

运用社交网络分析淹水空间信息之关联性

周天颖¹, 赖顺仁², 黄崇典², 杨龙士², 叶美伶¹, 吴政庭¹, 方耀民¹

1. 逢甲大学 地理信息系统研究中心, 中国 台湾 40724;

2. 逢甲大学 土木及水利工程博士学位学程, 中国 台湾 40724

摘要: 随着社交网络的普遍发展, 大量的讯息透过智能手机发布在个人的微博或其他社交网站。台湾地区的社交网站以脸书(Facebook)的使用量最大, 平均每天有近千万笔的讯息量, 大多数的讯息多以衣食住行或个人讯息为主, 但从本研究所撷取自2010年至2015年的数据中显示, 公众在社交网站所分享的信息中具有降雨、淹水或相关灾情的讯息, 而这些讯息具有极高比例的正确性。由于社交网站无法提供私人讯息, 故本研究将从社交信息中, 以地点为单位撷取大量的数据信息再辅以语意关键词萃取出有关可作为淹水预判的讯息数据。为检核资料的可信度, 本研究透过历史台风数据FLO-2D仿真重建淹水之空间信息进行检核。从研究比对分析中发现, 经萃取后的公众信息其与灾害的关联性正确性相当显著, 故透过社交网站中大量的非结构讯息, 透过语意及空间的转换, 可萃取转化为防灾信息, 对广域的都市治理而言, 此一讯息将可作为预判区域淹水或防救灾情报之有效参考。

关键词: 水灾, 语意探勘, 社交淹水测量, 志愿者空间信息

中图分类号: TP79 **文献标志码:** A

引用格式: 周天颖, 赖顺仁, 黄崇典, 杨龙士, 叶美伶, 吴政庭, 方耀民. 2016. 运用社交网络分析淹水空间信息之关联性. 遥感学报, 20(5): 1299-1307

Chou T Y, Lai S R, Huang C D, Yang L S, Ye M L, Wu Z T and Fang Y M. 2016. Social crowd sourcing application in spatial information analysis. Journal of Remote Sensing, 20(5): 1299-1307 [DOI:10.11834/jrs.20166204]

1 引言

台湾地区地处低洼地区的县市, 每逢台风或豪雨侵袭时, 常容易发生淹水灾害, 造成重大的人民财产甚至生命损失。经统计, 自2012年以来共发生350次台风、上千次豪雨侵袭, 平均每年发生3.5次台风与多次豪雨, 是台湾地区最严重的天然灾害。台风豪雨灾损及设施投入每年已超过300亿, 且灾害损失有逐年加大的趋势。在换算成金钱的灾损上, 随着人口的增加, 经济的成长, 台湾地区历年来的洪灾损失有逐年上升的趋势, 而且上升得非常迅速。而现行地方政府对于淹水灾害的处理方式, 多藉由民众主动通报(以电话或APP等方式)获知灾情, 针对少数易淹水区域则透过监测系统预警。然而, 当灾情发生时, 若政府能藉由社交网络等平台, 分析民众提供的信

息, 主动搜寻出淹水信息并能预判出灾情, 将可加速政府对于灾害的应变反应时间, 并对于有助于救灾资源的分派, 降低灾害的冲击。近年来在几次的急降雨下台湾地区的市亦都发生淹水事件, 在2012年—2014年包含苏拉台风、潭美台风与苏力台风3场的多次强降雨下, 台中沙鹿地区均成为淹水的灾区, 政府近年亦投入经费进行淹水的整治, 并建置无线淹水感测设备实时传递灾害信息之措施, 然而要建立完整的监测网恐要耗费庞大的资源, 实务上也不太可能, 因此如何运用创新的工具加强都市水患的调适能力成为重要的课题。

Surowiecki(2005)提出公众智慧的概念, 其中心论点即多样化的收集个人自主的决定, 可能使某些类型的决策和预测与统计抽样, 甚至比专家的更好, 其信息的特质可以更贴近真实。Goodchild

收稿日期: 2016-05-30; 修订日期: 2016-06-02; 优先数字出版日期: 2016-07-09

第一作者简介: 周天颖(1959—), 男, 特聘教授, 逢甲大学地理信息系统研究中心主任, 研究方向为地理资讯系统、资源保育、遥感探测、土地管理、防灾资讯。E-mail:jimmy@gis.tw

(2007)提出“Citizen as sensors”的概念,其概念是指当所有使用者所能提供互联网服务进行信息交流时,透过手持装置GPS获取坐标使用于开放空间信息平台上应用以及透过每一位用户的感官信息如视觉、听觉等,并结合居民的地理知识信息做为参考并记录于平台使用,使平台信息能相互交流,每一位使用者都能透过此平台达到建立、更新、管理、维护目的的地理信息。目前热门社交网站如微博、Facebook、Twitter、Flickr或Plurk等系统中,公众可藉由智能型装置进行拍照、签到(Check in)并上传到社交分享的动作,这些分享的事物均带有坐标点位,或是有意无意地透露所在地点的地理讯息,均是公众信息提供的表现形式。

在台湾地区近年来的灾害,如莫拉克风灾或高雄气爆时,均有民众主动将灾害的情况上传到平台上,使灾情与搜救情况可以快速分享。社交信息个别公众之集合信息符合此特性,加上网络及定位行动装置的迅速普及,这些技术的发展提供了公众智能议题的发酵。依此念之应用,例如,联合国开放空间信息协会OGC(The Open Geospatial Consortium)组织,2015以美国旧金山为研究案例,透过开发Sensor Observation Service (SOS)系统,搜集Twitter、Flickr、Instagram、Facebook等社交平台上的民众噪音签到的应用;2013年欧洲学者Laura等人建置GeoCONAVI平台(Geographic CONTEXT Analysis for Volunteered Information),搜集与分析森林火灾或相关的灾害信息应用。Degrossi等人(2014)亦提出以市民感测众包方式应用于巴西的淹水风险管理,此外针对澳洲的布里斯班,提出志愿者地理淹水信息可信度的评估方法(Hung等,2016),讨论志愿者信息可信度的课题,而在2015年由林峰正等人提出应用公众智能提供灾情情报暨信息过滤机制—以水灾为例,探讨如何将社交信息过滤成有用信息。Rodolfo(2012)提出了分析Twitter语料的语意以实时判断用户地理位置的方法,实验由Amazon Mechanical Turk上的93位网络用户判定结果的正确性,作为公众智能之信息来源。故本研究将非任务型不特定地区内社交信息语意进行分析过滤,结合FLO-2D之模拟进行空间关联分析,探讨以一般社交信息作为淹水讯息来源的可能性。

2 FLO-2D模式应用

2维洪水灾害模拟FLO-2D模式普遍应用在都市淹水、洪水平原管理、工程风险设计、不规则形状河道、桥梁与涵洞水理演算,处理漫地流、都市水文、泥石流(O'Brien, 2006)。此模式为台湾地区常用淹水分析模式之一(彭思显等,2008),能够仿真流动速度与深度,以推估合理的淹水面积,对泥石流及淹水灾害有相当多的应用,如应用于敏督利台风造成南湖溪淹水之模拟(汤嘉芸等,2007),模拟羌黄坑集水区内泥石流流动特性之研究(詹钱登等,2015),及FLO-2D应用于中坜都会区的淹水分析(何宇申,2013)等。

就水文数据的给定,FLO-2D模式可以给定流量历线及降雨数据。使用者可在入、出流点给定入、出流量历线。根据需要可以模拟一场暴雨,亦可输入暴雨之雨量相关数据即可仿真(汤嘉芸等,2007),本研究采用此FLO-2D之特性,就历史淹水事件之针对不同降雨水强度进行模拟,以重现淹水空间信息。

3 志愿性地理信息趋势与分析

地理信息的发展已从过去的信息贫乏时代,走向巨量信息时代(Miller和Han, 2009)。Intel公司于2013曾做过一个统计,在一分钟网络发生了什么事,以下节录一些记录:每分钟,Google执行200万次搜索;Facebook增加350 GB数据,诞生180万个“赞”;YouTube增加72 h影片;新增域名注册70个;Snapchat分享照片10.4万张;Twitter发出“Tweets”27.8万条,预估到2017年资料量还将成长3倍,由此显示出现代人的生活已经离不开网络,其产生的数据量也是巨大的。这样巨量的信息及其搜集方式,已颠覆了传统的运作方式,其来源非由一己之力,因此可藉由大众的力量达成。而另一方面,在网络的力量下亦形成众包(crowdsourcing),全新的组织劳动力方式及志愿性地理信息VGI(Volunteered Geographic Information)由公众的使用来完成空间信息的建置、维护及应用现象的趋势(Howe, 2006)。由于社交透过智能型装置用户将各式各样的信息上传至网络的趋势,在社交的需求下,公众有更多的动力不断地更新传递的特性,使各项信息更新的速度比以往还要快,更贴近真实。因此,如能运用公众的行动装置之集体智能来完成任务,透过公众的感知来提供气

候调适,进而提供正确的防灾信息,可弥补数理模型及感测系统信息的不足,掌握灾情迅速达成防救灾及灾后复原的任务。但社交信息是非结构的、巨量的,如何有效地将非结构文字描述结构化、空间化,探讨其空间分布关联性及其可应用性,成为值得探讨的重要议题。

空间分析的主要任务是对空间构成进行描述和分析,以达到获取、描述和认知空间资料,理解和解释地理图案的背景过程,空间过程的模拟和预测,调控地理空间上发生的事件等目的。由于学者对空间分析的内容定义相差较大,所以有很多称谓,如地理信息分析或地理空间分析、地理信息统计分析、地理分析、空间资料分析、空间统计学、空间统计与建模、地统计学等,但大部分的文献中概称为空间分析(赵永和王岩松,2011)。迭合分析是地理信息中的一项非常重要的空间分析功能。多层资料的迭置分析,可以产生新的空间及属性特征关系,能够发现多层空间数据间的相互差异、关联和变化等特征。陶翼煌和孙志鸿等人(2007)认为就灾害应变信息而言,必须认清它具有强烈的空间特性。虽然不同的应用领域依据关注重点给出不同的定义,但都是从不同面向领域对空间分析的内涵进行阐释,其最终目的是解决人们所涉及的地理空间的实际问题,获取和传递地理空间信息,以支持决策,特别是空间关系隐含的信息。

从文献回顾中,尚未发现结合FLO-2D重现淹水情境并与社交语意之信息进行空间迭合分析,就本研究淹水有关之防灾信息而言,将以交叉的空间信息迭合分析进行空间及其属性关系的比较,以期能获得容易忽略隐含的关联信息。

4 研究范围与研究方法

本研究以台中市沙鹿地区为研究范围(图1),分为3阶段探讨社交空间信息与淹水模拟空间信息的关联程度,第1阶段进行社交的感测信息空间化,第2阶段进行FLO-2D模拟重建淹水空间信息,第3阶段将第1、2阶段之成果进行空间迭合分析,分析其空间关联性。其流程如图2所示。

5 搜集社交信息

目前热门网络社交,如微博、Facebook、

Twitter、Flickr或Plurk等相关社交中,公众可藉由智能型装置进行拍照、签到并上传到社交平台与成员进行分享动作,这些分享的事物均带有坐标点位。公众参与者提供的资料可以被筛选整合起来成为重要的淹水情报。Facebook在2013年该单位发布公告中提到,每天活跃使用者高达1000万人,以Facebook 2300万用户,使用比率相当之高;另从台湾地区资策会的产业情报研究所(“网络社交使用现况分析报告”2014),在社交网站的使用程度分析,提到在台湾地区以Facebook使用量最高,因此本研究采用台湾地区使用较广泛的热门社交网站Facebook的开放社交数据,做为数据分析来源。本研究方法将从历史淹水事件中分析Facebook网络社交信息与其信息的空间分布。

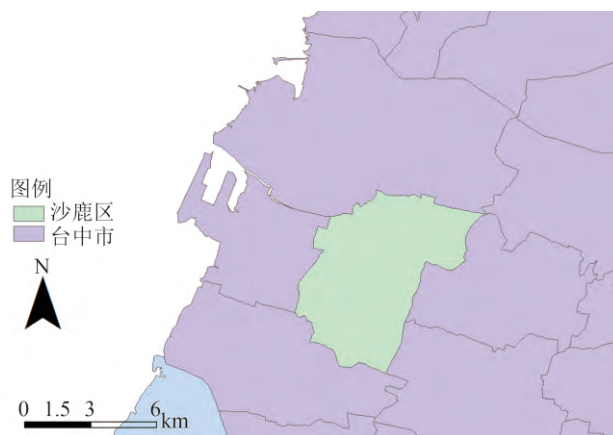


图1 研究范围-台中市沙鹿区

Fig. 1 The study area - Taichung City Shalu District

将搜集以人为感测基础的Facebook社区透过行动装置在研究区内的签到信息,对分析得出“淹水”、“水灾”、“积水”、“大水”...等相关字词,获取该发布者的空间点位与属性数据。另一方面以气象水文及地文数据透过FLO-2D软件的仿真,得出以物理环境之降雨淹水空间信息,最后将社交空间信息FLO-2D进行空间关联分析,得出以人为感测之网络社交在淹水信息利用的可能性。

5.1 信息过滤与分析

水患信息目前过滤的机制主要有2种,一种是地理位置的过滤,即抓出正确的点位数据,另一种是叙述文字中有水灾相关信息过滤。

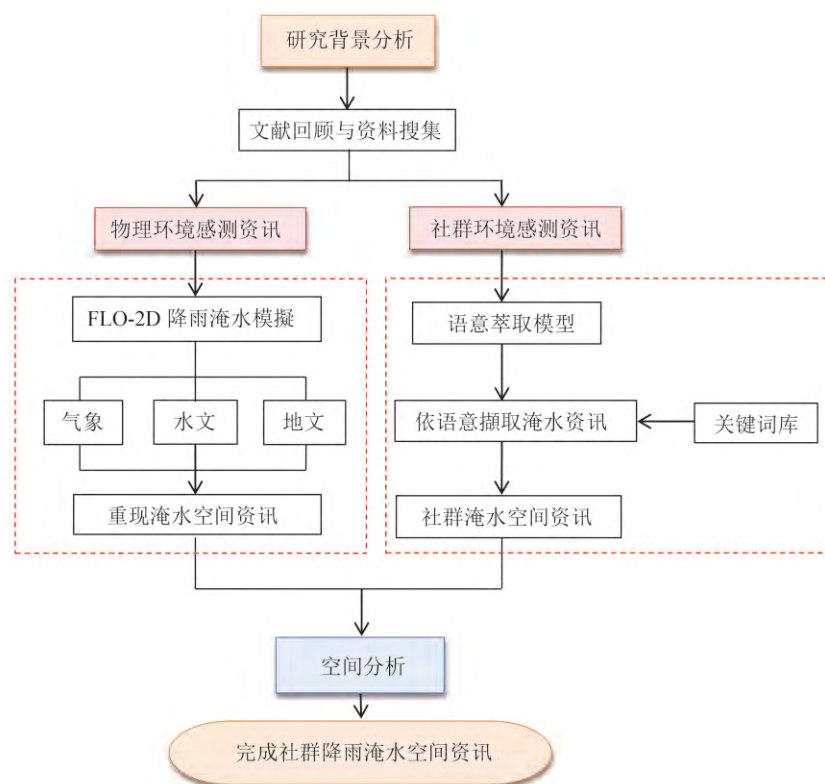


图2 研究流程架构

Fig. 2 Study process architecture

在台中市沙鹿地区,数据搜集过滤流程共有3个阶段(图3): (1)撷取社交网站资料:利用Facebook提供的API(Application Programming Interface),配合沙鹿地区的强降雨历史资料,针对2011-11-23到2016-01-12所取得研究区内之所有签到位置(Place ID)并读取留言内容; (2)读取签到ID之内容:藉由Visual Studio软件撰写网络数据撷取程序,读取HTML卷标内的留言,针对留言时

间、文字内容、坐标点位等数据汇入数据库中; (3)数据过滤:以关键词筛选内容,同时交互检视完成过滤撷取之关键词数据库,关键词的建置内容包括降雨、淹水及表达程度上的用语3个方面,经过过滤、分析比对文字之语意信息,过滤后的信息可以得到包含淹水语意、签到之点位之社交淹水空间信息,其中数据的过滤分析机制是决定数据可用性之关键。

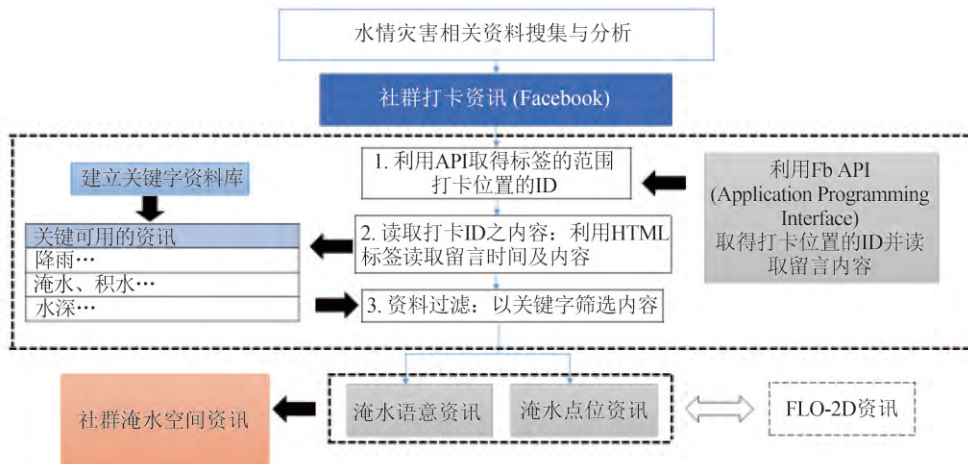


图3 语意分析流程

Fig. 3 Semantic analysis process

语意的过滤，通过语意分析技术来分析非结构化的文本数据。先从网络上寻找，结合MIT stopwords反复的过滤和验证得出关键词后，再从搭配普林斯顿中文网汇网络(Chinese WordNet)找出它们的同义词、相似词字等，并检视内容依地域及社交之用语修编增加相关具地方特性之用语关键词，产生出更大的能判断淹水的相关词库，之后采用关键词库以Apache Solr程序执行语料库的搜寻。在语意分析的流程中，先对数据进行断词来取得完整正确的语意，即断词分析上的歧义性及未知词的判断。使用Lexicon-based以关键词进行断词，并透过自然语言处理模型N-gram language model来分析社交语意信息，针对网站社交数据产生断词时使用的关键词库与进行断词，并筛选出能够代表“淹水事件”的关键词，进行过滤。其关键词如表1。

表 1 语意的过滤分析关键词
Table 1 Keywords for semantic filtering and analysis

关键词1	关键词2	关键词3
水灾	涨	泡水
泛滥	淹	水深
山洪	雨	豪雨
防洪	台	冲走
洪水	水头	漂流
洪灾	漫溢	积水
洪流	蓄洪	沙包
滚滚洪流	暴洪	河流
凌汛	落雨	喷水
洪峰	大水	淹到
溢	泥流	

5.2 FLO-2D淹水模式说明与设定

- (1)FLO-2D模式之假设，包括：浅水波假设；满足稳定流阻滞方程式(steady flow resistance equation)；静水压力分布；差分时间间隔内为稳定流(steady flow)；分析网格点断面与水力粗糙度为均值；每一个网格点的高程与曼宁值仅有单一值。
- (2)FLO-2D模式之限制条件，包括：假设为定床模式，而由于定床模式之限制，故本模式于非输砂模块无法模仿真实冲刷度深度之现象；无法模拟震波(shock wave)与水跃(hydraulic jumps)现象。
- (3)FLO-2D模式之参数条件，包括：区排初始

曼宁 n 值：采用0.014—0.015(参考区域排水整治及环境营造规划参考手册)；地表初始曼宁 n 值：采用FLO-2D模式内定值0.04，验算过程会自动会调整；地形高程：采用2006年5 m×5 m的DEM数据；雨型站：采用沙鹿地区横山雨量站资料；计算网格：采用40 m×40 m；计算时间间隔：共模拟24 h，并以每小时输出；其他：采用FLO-2D模式内定值。

本研究选取2013年苏拉台风实际降雨记录，利用FLO-2D模式进行降雨区间雨量等比例正推与反推降雨值后取整数值进行19场次之降雨模拟，以提供不同降雨情境下与社交签到空间信息；降雨正推最大一日累积雨量为1638 mm，其次为1556 mm，次之为1474 mm递减；降雨反推最小一日累积雨量为81.9 mm、163 mm与245 mm递增。当最大一日累积雨量为1638 mm，降雨量区间引起局部淹水之降雨强度约为43.7—53.6 mm/h之间；最小一日累积雨量为81.9 mm，降雨量区间引起局部淹水之降雨强度约为6.7—5.5 mm/h之间。

由于该淹水点位淹水灾情，主要为较强之瞬间降雨，所造成历史淹水点位之最高之淹积水现象。因此，由各场降雨量选取最大值之降雨强度为代表。

$$R_{\max}=\max(R_1, R_2, \dots, R_n)$$

式中， $R_{\max}$ 为最大降雨量(mm)， $n=1, 2, 3, \dots$ 。

利用此19场仿真数据所代表之降雨仿真，提供社交降雨淹水信息空间分析所需之数据。

6 结果与讨论

6.1 社交签到信息与历史灾情结果比对

本研究以3场历史台风事件中，确定研究区之具空间点位之社交文字信息，先以关键词初步萃取出与淹水语意有关之描述。Facebook的数据无法直接从后台取得个人的讯息，因此须从地点Place ID中撷取所有的签到公开信息，本研究以较大范围的中部地区进行资料的搜集，共下载了104964条信息，再筛选出位于研究地区内的信息共有10128条，透过接续从信息内容中撷取具有关键词的信息共有237条，然而即使透过关键词仍有许多的信息是错误、重复或无关于水灾的信息，故最后透过人工作最后挑选，取得110则与水灾相关的信息，操作流程如图4所示。

脸书文字讯息搜寻:台湾中部地区

Place ID	Fb留言笔数
3700点空间位置	10万4964则讯息

研究范围: 沙鹿地区

Place ID	Fb留言笔数
189点空间位置	10128则讯息

语意过滤: 关键词

Place ID	Fb留言笔数
70点空间位置	237则讯息

Place ID	Fb留言笔数
49点空间位置	110则讯息

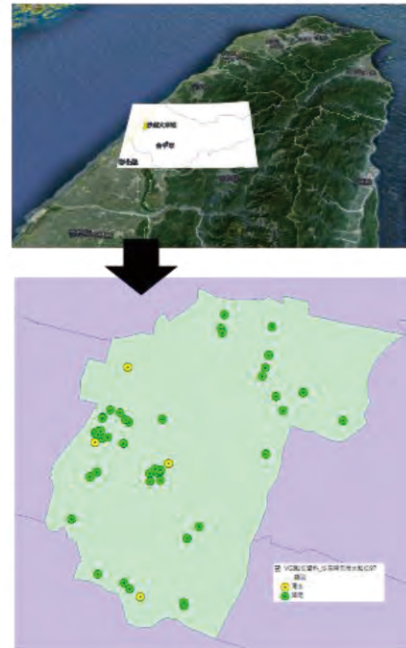


图4 公众签到信息筛选流程与空间点位分布

Fig. 4 The process of filter and VGI points location

为了解这110则信息是否可作为水灾预警之用,本研究以3个台风的时间点(表2)过滤与淹水相关之讯息,过滤后与淹水语意有关之部份内容如表3、表4所示。

表2 沙鹿地区台风数据表

Table 2 Shalu area typhoon data sheet

中文名称	英文名称	强度	时间
苏拉	SAOLA	中度	2012-07-30 20:30
			2012-08-03 14:30
潭美	TRAMI	轻度	2013-08-20 11:30
			2013-08-22 08:30
康芮	KONG-REY	轻度	2013-08-27 11:30
			2013-08-29 20:30

注:数据来自中国台湾气象部门。

总计在这3个时间点内共有10则讯息,其中9则明确叙述的降雨事件,其中8则明确陈述了淹水事件,其中3则明确叙述灾害发生。进而比对沙鹿地区在这3个台风中,从官方所建置的实际淹水监测数据中比对,其该地点均有发生淹水情形,故可显示民众提供的信息其正确率达80%,且其中有37.5%的比例可得出该信息地点的淹水程度。

表3 台风期间内公众签到讯息总表

Table 3 The contents of VGI during the three typhoon

台风	编号	信息内容	坐标(°E, °N)	讯息日期
苏拉	A1	落大雨	(120.608, 24.246)	2012-08-02
	A2	狂风暴雨不停~都淹水啦	(120.574, 24.229)	2012-08-02
潭美	B1	水淹半个轮胎了!赶紧把车开走...隔壁台车要泡水了	(120.559, 24.237)	2013-08-22
	B2	该改成"台湾地区大河"了	(120.580, 24.214)	2013-08-22
	B3	出去送货比回到公司还安全。这边水是用倒的。	(120.553, 24.218)	2013-08-22
	B4	下交流道就看到两边机车道已完全淹水,三个快车道也淹掉半个,沙鹿市区更是全淹了	(120.562, 24.233)	2013-08-22
康芮	C1	现在下班经过沙鹿天桥往下看满路黄泥,真的很恐怖,有关单位真的要来关心一下,阿弥陀佛!天佑台中沙鹿区	(120.560, 24.240)	2013-08-29
	C2	夸张房子都淹2/3	(120.560, 24.240)	2013-08-29
	C3	沙鹿天桥的地下道淹成这副德性!!!	(120.560, 24.240)	2013-08-29
	C4	沙鹿市场淹大水,恐怖的水量10分钟内淹到40公分深	(120.558, 24.237)	2013-08-29

6.2 社交签到信息与模拟淹水范围结果比对

进一步将签到信息与由FLO-2D的淹水模拟进行比对,首先淹水模拟范围图可得到不同降雨强

度的淹水范围，每小时强度分别为20 mm/h、26 mm/h、33 mm/h、...等共19个代表的降雨强度，以推估出各种不同降雨强度之情境，如图5中浅绿色部分表示模拟淹水范围随着雨量扩大之结果。藉由模拟结果可以得到每一个不同的降雨强度的淹水空间信息图层，可以提供社交信息对应的降雨情境，并可同时得到任一签到点位淹水范围及深度值。

表 4 公众讯息对灾害的表达程度分析表
Table 4 The table of analyst the VGI for disaster

编号	累积降雨量/mm	是否有降雨发生	是否有淹水(灾情)	是否说明灾情程度
A1	449	有	无	无
A2	449	有	有	无
B1	211	有	有	有
B2	126	有	有	无
B3	126	有	无	无
B4	126	有	有	无
C1	151	无	有	无
C2	151	有	有	有
C3	151	有	有	无
C4	151	有	有	有

6.3 空间分析结果

从气象资料时间范围内查得以3次台风中最大3次之时雨量为2012-08-02 12:00强度为88 mm/h，其次为2012-08-02 19:00为86 mm/h和2012-08-02 18:00 79 mm/h，前3名均为苏拉台风的降雨强度(表5)。本研究将FLO-2D模拟信息与不同的降雨强度进行迭合图层分析，降雨强度较小之迭合分析，空间重

迭的社交点位相对的少，而以88 mm/h最大时雨量之相近之87 mm/h为代表降雨强度所得到的FLO-2D模拟淹水模拟范围图层，并与经由语意过滤所得之社交签到空间信息进行图层迭合，迭合结果如图6。

研究发现，由空间迭合分析，显示与淹水相关之关键词所得到的空间信息与重现87 mm/h最大时雨量FLO-2D淹水模拟信息之空间关联度相当高。研究选取研究期间之最大降雨强度87 mm/h之FLO-2D淹水空间信息与同时期社交有关淹水签到信息，以空间迭合法进行套迭。迭合结果显示49个签到点位信息中有38点位之空间信息完全吻合，与其空间之吻合程度高达78%，其中针对上述10个点位不在模拟结果之范围内，但仍显示紧邻仿真范围之边缘，如图6中白色圆圈所圈示。10个点位中与拟模淹水范围边缘最大之直线距离为270 m，如果定义300 m设定为小区生活范围，签到者用以反映小区附近之淹水情形之观点，则显示所有空间数据完全吻合，空间信息呈现完全相关。而相对的，其他区域没有发现任何与淹水语意有关之签到信息。

此外，针对全部的讯息数据，分析具有“降雨”有关描述的110个签到留言所呈现的49个空间点位数据中，签到数据中几乎与FLO-2D直接与“淹水”有关之空间信息模拟，空间吻合情形相当显著。由于资料撷取研究时间长达3年，在历经多次强降雨事件，多次验证下，均有一致性的结果，显示在研究区内以社交签到信息提供淹水有关的信息具一定程度的可靠性。

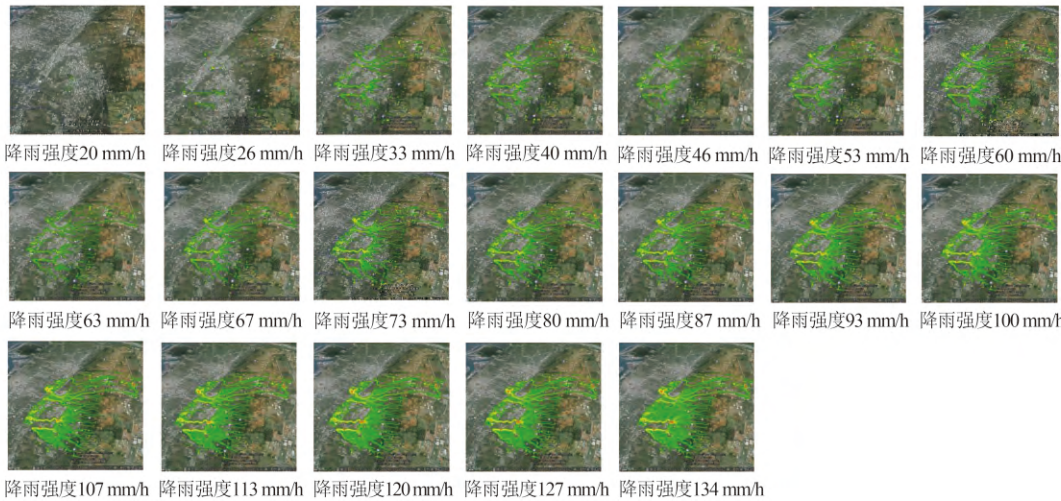


图 5 FLO-2D不同降雨强度淹水空间信息模拟

Fig. 5 FLO-2D simulation of flood spatial information under different rain intensity

表5 沙鹿地区3次台风每小时最大强降雨强度排序

Table 5 Order of per-hour rainfall in three typhoons in Shalu area

时间	降雨量/(mm/h)	台风名称
2012-08-02 12:00	88	苏拉
2012-08-02 19:00	86	苏拉
2012-08-02 18:00	79	苏拉
2012-08-02 11:00	65.5	苏拉

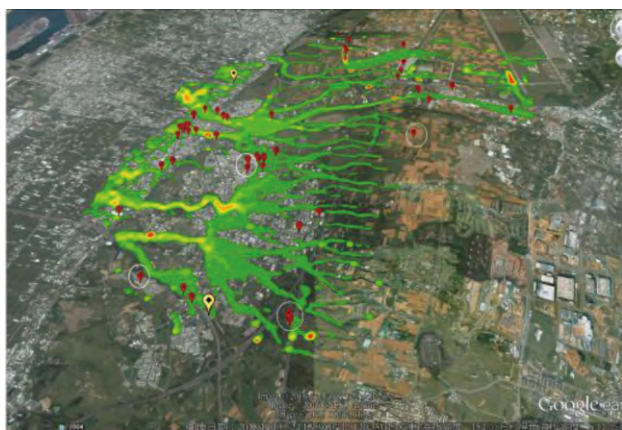


图6 降雨87 mm/h FLO-2D模拟与社交空间淹水信息迭合

Fig. 6 Under 87 mm/h rainfall the overlapping of FLO-2D simulation and community's flood spatial information

7 结 语

本研究发现, 公众提供的信息包含降雨、淹水等数据具有极高的正确性, 其签到的空间位置套选FLO-2D淹水模拟范围, 其空间关联性分析结果相当显着。由于本次研究所取得的数据信息是以开放、不是特定任务、或特定目的的一般社交, 虽数据多样化且不具规则性, 但所呈现之社交信息仍具有高度的可靠性。此外, 本研究发现空间分析所隐含的信息, 社交描述之“降雨”信息与历史的淹水事件之空间信息有显着的关联性, 印证了公众所提供的讯息可作为预警效益的观点, 研究结果显示, 非任务型不特定的社交信息是可以弥补感测设备之不足, 其进一步提供淹水信息是可行的。其他地区可以依此念框架探测降雨情境, 由Facebook社交签到信息探测可能的淹水范围。

然而, 由于FLO-2D模拟是以 40×40 均质化之网格, 加上19个模拟之降雨强度均以24 h内之最大降雨量为代表强度, 因此研究仍然有系统性的误差。相对的, 以人为感测之淹水信息虽有语意的不确定因素, 但其信息是有可能更贴近真实的,

如能精进语意之分析, 提高社交信息之准确性与可信度, 将网络社交之信息撷取过滤分析成可靠之防灾信息, 并可应用至未布设感测系统之区域, 可以为都市的水患问题提供相关决策之参考。

语意的萃取滤分析是数据正确的关键, 关键词库的建立过程, 发现社交特有的表达方式, 例如“大水”是以闽南发音用国语呈现之特有表达方式, 或有以注音符号取代或发音相近的同音字表述, 都值得未来在社交语意萃取语方法上进一步研究。

此外, 在大台中区非研究区的社交资料中, 发现民众更进一步述淹水深度的有少数几笔, 如直接以量尺描述淹水深度(40 cm), 或以对象为尺度描述淹水深度(半个轮胎、车子灭顶), 或以人为尺度描述淹水深度(水深及膝)等, 但均因语意表达的限制难以精准。未来如能研究精进语意之分析, 或在社交网络平台提供符号化之量尺, 以更迅速有效的表达灾害信息, 应该可以更充分发挥社交网络公众智能的优势, 对灾害之因应做出更多的贡献。

参考文献(References)

- Degrossi L C, de Albuquerque J P, Fava M C and Mendiondo E M. 2014. Flood Citizen Observatory: a crowdsourcing-based approach for flood risk management in Brazil//Proceedings of the 26th International Conference on Software Engineering and Knowledge Engineering. São Carlos, Brazil: SEKE: 570–575
- Goodchild M F. 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4): 211–221 [DOI: 10.1007/s10708-007-9111-y]
- Ho Y S. 2013. Application of FLO-2D on the Flooding Analysis in Chungli Metropolitan Area. Taipei, China: Taiwan University (何宇申. 2013. FLO-2D应用于中坜都会区的淹水分析. 台湾地区, 中国: 台湾大学)
- Howe J. 2006. The rise of crowdsourcing. *Wired Magazine*, 14: 6
- Hung K C, Kalantari M and Rajabifard A. 2016. Methods for assessing the credibility of volunteered geographic information in flood response: a case study in Brisbane, Australia. *Applied Geography*, 68: 37–47 [DOI: 10.1016/j.apgeog.2016.01.005]
- Jan Q D, Hsiao K W, Hsu Y C and Tseng K H. 2015. Applying FLO-2D and debris 2D model to simulate characteristics of debris flow in Qianhuangkeng watershed. *Journal of the Taiwan Disaster Prevention Society*, 7(2): 239–247 (詹钱登, 萧凯文, 徐郁超, 曾国训. 2015. 应用FLO-2D及Debris 2D模拟羌坑集水区土石流流动特性差异之研究. *中华防灾学刊*, 7(2): 239–247)

- Miller H and Han J W. 2009. Geographic Data Mining and Knowledge Discovery. 2nd ed. New York: CRC Press
- O'Brien J S. 2006. FLO-2D User's Manual, Version 2006. 01. Nutrioso: FLO-2D Software, Inc
- Peng S H, Tsai C and Zeng Z H. 2008. Study on flood disaster and scenario simulation in Nantou County. Journal of Chienkuo Technology University, 27(2): 1-16 (彭思显, 蔡志崇, 曾智弘. 2008. 南投县洪水灾害规模设定及其境况仿真研究. 建国科技大学报, 27(2): 1-16)
- Rodolfo C. 2012. Tweelocator: a non-intrusive geographical locator system for twitter. Beijing: National Tsinghua University Institute of Information Systems and Applications
- Surowiecki J. 2005. The Wisdom of Crowds. New York: Anchor Books
- Tang C Y, Hsieh P C and Lin L L. 2007. Application of FLO-2D and HEC-GeoRAS to the simulation of Na-Hu creek flooding by Typhoon Mindulle. Journal of Chinese Soil and Water Conservation, 39(1): 87-96 (汤嘉芸, 谢平城, 林俐玲. 2007. FLO-2D与HEC-GeoRAS应用于敏督利台风造成南湖溪淹水之模拟. 水土保持学报, 39(1): 87-96)
- Tao Y H and Sun C H. 2007. A methodology for spatial information requirement analysis—an approach of disaster research in geography. Geographic Information System, 1(2): 19-25 (陶翼煌, 孙志鸿. 2007. 空间信息需求分析方法之建构—地理学中灾害研究之新主张. 地理信息系统季刊, 1(2): 19-25)
- Zhao Y and Wang Y S. 2011. Progress of spatial analysis. Geography and Geo-Information Science, 27(5): 1-8 (赵永, 王岩松. 2011. 空间分析研究进展. 地理与地理信息科学, 27(5): 1-8)

Social crowd sourcing application in spatial information analysis

CHOU Tianying¹, LAI Shunren², HUANG Chungtien², YANG Longshi², YEH Meiling¹, WU Chengting¹,
FANG Yaoming¹

1. GIS Research Center, Feng Chia University, Taiwan 40724, China;

2. Ph.D. Program in Civil and Hydraulic Engineering, Feng Chia University, Taiwan 40724, China

Abstract: Short-term heavy rainfalls have become the norm as the climate changes. However, previous city drainage design is often unable to cope with this rainfall pattern. Thus, the damage caused by urban flood disasters is high, especially because of relatively high population densities. Several sudden rains in Taiwan have caused flooding in the cities in recent years. Taichung's Shalu area experienced flooding because of the heavy rains brought by typhoons Sura, Tam Mei, and Su Li from 2012 to 2014. The Taichung City Government has made flood remediation investments and has built wireless sensing devices to convey real-time flooding information. However, a considerable amount of resources are required to establish a complete monitoring network, and such requirement is unachievable in practice. Therefore, how to use innovative tools to enhance cities' flood adaptability has become an important issue. As a result of the popularity of smart phones, an increasing number of people publish personal information on popular Internet communities, such as Facebook, Flickr, Twitter, and Plurk. Using their smart devices, people can take pictures and post information to share with community members. The shared information is tagged with coordinate points. Information from a large group of people can thus be screened and integrated for use as valuable flood information.

This study investigates the online descriptions of direct experiences during flooding events to obtain the spatial information of floods through semantic retrieval and filtering analysis and thereby identify flood patterns. Sensed disaster data, credible information extracted from Internet communities, and the use of such information as city flooding information can effectively support the assumptions and limitations of physical sensing mathematical models, particularly in terms of the degree of effective operation and adequate maintenance of urban drainage systems and grid mode homogenization. This study, which on urban flooding events in history, extracts real-time flood-related information from online communities, such as Facebook, and compares the actual values of the flood for spatial analysis. Moreover, this study filters VGI on the basis of semantic meaning and obtains 49 "rain"-related descriptions. The descriptions are matched to the flood spatial information reconstructed with FLO-2D simulation, with the result indicating that the distribution correlation is significant. The method of converting human-sensed non-structured information from Internet communities into usable spatial information, extracting usable information available in online communities, converting this information into disaster prevention information, and using physical-sensing FLO-2D simulation to reconstruct flood spatial information for a correlation analysis is unique and innovative. Results show that a non-task-based, non-specific community can compensate for the insufficiency of detective equipment and further provides flood information. Rainfall situations in other areas can be detected with this method and be framed using Facebook's community check-in information to detect possible flooding ranges. Using flood information from online communities to provide initial flood information and to govern cities with broad areas is a feasible method.

Key words: flood, semantic mining, social flooded measurement, volunteers space information