冠肺炎相关的信息归类为情境信息：

- 1) 注意事项和建议；2) 通知和采取的措施；3) 捐赠金钱、货物或服务；4) 情感支持；5) 求助；6) 质疑铸造与批判；和
- 7) 反谣言。图 1 展示了内容类型。

具体来说，警告和建议信息通知公众，以帮助他们保护自己免受病毒的危害。新冠肺炎的通报(情况更新或伤亡和损失)告诉公众疫情的详细情况，分享这类信息有助于他们了解情况，并帮助他人缓解信息不足的焦虑[13]，[15]。捐赠(提供帮助)信息帮助公众，尤其是那些需要帮助的人



图一。新浪微博上与新冠肺炎相关的信息类型。

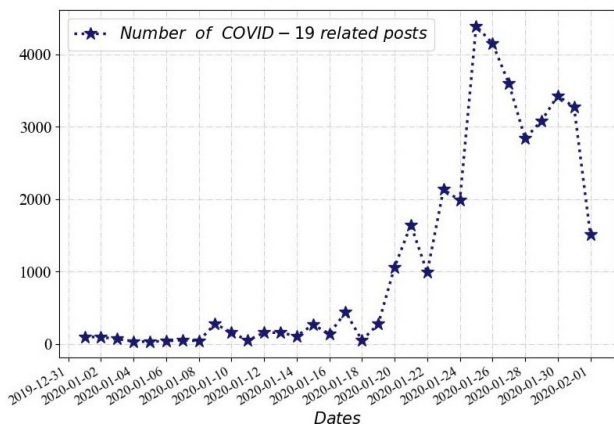
了解可以获得哪些帮助[15]。情感支持信息对受害者从疫情造成的情感伤害中恢复有积极作用，分享这类信息有助于他人获得集体支持并感受到同理心[4]。求助相关的帖子在疫情期间传播有关紧急帮助或援助的信息，分享这类信息有助于当局和个人获得帮助或援助[17]。此外，质疑和批评信息通常会讨论危机的社会政治原因、影响和责任，分享此类信息有助于他人验证信息的有效性或增强他们对疫情的了解[16]，[23]。反谣言信息也有助于公众了解真相，减少谣言造成的混乱[4]。鉴于分享这七种类型的信息有利于提高灾难/危机救援过程的效率，我们将它们归类为情景信息。

## III. 预测传播的特征

### 社交媒体中的危机信息

应用多个特征来预测危机中社交媒体内容的传播规模，包括内容特征和用户特征。内容特征主要包含以下内容：是否包含 URL/Hashtag、内容的发布时间以及内容的长度已被认为会积极影响社交媒体内容的转发量[18]–[20]。此外，内容类型也可能显著影响转发的信息量，因为之前的研究发现，用户在共享不同类型的信息时表现出不同的信息需求和互动[21]。此外，内容的情感也影响信息的信息传播规模[20]。例如，Berger 和 Milkman [19]描述了内容中的情绪化、积极、敬畏、愤怒、焦虑和悲伤词汇对社交媒体内容的病毒式传播的影响。

从社交媒体用户的角度来看，用户是否拥有较高数量的关注者或被关注者(社会资本的代理人)以及是否是验证用户也正向影响帖子在社交媒体中的传播规模[26]，[27]。另一个与用户相关的特征是用户的位置，如果用户来自发达城市，他们更有可能吸引更多的社会资本，这将扩大他们对他人的影响[28]。此外，在发生灾难的情况下，人们更有可能分享来自事件目击者和位于事件附近的用户的帖子[29]。此外，人们的从属关系和看法也会影响用户的信息共享行为[22]–[25]。如建议的那样



图二。每天与新冠肺炎相关的帖子数量。

根据李等人[20]的研究，表明个人与其他人之间的密切和友好关系的从属关系，可能正与人们与更大的受众分享的内容有关。用户的感知(看、听、感觉)也会影响用户的信息分享行为[25]。

#### IV. 数据收集和处理

##### A. 数据收集

本文在武汉新冠肺炎疫情开始后，使用微博 API 收集了使用关键词“新型冠状病毒()”和“不明肺炎()”的微博帖子。根据当局 2020 年 2 月 17 日的报告，这种病毒已经夺走了 1772 人的生命，超过 70 000 人受到感染。我们收集了从 2019 年 12 月 30 日到 2020 年 2 月 01 日的微博数据的以下属性:1) 帖子总数为 36746; 2) 帖子的属性，包括帖子的创建时间、帖子的独特 ID、文本内容、帖子的长度、转发/评论/喜欢以及帖子的数量; 以及 3) 用户属性，诸如用户 id、标记位置、验证、追随者计数和被追随者计数。此外，我们按日期绘制了时间跨度内与新冠肺炎相关的帖子数量，以揭示公众意见的演变趋势，如图 2 所示。结果显示，民意在 1 月 19 日开始加速传播，在 1 月 25 日达到顶峰，从 2 月 1 日开始下降。这表明当局需要在疫情早期更多地关注内容类型，以了解公众的需求并调整他们的信息恰当地发布策略。

##### B. 情境信息分类

各种自然语言处理方法被用于将社交媒体内容分为几种类型[30]、[31]。具体来说，我们使用监督学习方法，如支持向量机(SVM)、朴素贝叶斯(NB)和随机森林(RF)，根据标记数据学习未标记数据的类型[21]、[30]、[31]。分类过程如下。

表一  
样本数据的人工标记结果

Types	Names and definitions	Manual label#
Type 1	Caution and advice: precautions to explain the face of the epidemic should pay attention to what aspects, such as frequent hand washing, wearing masks, less out of the door or the recommendations of responding to the outbreak of the crisis.	526
Type 2	Notifications or measures been taken: outbreak announcements or the measurement already taken by the relevant departments, such as how many cases have occurred, the characteristics of the virus, material reserves, etc.	957
Type 3	Donation of money, goods, or services: donations or wishes to donate materials, money, or services for outbreak prevention and control.	182
Type 4	Providing emotional support: praise or show sympathy to others such as medical team, public welfare organizations, celebrities, and the masses who supporting Wuhan.	255
Type 5	Help seeking: (a) Medical institutions, individuals, etc. to seek support needs, etc. (b) Seek emotional support such as to seek comfort, or to express depression, etc.	262
Type 6	Doubt casting and criticizing: to question local government officials for inaction, the central government, the Red Cross and other related initiatives, or some of the public to mislead others.	522
Type 7	Refute rumors: in response to recent rumors.	52
Type 8	Non-situational information: information that are not related to the crisis.	244

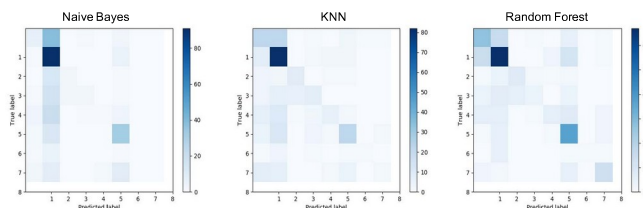


图 3。不同分类器的混淆矩阵。

首先，我们从收集的数据集中随机抽取了 3000 篇与新冠肺炎相关的帖子。

第二，三个研究生手动标注 3000 个采样数据的信息类型。

第三，我们计算了三个编码者的 Cohen Kappa 值，该值为 0.81，大于 0.8，表明标记结果是可信的。表 I 给出了标记结果和每种类型信息的定义。

第四，使用随机采样的 85% (2550) 数据来训练 SVM、NB 和 RF 分类器，并使用剩余的 450 个来测试每个分类器的准确度，这些分类器的平均准确度分别为 0.54、0.45 和 0.65。由于我们有八种类型，我们进一步绘制了每个分类器的混淆矩阵，如图 3 所示。结果表明，RF 分类器的性能最好。

第五，我们使用所有 3000 个标记的数据作为训练集，选择 RF 分类器来自动标记剩余的数据。第六，我们总结了表二所示的五个属性，以揭示一般公众的信息需求。这些属性包括帖子的总(平均)数量、已验证用户的数量(已验证用户的比例



表二  
每种信息的属性摘要

Features	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7
affect	5.324	4.417	3.870	6.269	7.021	6.791	3.802
posemo	2.536	1.964	2.323	3.867	2.769	2.494	1.308
negemo	2.369	2.051	0.831	1.826	1.369	3.470	2.152
anx	0.450	0.454	0.236	0.373	0.339	0.677	0.218
anger	0.232	0.215	0.115	0.303	0.150	0.703	0.298
sad	0.159	0.148	0.115	0.208	0.150	0.427	0.063
percept	2.377	1.865	1.355	2.429	1.243	2.601	1.433
see	1.218	0.775	0.668	1.175	0.460	0.799	0.369
hear	0.451	0.471	0.326	0.625	0.375	0.921	0.698
feel	0.391	0.350	0.188	0.309	0.212	0.458	0.172
affiliation	1.796	1.501	2.328	2.503	6.350	1.906	2.668
reward	2.297	2.034	2.041	2.270	1.663	2.421	1.683
risk	3.410	3.113	3.171	3.785	2.563	5.304	3.254
drives	0.308	0.331	0.264	0.562	0.306	0.425	0.135
achieve	1.635	1.230	0.956	1.264	1.624	1.558	0.986
power	1.796	1.501	2.328	2.503	6.350	1.906	2.668
Verified	0.712	0.623	0.800	0.573	0.681	0.274	0.624
Followers(log)	10.783	10.004	11.742	9.510	10.298	7.367	10.153
Followees (log)	6.131	6.084	6.268	5.994	6.059	5.768	6.065
BigCity	0.145	0.133	0.172	0.131	0.093	0.113	0.133
NearCity	0.040	0.042	0.053	0.060	0.096	0.066	0.024
Hash	0.593	0.687	0.650	0.694	0.594	0.423	0.412
URL	0.000	0.000	0.000	0.000	0.000	0.000	0.000
length	54.798	53.135	58.083	42.649	49.181	47.742	51.364
hours	655.32	650.725	692.757	651.789	689.316	664.554	626.003

表三  
每种情景信息提取特征的平均值

Types	Amounts	Verified	Followers amount	Followees amount	Reposted amount
Type 1-Caution and advice	5669	4109 (72.4%)	1.46e+11 (2.57e+6)	531590 (937.7)	464457 (81.9)
Type 2-Notifications or measures been taken	14419	9150 (63.5%)	3.45e+11 (2.40e+6)	12850699 (891.2)	874427 (60.6)
Type 3-Donations of money, goods, or services	1826	1463 (80.1%)	7.13e+10 (3.90e+6)	1801596 (986.6)	170975 (93.6)
Type 4-Emotional support	2022	1154 (57.1%)	3.03e+10 (1.50e+6)	1722102 (851.7)	220073 (108.8)
Type 5-Help seeking	1474	1003 (68.0%)	3.15e+10 (2.14e+6)	1183222 (802.7)	239573 (162.5)
Type 6-Doubt-casting and criticizing	8297	6048 (72.9%)	5.99e+10 (7.22e+5)	4912385 (592.1)	747664 (90.1)
Type 7-Counter-rumors	173	110 (63.6%)	6.87e+9 (3.97e+6)	143641 (830.3)	38453 (222.3)
Type 8-Non-situational information	2866	1655 (57.7%)	4.65e+10 (1.62e+6)	2230235 (778.2)	2937433 (1024.9)

用户)、用户的关注者和被关注者的总(平均)数量,以及每种类型的信息中帖子的总(平均)转发量。

表二显示,在新冠肺炎疫情期间,2类信息传播过程中涉及的已核实用户较多(通知和采取的措施)。经过验证的用户主导了第三类信息(金钱、商品和服务的捐赠)的传播。类型

5信息(寻求帮助)吸引了最多的平均态度。类型6信息(质疑和批评)在我们的数据集中数量第二大,并且有更大比例的用户参与这类信息的传播。这些发现表明,如 Atlani-duault 等人[31]所建议的,有必要确定此类信息的进一步信息发布策略。

此外,在我们的数据集中,第8类信息(非替代)的转贴总量最大。通过追溯数据集,我们发现占主导地位的帖子是由“央视新闻”发布的,该新闻在2020年1月24日(农历除夕)发布了这条信息,向所有用户发送问候,吸引了2 774 936次转发量(占有转发量的94.5%)。删除这个异常值后,剩下1 62 497个总转发量和56.9个平均转发量。那么与情境信息相比,这种非情境信息成为最不受欢迎的信息。这表明在这次危机中,用户更倾向于转发情境信息而不是非情境信息。

### C. 态势信息传播规模预测

为了找到能够准确预测每类情境信息传播规模(转发量)的关键特征,我们提取了文献[33]、[32]中总结的五组特征。提取的特征(名称为斜体)如下。

- 1) 情绪因素:帖子里的情感、负面情绪(negemo)、正面情绪(posemo)、焦虑(anx)、愤怒、悲伤词(sad)。
- 2) 感知相关因素:感知、看见、听见和感觉。
- 3) 从属关系相关因素:从属关系、成就、权力、回报和风险。
- 4) 用户相关特征:关注者量(关注者(日志))、被关注者量(被关注者(日志))、活动附近与否(附近城市)、是否居住在发达城市(大城市)、是否验证用户。
- 5) 内容相关因素:是否包含 Hash tag/URL(URL, Hash),帖子的长度(length),帖子的发布时间(hours)。

具体来说,情感因素、感知因素和从属因素是使用语言查询和字数统计(LIWC)提取的,这是一种从内容中提取语言信息的常用工具[32], [33]。因为人们在文本中使用的词汇和语言模式可以揭示他们的情感、感知以及他们对归属感、权力、成就等的需求[32]。

追随者/被迫追随者的数量是对数转换的,1是为了避免零。用户的验证状态直接从数据集获得。用户的位置通过以下定义获得:1)用户是否位于危机附近定义为:如果用户位于湖北省,将变量赋1;否则,分配0和2)用户是否位于发达地区,赋值1,即如果用户位于“北京、上海、深圳、广州”;否则,赋值0。

对于与内容相关的因素URL/Hashtag,如果它包含至少一个URL或Hashtag,则为其分配1;否则,分配0。

帖子的长度是帖子的总字数。发帖时间定义为发帖时间与2019年12月30日之间的小时数。表三总结了所有类型的情况信息的所有特征的平均值。

表四  
负二项式和线性 RMSE  
预测回归

RMSE	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7
Linear Regression	414	381	415	576	1517	2369	805
Negative Binomial	420	385	420	598	1537	2384	854

表五  
最佳回归的特征和系数  
情境信息模型

Features	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7
affect					-0.03**		
negemo						-0.01**	-0.02
anx							-0.12
anger							-0.006
sad		0.03*					
percept		0.02***					0.11
sec	0.06***					0.03***	-0.24
hear							-0.09
feel		-0.03*			-0.09.		
affiliation		0.02**	0.04**				
reward		0.04**			0.13**		
risk		0.04***			0.10***		0.038
drives		-0.03***		0.01*	-0.06*		
achieve	-0.04***			-0.05***			
power		0.03***	0.02		0.07**		
Verified	-0.65***	-0.57***	-0.5***	-0.4***	-0.68***	-0.29***	
Followers(log)	0.31***	0.25***	0.28***	0.28***	0.35***	0.28***	0.29***
Followees(log)				-0.06*		-0.05***	
BigCity	0.31***	0.56***	0.56***	0.29***		0.23***	0.64.
NearCity				-0.2.			
Hash		0.16***	0.23**			0.26***	
length	0.004***	0.001**	0.004***	0.001*	0.002*	0.002***	
hours		-0.001***	-0.005***		-0.002***	0.0003*	
RSE	1.376	1.25	1.43	1.275	1.565		1.267
Adjusted R2	0.369	0.35	0.377	0.364	0.393		0.585

Significant as 0.1 level; \*Significant at 0.05 level; \*\* Significant at 0.01 level; \*\*\*Significant at 0.001 level.

使用上面定义的每种情况信息的特征，我们首先使用线性回归、RF 和逐步方法选择特征。通过比较每种类型信息的所选特征的性能，我们最终为类型 1-类型 6 信息选择逐步方法，但为类型 7 信息(反谣言)使用 RF 特征选择结果，因为其拟合优度较高(R2)。此外，我们使用 85%的随机采样数据作为训练数据，使用选择的特征进行多元线性回归和负二项式回归，以预测每种类型信息的对数转换转贴量。表四显示了均方根误差(RMSE)每个预测模型。

由于线性回归模型的 RMSE 较低，我们最终选择线性回归来预测新冠肺炎相关帖子的转发量。此外，表五显示了所选特征对所有七种情况信息的影响。具体来说，第一列显示了潜在的特征，而左边的列显示了所选择的特征及其对每种类型的情境信息的影响。例如，对于类型 1 信息(注意和建议)，主要选择六个特征，即，查看、实现、验证、关注者(日志)、大城市和小时。

表五显示了以下情况。

- 1) 对于类型 2(通知和已经采取的措施)、类型 3(金钱、商品或服务的捐赠)和类型 5(寻求帮助)信息，散列标签的使用将扩大其转贴金额的转贴金额。对于这些类型的信息，通常 it

出版它是当局的责任。如果当局希望它被大量转发，使用标签将是一个不错的选择。

- 2) 未经核实的用户的新新冠肺炎相关帖子转发了大量几乎所有类型的灾害相关信息。因此，当局需要更加关注未经验证的用户的可信度，因为他们的影响力很大。由于忽视了他们职位的可信度，这将对当局造成伤害。
- 3) 如果用户有更多的追随者，来自发达城市，或使用较少的负面词汇，类型 6 信息(质疑和批评)将有更大的转发量。此外，与其他类型的情景信息相比，怀疑铸造信息在危机后期变得流行。对于当局来说，最好是更多地关注那些拥有更多追随者的理性批评者(他们可能是意见领袖)，并验证他们的观点是否有价值。如果是这样，接受他们的建议来改进现有的危机应对策略。
- 4) 对于所有类型的情境信息(除了类型 7，反谣言)，内容包含的字数越多，其转发量越大。对于除反谣言外的各类情境信息，增加长度可以放大其传播规模；否则，使用相反的策略。至于反造谣，需要放大样本数据来验证是否用较少的文字来放大其传播尺度更好。
- 5) 对于 7 类信息(反造谣)，如果来自拥有较高关注人数的用户，且来自发达城市，则会放大其转发量。如果当局想扩大反谣言的传播规模，最好通过提及或回复的方式针对高关注人数用户。

总之，当局可以应用本文的结果来了解公众，如他们对当局当前疫情应对策略的态度，以识别和放大公众可能需要的求助、捐款和通知信息，并识别和反击指责或谣言的企图，以改进当局未来的危机信息发布策略。

## V. 结论

本文的发现表明了对不同类型的情境信息使用不同的信息发布策略的必要性。为不同类型的情况信息选择的特征也可以帮助当局学习如何组织他们的与新冠肺炎相关的帖子，以增加或减少他们的帖子的转发量。此外，情境信息的定义将有助于旨在建立有效的基于社交媒体的应急响应计划和危机信息系统的研究人员或从业人员。

这篇文章有局限性。首先，由于新浪 API 的限制，我们只获得了社交媒体数据的子集。将来，

我们将与数据提供商合作,以获取更多的数据。第二,在训练分类器以识别情境信息的内容类型时,由于数据有限,我们只训练了三个传统的基于自然语言处理的分类器。未来,我们将使用更多的数据和训练深度学习方法,以更高的准确性识别内容类型[34],[35]。第三,危机数据的人工标注非常耗时,可能会影响危机信息共享的效率。未来,我们计划应用自动标记方法来避免这种限制[6],[36]。

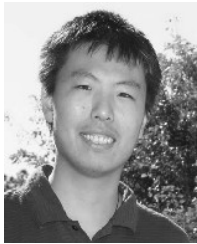
### 参考

- [1] J. T. Wu, K. Leung and G. M. Leung, “对源自中国武汉的 2019-nCoV 爆发的潜在国内和国际传播的临近预报和预测:建模研究”,《柳叶刀》,第 395 卷,第 10225 期,第 689-697 页,2020 年 2 月。
- [2] 页 (page 的缩写) Burnap 等人,“推特上的恐怖:模拟社交媒体对伍尔维奇恐怖袭击的反应”。肚门。《采矿》,第 4 卷,第 1 期,第 1-14 页,2014 年 12 月。
- [3] 南 Vieweg, A. L. Hughes, K. Starbird and L. Palen, “两次自然灾害事件期间的微博”,正在进行中。第 28 国际。糖膏剂。因子计算。系统。(池),2010 年,第 1079 页。2010
- [4] A. Mukkamala 和 R. Beck, “社交媒体在应对自然灾害的集体行为发展中的作用”。26 日欧元。糖膏剂 Inf. 系统。超越数字。刻面。(ECIS),2018 年,第 1-18 页。
- [5] 米 (meter 的缩写) Martínez-Rojas, M. D. C. Pardo-Ferreira 和 J. C. Rubio-Romero, “推特作为管理和分析紧急情况工具:系统的文献综述”,Int. J. Inf. 管理。第 43 卷,第 196-208 页,2018 年 12 月。
- [6] 长度严和 A. J. 佩德拉萨-马丁内斯,“灾害管理的社交媒体:社会对话的操作价值”,产品。操作。管理。第 28 卷,第 10 期,第 2514-2532 页,2019 年 10 月。
- [7] K. Rudra, S. Ghosh, N. Ganguly, P. Goyal and S. Ghosh, “在灾难事件中从微博中提取情景信息:一种分类-总结方法”,正在进行中。第 24 届国际会议。糖膏剂 Inf. 知道了。管理。(CIKM),2015 年,第 583-592 页。
- [8] F.-王,曾,张,韩德乐,曹军,“中国人的人肉网络:第一个十年及以后”,中国。Sci. 公牛。第 59 卷,第 26 期,第 3352-3361 页,2014 年 9 月。
- [9] 张,王凤英,曾丹,王,“理解群体驱动的搜索群体:社会网络视角”,《公共科学图书馆·综合》,2012 年第 7 卷第 6 期,论文编号 e39749。
- [10] 张,曾德辉,王凤英,王东平,“理解群体驱动的搜索群体:一个社会网络的视角”,PLoS ONE. 第 7 卷,第 6 期,2012 年 6 月,货号 e39749,doi:10.1371/journal.pone.0039749。
- [11] F.-Y. Wang 等,“人肉搜索引擎研究:网络知识的众包扩张”,《计算机》,第 43 卷第 8 期,第 45-53 页,2010 年 8 月。
- [12] 张,曾迪,王凤英, R. Breiger 和 J. A. 亨德勒,《经纪人还是桥梁?探索众包系统中的结构漏洞》,第 49 卷第 6 期,第 56-64 页,2016 年 6 月,doi:2016.166。
- [13] R. Dutt, M. Basu, K. Ghosh 和 S. Ghosh, “通过匹配资源需求和可用性利用微博协助灾后救援行动”,Inf. 过程。管理。第 56 卷,第 5 期,第 1680-1697 页,2019 年 9 月。
- [14] 南 E. Vieweg, “群体突发事件中的情境意识:微博交流的行为和语言分析”,博士论文,ATLAS Inst., 科罗拉多大学博尔德分校,博尔德,科罗拉多州,美国,2012 年,第 1-300 页。
- [15] 米 (meter 的缩写) Imran, C. Castillo, P. Meier 和 F. Diaz, “从社交媒体的灾害相关消息中提取信息金块”,正在进行中。Iscram, 2013 年 5 月,第 791-800 页。
- [16] B. Takahashi, E. C. Tandoc, C. Carmichael, “灾难中的推特交流:菲律宾台风海燕期间的推特分析”。哼。贝哈夫。第 50 卷,第 392-398 页,2015 年 9 月。
- [17] B. 许立红,皮罗利和迟浩田,“想被转发吗? Twitter 网络中影响转发因素的大规模分析”,正在进行中。IEEE 第二届国际会议。糖膏剂社会计算。2010 年 8 月,第 177-184 页。
- [18] O. Tsur 和 A. Rappoport, “标签中有什么? 基于内容的微博社区思想传播预测”,正在进行中。第五届国际会议。糖膏剂网络搜索数据挖掘(WSDM),2012 年,第 643-652 页。
- [19] J. Berger 和 K. L. Milkman, “是什么让在线内容像病毒一样传播?” J. 马克。《决议》,第 49 卷,第 2 期,第 192-205 页,2012 年。
- [20] 长度李,张庆,田军,王,“突发事件中信息传播模式的特征:以宜良地震为例”,《国际传播研究》2002 年第一期。J. Inf. 管理。第 38 卷第 1 期,第 34-41 页,2018 年 2 月。
- [21] R. M. Seyfarth 和 D. L. Cheney, “从属关系、移情和心理理论的起源”,Proc. 纳特。阿卡德。Sci. 美国,第 110 卷,第 2 期,第 10349-10356 页,2013 年 6 月。
- [22] C. 舍雷尔和赵,“风险认知的社会网络传染理论”,《风险分析》,第 23 卷,第 2 期,第 261-267 页,2003 年 4 月。
- [23] 名词 (noun 的缩写) Bhuvana 和 I. Arul Aram, “脸书和 Whatsapp 在 2015 年钦奈 (印度) 洪水期间作为灾害管理工具”,Int. J. 减少灾害风险,第 39 卷,2019 年 10 月,第 101135 号。
- [24] 长度 Bui, “波多黎各的社交媒体、谣言和飓风预警系统”,正在进行中。第 52 届夏威夷国际电影节。糖膏剂系统。Sci., 2019 年,第 2667-2676 页。
- [25] 南 Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi 和 M. Tesconi, “出售名声:有效检测虚假 Twitter 追随者”,Decis. 支持系统。第 80 卷,第 56-71 页,2015 年 12 月。
- [26] 米 (meter 的缩写) Hofer 和 V. Aubert, “推特上感知的桥梁和结合社会资本:区分追随者和被追随者”。哼。贝哈夫。第 29 卷,第 6 期,第 2134-2142 页,2013 年 11 月。
- [27] 米 (meter 的缩写) j. 斯特恩, A. E. 亚当斯和 s. 埃尔泽塞尔,“数字不平等和场所:技术扩散对农村、郊区和城市互联网熟练程度和使用的影响”。《调查》,第 79 卷,第 4 期,第 391-417 页,2009 年 11 月。
- [28] 南亚迪和 d. 博伊德,“从城镇广场发微博? :测量地理局域网,《方法》,第 77 卷,第 3 期,第 194-201 页,2010 年。
- [29] A. C. E. S. Lima, L. N. de Castro 和 J. M. Corchado, “推特信息的极性分析框架”,应用数学。计算机。第 270 卷,第 756-767 页,2015 年 11 月。
- [30] Z. Chu, S. Gianvecchio, H. Wang 和 S. Jajodia, “检测 Twitter 帐户的自动化:你是人类、机器人还是电子人?” IEEE Trans. 可靠的安全计算。第 9 卷,第 6 期,第 811-824 页,2012 年 11 月。
- [31] 长度 Atlani-Duault, J. K. Ward, M. Roy, C. Morin 和 A. Wilson, “追踪流行病中的在线英雄行为和指责”,《柳叶刀公共卫生》,第 5 卷,第 3 期,第 e137-e138 页,2020 年 3 月。
- [32] Y. 单词的心理意义:LIWC 和计算机化的文本分析方法。社会心理学。第 29 卷,第 1 期,第 24-54 页,2010 年 3 月。
- [33] C. 钟和彭尼贝克,“语言调查和字数统计 (LIWC)”,应用国家。郎。过程。第 2015 卷,第 206-229 页,2012 年 2 月。
- [34] C. 李,易建联,吕燕,段,“语言动态系统数据驱动设计的混合学习方法”,IEEE/CAA 自动化学报,第 6 卷,第 6 期,第 1487-1498 页,2019 年 11 月。
- [35] Y. 夏,于,王福友,“通过完全卷积网络实现精确鲁棒的眼中心定位”,IEEE/CAA J. Automatica Sinica,第 6 卷,第 5 期,第 1127-1138 页,2019 年 9 月。
- [36] C. 孙等,“基于邻近度的自动驾驶数据标注”,IEEE/CAA J. Automatica Sinica,第 7 卷,第 2 期,第 395-404 页,2020 年 3 月。



李丽芳于 2019 年获得香港城市大学数据科学博士学位。  
她是中国广东风险治理研究中心的成员,也是华南理工大学公共管理学院的助理研究员之一。她的研究兴趣包括危机管理和社会计算。





张庆鹏 (IEEE 成员) 于 2009 年获得中国武汉华中科技大学自动化专业学士学位，并于 2012 年获得美国亚利桑那州图森市亚利桑那大学系统与工业工程专业博士学位。

他是美国纽约州特洛伊市伦斯勒理工学院计算机科学系无线世界星座的博士后研究员。他目前是城市大学数据科学学院的助理教授。他的研究兴趣包括社会计算、复杂网络、数据挖掘和语义

通系统汇刊》、《IEEE 计算社会科学汇刊》和汇刊的副主编。



小王 (IEEE 成员) 于 2011 年获得中国辽宁大连市大连理工大学网络工程学士学位，并于 2016 年获得中国北京中国科学院大学社会计算博士学位。

她现在是中科院自动化研究所复杂系统管理与控制国家重点实验室的副教授。她的研究兴趣包括社会运输，网络运动组织-

社会网络分析。她撰写或合著了十多篇 SCI/EI 文

术书籍 (英汉互译)。

系统汇刊，IEEE/CAA 自动化学报

ACM 智能系统与技术学报，作为同行评审者享有良好声誉。



张军 (IEEE 资深会员) 获得了

2003 年和 2005 年分别获得中国武汉华中科技大学电气工程学士和硕士学位，2008 年获得美国亚利桑那州凤凰城 Arizona 州立大学电气工程博士学位。

他目前是美 国丹佛大学电子和计算机工程副教授。他撰写/合著了 80 多篇同行评议的出版物。他的研究专长是

能、知识自动化及其在智能电力和能源系统中的应

张博士是《IEEE 计算社会系统汇刊》和《自动化学报》的副主编，也是 IEEE RFID 委员会物联网技术委员会的联合主席。



王海目前是中国长沙国防科技大学系统工程学院的助理教授。他的主要兴趣包括社会计算和人力资源分析。



田毕业于美国丹佛大学电气与计算机学院，获硕士学位。

他目前是中国武汉市武汉大学电气与自动化学院的项目工程师。他的主要研究兴趣是人工智能、自然语言处理和区块链在电力系统运行和控制中的应用。



段威于 2014 年获得中国长沙国防科技大学控制科学与工程博士学位。

他的研究兴趣包括复杂网络、流行病建模、信息传播、基于代理的模拟和社会计算。



蔡锦辉 (Kelvin Kam-fai Tsoi) 获得香港中文大学统计学系学士学位，香港中文大学公共卫生学院博士学位，并在香港中文大学内科及药物治疗学系胃肠病学及肝病系接受博士培训。

他还被任命为香港中文大学 JC 肠癌教育中心主任，以推广结肠癌筛查-

中文大学 JC 公共卫生与基层医疗学院、SH 大数据决策与老龄研究所的副教授。



王飞跃 (IEEE 研究员) 于 1990 年获得美国纽约州特洛伊市伦斯勒理工学院的计算机和系统工程博士学位。

他于 1990 年加入美国亚利桑那州图森市的亚利桑那大学，成为一名教授，并担任机器人和自动化实验室以及复杂系统高级研究项目的主任。1999 年，他创建了中国科学院自动化研究所智能控制与系统工程中心。

在外杰出人才计划和中科院百人计划的支持下，中国加入中国科学院复杂系统与智能科学实验室，担任实验室主任。2006 年至 2010 年，他担任自动化研究所研究、教育和学术交流副总裁。2011 年被任命为国家特聘专家，北京复杂系统管理与控制国家重点实验室主任。他目前的研究兴趣包括并行系统、社会计算、并行智能和知识自动化的方法和应用。