

多源地理大数据支撑下的地理事件建模与分析

杜云艳^{1,2}, 易嘉伟^{1,2}, 薛存金^{2,3}, 千家乐^{1,2}, 裴 韬^{1,2}

(1. 中国科学院地理科学与资源研究所, 北京 100101; 2. 中国科学院大学, 北京 100049;
3. 中国科学院空天信息创新研究院, 北京 100101)

摘要: 地理事件作为描述地理过程的基本单元, 逐渐成为地理信息科学(GIS)核心研究内容。由于受人类活动数据获取限制, GIS对地理事件的建模和分析主要关注事件所引起的地理空间要素变化及要素之间的相互影响与作用机制。然而, 近年来随着基于位置服务数据(LBS)爆炸式的增长和人类活动大数据定量刻画手段的快速发展, 地理事件对人类活动的影响以及公众对地理事件的网络参与度都引起了多个领域的广泛关注, 对地理事件的时空认知、建模方法和分析框架提出了巨大的挑战。对此, 本文首先深入分析了大数据时代地理事件的概念与分类体系; 其次, 基于地理事件的时空语义给出了基于图模型的事件数据建模, 建立了事件本体及其次生或级联事件的“节点一边”表达结构, 开展了事件自身时空演化及其前“因”后“果”的形式化描述; 第三, 从时空数据分析与挖掘的角度, 给出了大数据时代地理事件建模与分析的整体框架, 拟突破传统“地理实体空间”事件探测与分析方法的局限性, 融合“虚拟空间”事件发现与传播模拟思路, 实现多源地理大数据支撑下的面向地理事件的人类活动多尺度时空响应与区域差异分析; 最后, 本文以城市暴雨事件为例诠释了本文所提出的地理事件建模与分析方法, 从城市和城市内部两个尺度进行了暴雨事件与人类活动的一致性响应及区域差异分析, 得到了明确的结论, 验证了前文分析框架的可行性与实用性。

关键词: 地理事件; 图数据模型; 地理大数据; 人类活动响应

DOI: 10.11821/dlxb202111018

1 引言

百科全书中事件被定义为比较重大且对一定人群产生一定影响的事情, 包含人类有目的的行为及其与物体间的相互作用, 具有复杂的动态多变性^[1]。从认知科学角度看, 事件发展过程也是人类认知世界的过程, 作为知识的一种表示形式, “事件”被认为是人类认识世界的基础单元^[2]。由此不难推断, 发生在客观世界中具有一定时空范围的地理事件是人类活动与地球表层相互作用且产生重大影响的事情, 是地理信息科学(GIS)机理认知的基本单元, 因此, 对其深入研究是了解人地系统各要素间相互作用机理与动态演化趋势的重要突破口。

然而, 由于早期GIS系统受数据获取手段的限制, 无法定量刻画人的特征及其活动, 因此, 以往对地理事件的研究重点关注事件引起的客观世界变化的表达及其影响的评估与预测。具体包含两方面: 其一是从时空数据模型角度, 把事件作为引起客观世界变化的原因或结果提出的时空演化模型研究; 其二是地理事件的过程分析与模拟。在时

收稿日期: 2020-07-20; 修订日期: 2021-04-19

基金项目: 国家重点研发计划(2017YFB0503605); 国家自然科学基金项目(41901395, 41671401) [Foundation: National Key R&D Program of China, No.2017YFB0503605; National Natural Science Foundation of China, No.41901395, No.41671401]

作者简介: 杜云艳(1972-), 女, 研究员, 博士生导师, 主要从事地理事件的人类活动感知。E-mail: duyuy@lreis.ac.cn

空数据模型方面, 自 Langran 等提出“状态—事件—依据”3种时空数据表达类型后, 基于事件的GIS时空数据模型被提出并不断地发展^[3]。该类模型的主要特征是同时描述时空变化过程和触发这种变化(事件)的原因与结果, 显式存储事件序列, 顾及状态与因果关系, 最具代表性的模型是ESTDM(Event based Spatio-Temporal Data Model)^[4]。在该模型中地理现象的时空变化通过一连串的事件序列和基础网格图层来表达, 每一个事件包含了事件发生时间和若干记录不同地理位置上属性变化的组件, 通过事件记录地理现象的每一次重要变化, 用事件的多个组件来反映现象在空间上的演化结果。后期不同学者不断地改进和完善该类模型^[5-7]以进一步增强地理现象变化的原因及空间语义的表达。但上述研究依然侧重于地理状态变化或起因的描述, 缺乏对地理事件时空过程建模及对其所涉及人类活动响应的相关表达与分析。

在地理事件过程分析与模拟方面, 由于自然灾害类事件对地表要素的改变在短时间内容极为剧烈, 因此以往研究多关注自然灾害类事件的影响评估与模拟预测。上世纪90年代以来, 自然灾害的遥感监测和高精度灾情评估与预警成为地理事件过程分析的研究核心^[8-10]。基于卫星遥感的应急制图(Satellite-based Emergency Mapping Systems, SEM)系统被广泛用于自然灾害监测与应急管理中, 在海啸、地震、飓风、洪水等灾害的动态过程监测、协调救灾等应急管理中发挥着不可替代的作用^[11-14]。但是由于SEM系统无法提供灾害过程中人类活动和情绪变化的定量监测, 依然无法监测灾害事件下的人类活动, 但这些信息往往对灾害的紧急救援与管理是至关重要。而关于公共卫生事件的研究, 由于其主要依靠流行病调查数据, 传统研究主要在GIS时空模型支撑下开展流行病传播链条及区域传播时空差异的分析, 2003年SARS流行病研究中GIS的空间分析与模拟发挥了巨大作用^[15-17]。但这些研究都未能关注流行病对社会公众的时空行为模式、舆论和情感等带来的影响分析, 更谈不上对其进行模拟与预测。

随着基于位置服务(Location Based Service, LBS)的数据及社交媒体等网络大数据的爆炸式增长和人类活动大数据定量刻画手段的快速发展, 地理事件对人类活动的影响以及公众对地理事件的网络参与度引起了多个领域的广泛关注, 进而也对地理事件的时空认知、建模方法和分析框架等提出了巨大的挑战^[18-20]。当前基于人类活动大数据开展的地理事件相关研究为特定地理事件数据的近实时获取^[21-25], 人类社交活动强度随地理事件的时空变化分析以及地理事件损失与影响的快速推断^[26-29]等。但上述研究大都针对某次特大事件进行具体地实例分析, 对人类活动在某类典型地理事件下响应的区域差异及普适性模式的分析相对缺乏, 更谈不上系统分析地理事件中人地关系耦合及其驱动机制。

另一方面, 图数据模型的发展和成熟使其逐渐成为应对LBS数据的一种新的表达与建模手段。图数据模型是采用图结构对现实世界进行数据特征抽象与描述的一组概念和定义, 是图数据库中数据的存储方式, 主要包括数据类型与结构描述、数据操作与查询语言, 以及数据完整性规则^[30-32]。该模型以离散数学的图论为数学基础, 通过节点和边来抽象客观世界中的对象和关系^[33-34]。目前常见的图数据模型包括属性图模型、超图模型和RDF三元组模型, 已成为新的地理时空数据建模手段, 并广泛应用于地理时空表达与建模领域^[35-37]。地理时空图模型是基于传统的图建模思想, 提出并不断地完善可以准确的抽象地理对象并表达对象之间空间、时间以及亲缘关系的数据模型^[33, 38-41]。相比于传统的地理数据模型, 时空图模型一方面继承了地理对象模型对空间、时间和属性整体表达的优点, 也实现了地理对象与对象关系的一体化表达。由于采用无索引邻接技术, 在处理复杂地理时空关系时具有级跳查询能力^[35], 相比于地理对象—关系模型具有明显的时空关系查询与分析优势^[37], 从而为分析复杂的地理事件奠定了基础。当前地理时空图模型的

应用主要针对的是具有复杂时空演化关系及行为的地理现象,发现其演化特征和时空模式^[34],但对地理事件所引起的人类活动或者级联响应事件的表达涉及较少,亟需加强研究。

因此,综上所述,在人类活动大数据日渐丰富的背景下,地理事件的认知与研究范畴发生了根本性的变化,地理事件不仅其自身演变同时其演变的内因与外因都得到了广泛的关注。但当前的地理事件模型对不同地理现象间的相互关系及作用的表达以及地理事件及事件所引起的人类活动的级联事件表达都涉及的较少,难以开展地理事件演变的前因与后果分析。因此,无论是从GIS数据模型还是从大数据背景下地理事件的分析与挖掘的研究需求出发,都亟需开展融合“人”类活动大数据的地理事件建模与分析,充分发挥地理时空图模型在复杂关系表达上的优势,结合当前地理事件自身特点构建基于地理事件图数据概念模型与分析框架。本文不仅为GIS数据模型研究提供新视角,还能地为地理事件级联效应及其人类活动响应的时空规律与区域模式挖掘奠定基础。

2 基于图模型的地理事件建模

2.1 地理事件概念及分类

从时空数据模型研究出发,当前的大数据背景下首先需要拓展的就是地理事件的概念。但目前为止地理事件还没有明确统一的概念,从目前国内外学者所关注的地理事件建模与分析研究对象来看,可归纳为两类:以地理变化为主题的事件和以触发变化为主题的事件。前者强调在某时间范围某空间区域地理对象发生了变化,如犯罪、传染病爆发、交通事故、异常变化等,该地理事件不关注发生到结束时间的长短,仅强调发生变化。比如,谢炯等利用地理变元把土地利用变化定义为地理事件^[42];郑宇把时空集体异常定义为地理事件^[43];薛存金等把海洋环境要素的异常变化过程定义为海洋异常事件等^[44]。

以触发变化为主题的事件认为地理实体或对象的产生、状态的变化、对象的进化和消亡都由特定事件触发,把引起这种对象变化的事件定义为地理事件。地理事件不仅仅是状态开始和结束的标志,同时也是地理实体或对象演化历程的关系或纽带^[45]。驱动对象发生变化的直接原因定义为原子事件,包括对象的产生、发展及消亡,而在宏观尺度下多个具有因果关系的原子事件组成形成复合事件,它对应于一个更宏观的对象变化^[46-48]。

无论是以地理变化为主题还是以触发地理变化为主题的地理事件表达了地理对象在一定的时空范围内相互关系,这种相互关系既包含了地理对象内部的演变关系^[49],也包含了地理对象之间的关联关系^[50]。目前地理大数据的发展促进了人地相互关系的研究,特别是地理异常事件对人类活动造成的影响和人类面对异常地理事件时如何快速响应,已成为地理科学研究的热点问题^[51],迫切需要新的地理事件模型支撑人地相互关系的研究。因此,本文把地理事件归纳为在一定的时空域内,地理主体对象(现象)的行为或空间结构、属性或组合发生了变化,且这种变化对人类或地理客体对象(现象)造成了重要影响的事件,包括5个基本要素:地理空间、时间、主题属性、对象变化和变化引起的相互关系。本文定义的地理事件以描述人类活动与地理客体对象(对象间)相互作用为主,重点刻画地理事件主体与客体的作用关系,强调主体对象的行为导致客体对象(现象)发生了重要变化并产生重要的影响。如:疫情爆发导致的人类时空行为模式的变化,台风到来引起的城市正常节律的异常,道路交通的管制导致的车流量的异常减少等。

2.2 地理事件语义与图表达模型

地理大数据背景下,迫切需求对人地关系开展研究来解决地理事件的前因后果,因此,地理事件由地理对象(现象)、对象变化和变化引起的相互关系组成。地理对象包括地理空间位置、时间和主题属性3个要素,用于描述地理事件的时空特征;地理对象变

化是地理属性、空间位置或空间关系随时间的变化,用于描述地理事件的时空变化,在不同的时间尺度下,地理事件的变化性质存在差异;地理事件的相互作用用于表达地理事件与外界环境的交互,刻画地理事件发生的前因后果:外界地理事件在何时、何地引起了本地理事件的发生发展和本地理事件的变化又如何驱动外界其他地理事件的变化,比如,城市暴雨事件在何时何地引起了城市内涝,而城市内涝又如何引起城市交通拥堵等。地理事件的语义描述框架如图1所示。

根据地理事件描述语义,面向大数据的地理事件可以采用属性图模型进行表达,由2类节点和3类边组成,如图2所示。

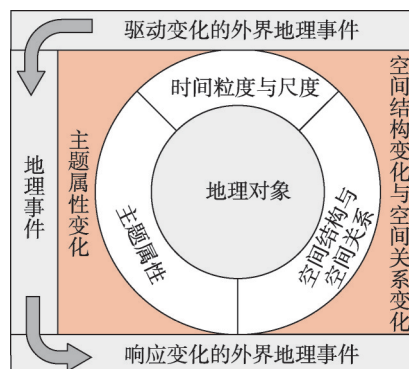


图1 地理事件描述语义

Fig. 1 Description semantics of geographic event

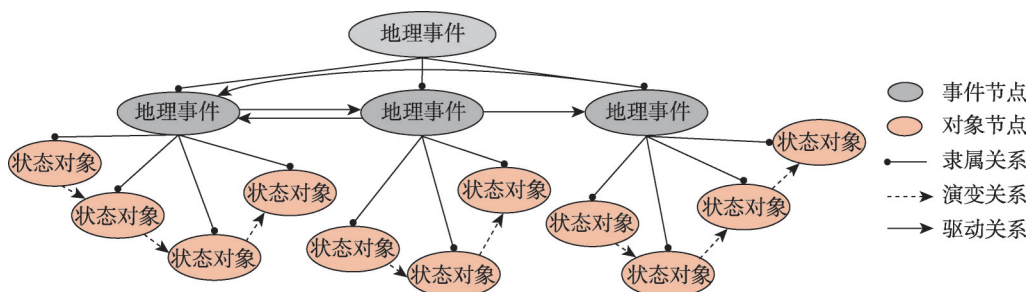


图2 地理事件图表达模型

Fig. 2 Graph representation model of geographic events

地理事件的两类节点包括事件节点和对象节点。事件节点是对象的父节点,用于表达地理事件的空间覆盖范围、持续时间、事件的主题属性和包含的地理对象;对象节点用于表达某时刻的地理对象的空间位置、空间结构、时间信息和主题属性。

地理事件的3类边包括事件—对象的隶属边、对象—对象的演变边和事件—事件的相互作用边。事件—对象的隶属边表达地理事件与地理对象之间的隶属关系,表达地理事件包含多少个地理对象,某个地理对象隶属于哪个地理事件;对象—对象的演变边表达地理对象在时刻间的关系,主要包括发展、合并、分裂、消弱关系,用以表达地理对象在地理事件范围内的演变;事件—事件的相互作用边表达地理事件间的相互作用关系,比如,事件之间的响应关系、驱动关系等。

2.3 地理事件图存储模型

本文采用Neo4j图数据库对地理事件图存储模型进行说明。Neo4j是一款基于图模型的开源数据库,在动态数据存储和关系查询方面比对象—关系数据库具有明显优势^[35, 37]。Neo4j由一个节点存储文件和一个关系存储文件存储,节点存储文件用来存储地理事件的两类节点信息:事件节点用于存储地理事件的基本信息,包括事件ID、空间覆盖范围、时间、主题属性类型、属性的基本统计信息等;对象节点用于存储时刻状态对象信息,包括对象ID、空间位置、空间形态、发生时刻、主题属性等。关系存储文件用来存储地理事件的3类边信息:隶属边存储边ID、边类型、父节点ID和子节点ID;相互作用边存储边ID、边类型、前节点ID和后节点ID;演变边存储边ID、边类型、前节点ID和后节点ID。节点存储文件和关系存储文件采用双向链表进行连接,图存储表达模型如图3所示。

3 基于大数据的地理事件分析与挖掘

大数据背景下地理事件的概念得到拓展,重点强调了事件之间的相互作用及其表达。相应地,在分析方法及分析思路,传统地理事件大都针对事件的自然要素变化进行时空模式挖掘,缺乏对地理事件前因与后果的级联事件分析与挖掘。因此,在上述地理事件概念拓展基础上,对大数据背景下的地理事件分析与挖掘方法体系也亟待深入地拓展。本节针对性地讨论与构建基于地理大数据的事件分析与挖掘框架,设计该框架下地理事件分析与模拟方法,并利用人类活动

与人地交互影响的地理演变模式和级联模式来修正和完善地理事件时空表达模型与方法。

3.1 地理事件分析与挖掘方法框架

基于地理大数据的地理事件分析与挖掘框架强调在突破传统地理空间对事件概念与认知局限性基础上,深入理解当前重大地理事件在地理活动空间和虚拟网络空间的传播与演化特征,开展地理大数据的事件分析与挖掘,主要包括地理事件多源数据感知与认知模型、地理事件重构、表达与地理事件库构建和地理事件演变行为和级联效应挖掘方法。地理事件多源数据感知与认知模型基于对地观测、位置服务和社交感知大数据等多源数据收集与处理技术,从“自然空间—地理活动空间—虚拟网络”多个空间开展事件时空特征、演化语义、人类情绪、社会反响的特征分析,实现地理事件的多源数据感知和认知。地理事件重构与表达采用传统地理空间的历史事件过程重构和基于大数据驱动事件发现方法两种思路开展典型地理事件案例获取,获取两大类地理事件,其一是具有细粒度演化过程的“天—地—人”交互耦合的事件(具有完整的历史演化过程和多维空间交互作用),其二是粗粒度的具有级联效应地理事件,重点关注事件直接或间接引起的人类各种活动响应事件;然后采用GIS时空演化模型开展地理事件时空过程描述,采用关系模型开展级联事件时空传播的描述与舆情话题时空表达,而整体上采用图数据模型进行地理事件地理空间和虚拟空间一体化表达;最后,基于图数据组织模型开展地理事件的表达和数据建模,构建地理事件库,实现地理事件的组织存储。地理事件演变行为和级联效应挖掘方法基于构建的地理事件数据库,分别从事件时空演化的人类空间行为—舆论—情感等耦合建模和事件所导致的人类行为等区域模式两个不同的视角开展人类活动定量感知方法研究,重点在事件级联效应时空模拟、人类行为、情绪及舆情定量评价方法、时空聚类分析、时空可视化挖掘等方法。基于地理大数据的地理事件分析与挖掘框架如图4所示。

3.2 地理事件发现方法研究

地理事件发现是实现地理大数据事件分析与挖掘的核心,而地理大数据支撑的地理事件发现与传统地理空间的发现具有较大的差异。地理事件发现主要采用两大类方法,其一是历史重大事件的重构方法,其二是基于数据驱动的事件发现方法。传统对地理事

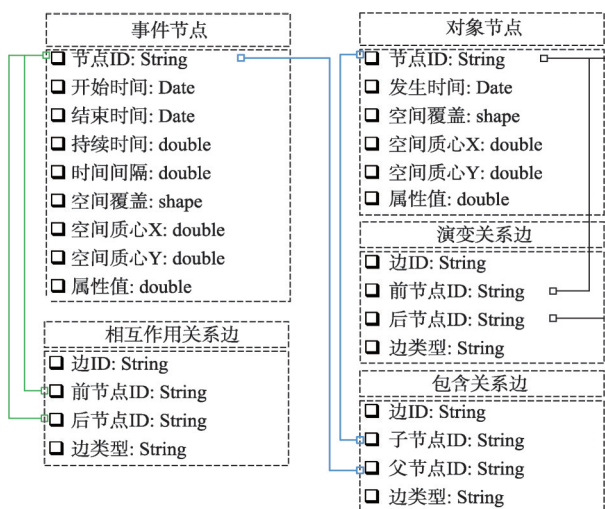


图3 地理事件图存储模型

Fig. 3 Graph storage model of geographic event

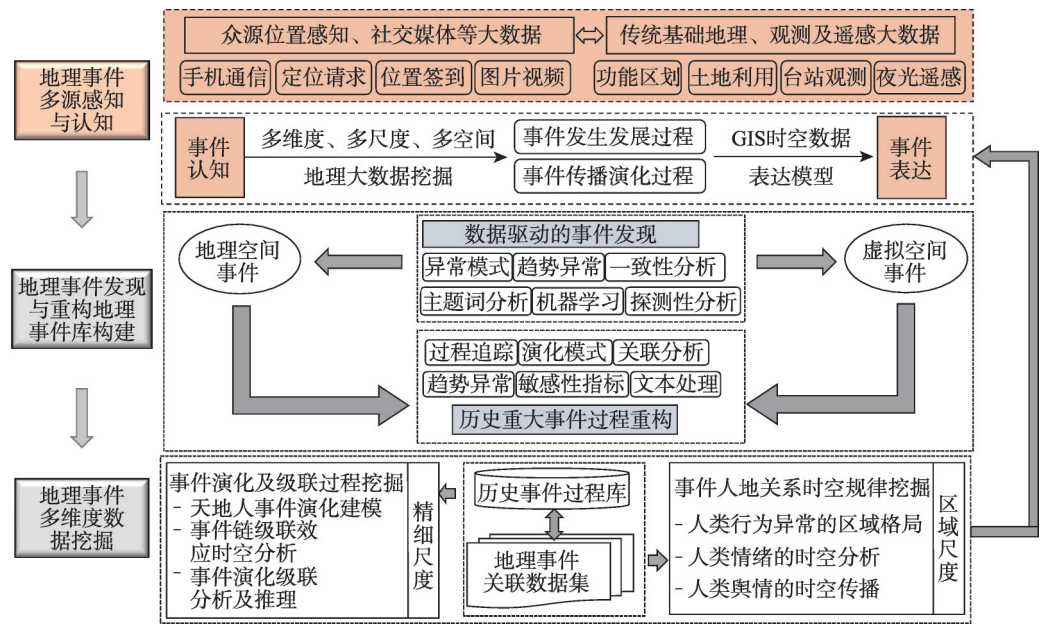


图4 地理大数据事件分析与挖掘基本框架

Fig. 4 Basic framework of geographic big data event analysis and mining

件研究更多采用事件重构方法，应用最多领域为自然灾害类突发性事件，分析与建模更多关注于事件发生后地理要素的变化及影响评估。但随着地理事件公众的网络参与度的提升及其演化过程中人类时空、情感和舆论等定量获取技术的发展，对地理事件发现则更多采用数据驱动方法^[52-54]。然而由于地理事件在地理要素和社交媒体数据呈现的不同时空特征，导致地理事件的数据驱动方法多种多样。总体上，地理事件发现方法可以归纳为基于统计模型、基于时空数据挖掘和基于人工智能的3大类（图5）。① 基于统计模型的事件发现方法：主要包括时空扫描统计、核密度、似然比检验等方法。其中最常用的是时空扫描统计方法，其基本思想是设定一个圆柱形的扫描窗口（圆柱底对应地理区域，圆柱高对应时间），利用扫描窗口内外的实际事件数量和预期事件数量构造检验统计量LLR，统计评价选出LLR最大窗口即为事件时空区域^[55-58]。② 基于时空数据挖掘的事

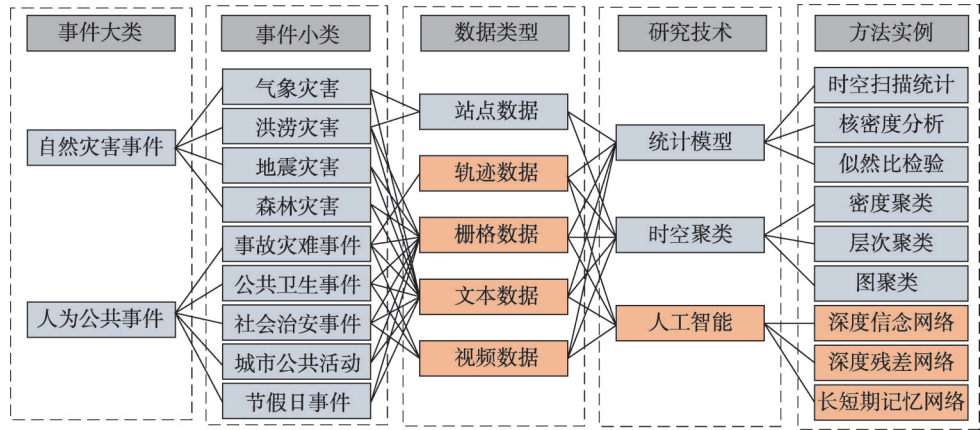


图5 基于大数据的地理事件发现方法分类

Fig. 5 Classification of geographic event discovery methods based on big data

件发现方法:主要包括时空聚类、时间序列分析、时空异常探测和复杂网络挖掘等。时空聚类在传统地理空间的事件发现中普遍使用,近年来主要是结合社交媒体数据等开展新的数据源事件发现与探测研究^[59-62],反映人类时空行为各种时序数据,则主要采用时间序列与时空异常相结合的检测方法^[63];而复杂网络时空挖掘则主要针对近年来出现的流数据开展的地理事件发现的常用方法^[64]。③基于人工智能机器学习的事件发现方法主要是处理反映人类舆情的文本与视频等信息,与自然语言处理所涉及到的方法多有融合,成为网络虚拟空间地理事件发现的热点方法,包括深度信念网络,深度残差网络,深度卷积神经网络等^[65-67]。

综上,在大数据背景下,地理事件发现方法的本质是面向众源数据,基于地理事件的空间分布特征、时间演变特征、属性特征、或其综合特征,引入地理大数据挖掘的思想,实现地理事件的识别与提取,包括地理事件的影响范围,持续时间及其带来的对地理客体对象(现象)的作用及影响。在地理事件的发现过程中涉及到两个关键问题:其一是如何解决众源数据的有偏及代表性问题;其二是如何解决不同空间中地理事件的一致性匹配与关联问题。这些是后面开展地理事件级联效应分析的技术基础。

3.3 地理事件级联效应分析

随着地理事件发展过程中与人类相关位置信息(主动或被动)的快速收集与处理,对地理事件的分析并不仅仅局限于对客观世界变化的分析与影响评估,更多的涉及人类空间行为、情感与舆情等时空响应及事件级联效应分析。如城市降雨所引起的交通拥堵、公共卫生事件所引发的人类活动变化、舆情关注等。虽然传统地理事件的研究也涉及事件链分析,比如台风引起的降雨和洪涝等,但以往研究由于没有直接涉及人的时空动态分布及情感、舆情等活动分析,其复杂性和跨学科特定并不突出。因此当前在地理大数据快速涌现的情况下,综合分析地理事件演化所引起的地理要素突变、人群时空聚集、舆情与情感的异常特征,用多种时空统计与挖掘方法解析事件的级联效应及其内在因果关系,会成为事件建模与分析必不可少的一部分。

基于上述地理事件发现方法获取到相关数据集开展地理事件级联效应分析(图6)。首先,需要对不同空间用不同事件探测方法获取到的地理事件集合进行地理事件的匹配分析。分两个方面开展事件匹配。如果不同空间探测到的是同一次历史事件,则选择时空匹配方法进行该次事件时空演化过程的重构;如果不是同一次历史事件,则从全国—城市—城市内部3个尺度分别开展同类型的相关事件时空语义匹配,构建关联事件多尺度数据集,以图数据模型开展事件及其关联事件的语义关系描述。其次,针对同一历史事件的过程重构,则在精细粒度采用时空可视化、时空一致性、时空关联规则挖掘或深度学习等模型方法开展“自然—地理—人类”多个空间的地理事件级联效应分析及演化模式分析与预测。第三,针对同种类型相互关联的地理事件多尺度数据集,则采用一致性检验、时空聚集模式、时空谱分析、机器学习、及贝叶斯统计分析等方法开展地理事件不同空间交互影响探索性分析,获取同一个或同一类地理事件在不同时空尺度上级联关系及时空聚类模式;上述不同尺度地理事件的级联效应分析为不同层面、不同视角下事件的人类活动感知奠定方法基础。

4 实例分析

当短时大量的降水超过城市的排水能力时会导致城市发生积水内涝灾害,严重影响人们的正常出行、工作与生活,甚至危害生命财产安全,因此城市暴雨事件不局限于暴

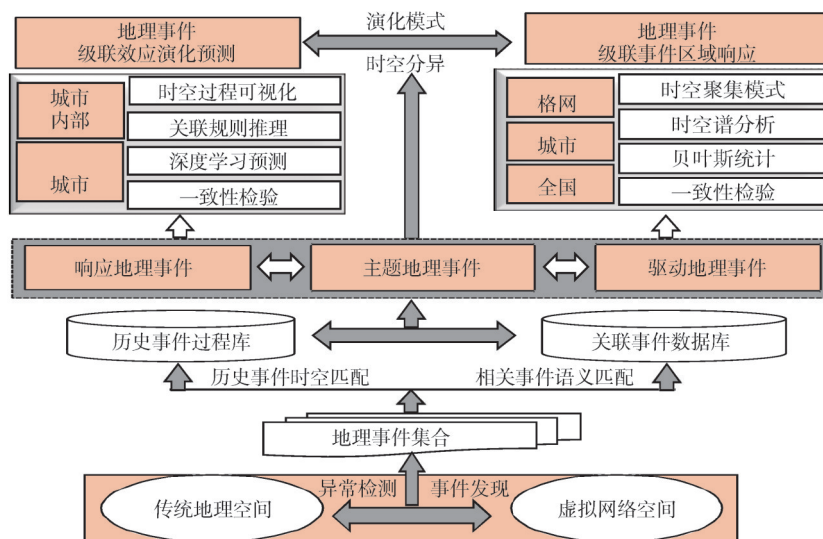


图6 地理事件级联效应分析框架

Fig. 6 Cascading effect analysis framework of geographic events

雨事件自身的演变，还会导致人类出行异常，引发城市交通拥堵等级联效应。本文以城市暴雨事件为例，通过含位置信息的大数据感知暴雨事件给城市人类活动带来的异常变化与影响，从事件表达、异常挖掘、级联效应三方面开展实例验证，诠释地理事件建模与分析方法在暴雨事件研究中的价值。

4.1 城市暴雨事件时空表达

城市暴雨事件的时空过程及其影响可以通过图数据模型进行表达。基于图模型，可以将最小时间分辨率观测到的降雨过程切片作为一种离散的“状态”节点，状态节点可以记录此刻降雨的强度和范围，而整个降雨事件通过一系列时间相邻的状态节点来表示其时空过程。降雨引发的道路积水、房屋及地下车库淹水等次生事件，也可以通过节点来记录事件发生的时间、损失程度、恢复时间等关键信息，并且可以通过图的连接来描述这些事件与降雨事件之间的事件链关系（图7）。

暴雨事件对城市的影响表现为城市日常节律的扰乱。图数据库中的节点可以记录异常发生的时间、位置及异常程度，从而实现暴雨事件影响的数据表达。由暴雨事件引发

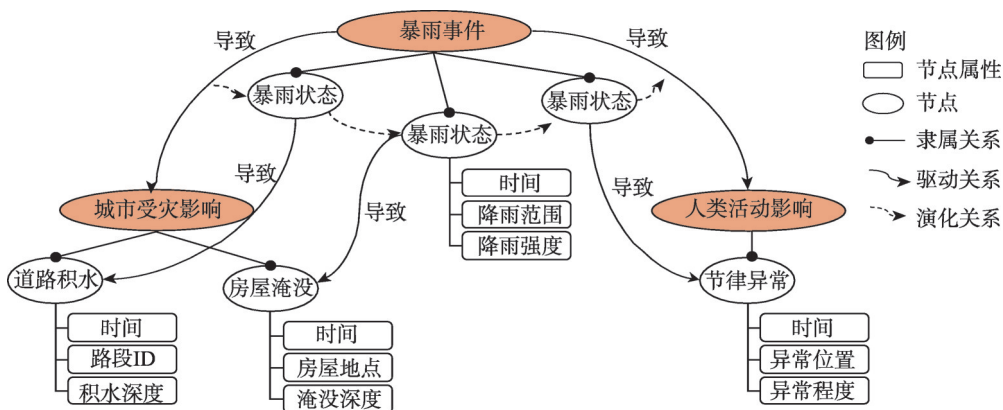


图7 城市暴雨事件的图模型表达

Fig. 7 Graphical model representation of urban rainstorm events

的节律异常通常会与暴雨过程在时空上具有一致性或具有时滞的一致性,通过图的连接关联暴雨事件与节律异常,可以表达暴雨事件引发的人类活动异常现象,从而有助于检索和发现事件的影响,为分析暴雨灾害的级联影响提供数据支持。

相比目前常用的关系数据模型,图数据模型更灵活、直观地表达了暴雨事件中降雨动态过程与人类活动异常、受灾状况等不同对象之间复杂的相互关系。与关系型数据库中通过建立暴雨状态表和节律异常表的主键—外键来描述对象关系不同,图模型更灵活地建立了数据对象之间的关系,而非整个数据集之间的关系,降低了数据表达冗余,如4.2节通过异常模式挖掘而发现的暴雨与人类活动之间的响应关系,图数据模型将节律异常和与之关联的降雨状态进行关联,而无需为所有时刻的降雨状态都建立关联。同时,相比传统数据库检索,图数据库在级联检索效率上的优势^[3]是实现快速检索与降雨关联的异常和灾情的重要,有效保障了应对灾害的快速响应能力。

4.2 暴雨事件中的人类活动异常特征发现

通过大数据异常挖掘可以识别暴雨事件引发的城市节律异常。以北京市2018年7月16日的暴雨事件为例,本文采集了2018年7月9—16日连续8天的位置请求数据。该数据是手机用户使用腾讯地图位置服务时的位置数据,来自腾讯位置大数据网站(<https://heat.qq.com/>),按1 km×1 km的网格进行了逐时聚合,可用于刻画不同位置人们使用LBS服务的节律。北京“7·16”暴雨事件是2018年北京入汛后最强的降雨过程。7月15日20时—16日20时,北京市平均降雨量为60.3 mm,城六区的平均降雨量达到93.8 mm。7月16日凌晨3时的小时降雨量达11 mm,白天降雨集中在早高峰8时和晚高峰18时两个时段。

在精细的网格尺度上定位请求量的异常节律检测,首先通过STL时间序列分解方法剔除趋势和周期变化,然后采用GESD统计方法判断残差分量中的正、负异常时间及程度^[68]。检测结果以正异常为主,负异常网格数量较少且与暴雨事件无明显响应特征。从时间上统计正异常网格数量与偏离量的逐时变化,发现7月16日8—9时、16—17时、19—20时的值超过一周统计的同期Q3水平,而夜间22—24时低于同期中值(图8)。从空间上统计正异常网格的核密度分布,发现上述时段的异常网格与降雨的空间分布十分一致。可见,暴雨事件引发了定位请求量的节律异常变化,并且这种变化在时间和空间上与降雨的时空过程是比较一致的。

4.3 暴雨事件的级联影响分析

基于城市节律的异常变化,本文进一步对暴雨事件对人类活动造成的级联影响及其与城市环境要素之间的关系进行了分析。我们选取了在2017年8月发生了严重暴雨事件的8个城市,构建了异常评分指标来综合每次暴雨事件引发的城市节律正异常在数量、程度和空间密度上的特征,进而分析异常评分指标与降雨强度之间的关系及这种关系在城市间的差异^[69](图9)。从降雨强度与异常评分指标的统计结果来看,暴雨事件的降雨强度越大,引发的节律正异常发生概率越高。但不同城市对雨强大小的敏感性存在差异,即引发城市出现人类活动异常的降雨强度在不同城市存在差异。通过拟合各城市雨强与异常评分指标的线性关系,可以发现哈尔滨的回归系数较大,而海口较小,表明同样强度的降雨对哈尔滨的影响要比对海口的影响更大。这种差异表明了理解暴雨事件对人类活动影响还需要考虑城市地理环境的差异。从城市的气候特征(如年均降雨量)和城市排水能力来看,海口的年均降雨量超过1600 mm,超过哈尔滨,同时海口的排水管密度为11.74 km/km²,大于哈尔滨。所以,相比哈尔滨,海口市的居民每年会经历更多的降雨,而城市有效的排水能力保障了居民的正常工作与生活。

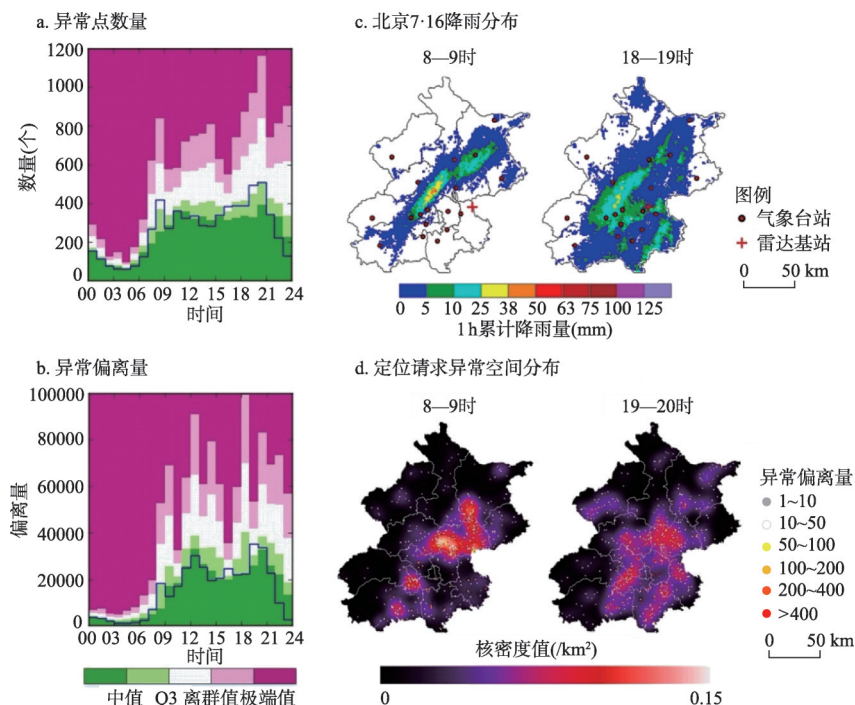


图8 北京市2018年7月16日暴雨事件及其人类活动影响

Fig. 8 The July 16, 2018 heavy rain event in Beijing and its impact on human activities

以上实证研究展示了地理事件的图模型表达和挖掘这套方法体系在描述和理解暴雨事件影响下的人类活动节律变化及其与城市环境因素之间关系的有效性与实用性。与目前暴雨事件的研究方法相比,该方法系统性地将时空图数据表达与异常模式挖掘相结合,不仅有效描述了地理事件的时空变化,更突出了地理事件对“人”的影响,为揭示事件视角下特殊的人地关系提供了有效的研究手段。

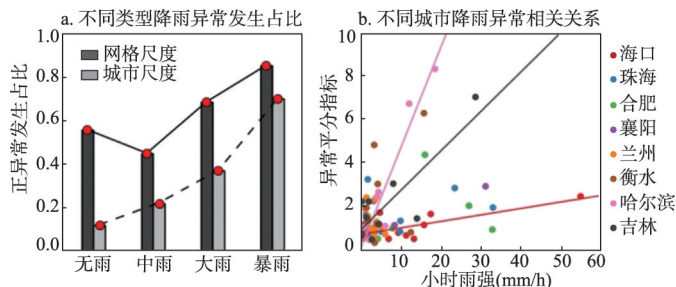


图9 暴雨事件影响的城市间差异

Fig. 9 Differences between cities affected by rainstorm events

5 总结

地理事件的建模与分析是地理信息机理研究的核心内容,但由于各种社交媒体大数据的出现,使得人类对地理事件响应与人—地理事件互影响的研究成为可能,也因此给地理事件的建模与分析带来前所未有挑战。具体为3个方面:①由于大数据时代,公众对地理事件网络参与度提升,地理事件在改变各种地理要素之外,还在一定程度上引起了人类各种社会、经济、情感等活动的变化,这些对如何界定地理事件提出了挑战,其概念内涵和外延都有很大程度地拓展;②面对新的地理事件概念及其组成,如何开展地理事件数据建模成为地理事件的关键研究问题。如何从地物要素变化的表达视角转向直

接表达“事件”本体及其前因和后果成为大数据时代地理事件数据建模的关键点;③如何有效融合“虚拟空间”的地理事件发现与传播模拟分析方法,构建地理大数据支撑下的地理事件时空分析与推理方法框架,实现地理事件下“人—地”交互关系的时空规律与区域模式挖掘成为新时代地理事件分析的核心内容。

本文针对上述3项具体挑战,本文详细阐述了地理大数据支撑下的事件建模与分析理论框架,并以城市强降雨为例进行了实例研究。①界定地理事件为在一定时空域内地理事物(现象)空间结构、属性或组合发生了变化,且对人类或其它地理事物或地理事物造成了影响的事件,从概念上拓宽了地理事物仅仅局限于传统地理空间要素变化的约束;②采用图数据模型开展地理事物“本体”及其关联事件的表达与建模,使得地理事物的数据建模可以有效地融合传统地理要素变化表达与大数据支撑的“人类活动”的表达;③给出了地理大数据支撑下的事件时空分析框架,以地理事物下“人类活动”为桥梁构建地理事物的“地理空间”与“虚拟空间”的一体化级联效应分析与建模;最后以多个城市的暴雨事件开展地理事物不同尺度的一体化分析实例研究,诠释了暴雨事件引起的人类活动响应的时空模式及区域差异。希望通过本文的研究成果,能抛砖引玉,激发大数据时代地理事物分析与建模新思路与研究热潮。

参考文献(References)

- [1] Nelson K. Event Knowledge: Structure and Function in Development. London: Psychology Press, 1986.
- [2] Zhou Wen, Liu Zongtian, Kong Qingping. A survey of event-based knowledge processing. Computer Science, 2008(2): 160-162, 184. [周文, 刘宗田, 孔庆平. 基于事件的知识处理研究综述. 计算机科学, 2008(2): 160-162, 184.]
- [3] Langran G. Issues of implementing a spatiotemporal system. International Journal of Geographical Information Science, 1993, 7(4): 305-314.
- [4] Peuquet D J, Niu D. An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. International Journal of Geographical Information Systems, 1995, 9(1): 7-24.
- [5] Garcia Alvarez M, Morales J, Kraak M J. Integration and exploitation of sensor data in smart cities through event-driven applications. Sensors, 2019, 19(6): 1372. DOI: 10.3390/s19061372.
- [6] Gebbert S, Pebesma E. The GRASS GIS temporal framework. International Journal of Geographical Information Science, 2017, 31(7): 1273-1292.
- [7] Zhao Y J, Zhou X G, Li G Q, et al. A spatio-temporal VGI model considering trust-related information. ISPRS International Journal of Geo-Information, 2016, 5(2): 10. DOI: 10.3390/ijgi5020010.
- [8] Xie Pan, Wang Yanglin, Liu Yanxu, et al. Incorporating social vulnerability to assess population health risk due to heat stress in China. Acta Geographica Sinica, 2015, 70(7): 1041-1051. [谢盼, 王仰麟, 刘焱序, 等. 基于社会脆弱性的中国高温灾害人群健康风险评价. 地理学报, 2015, 70(7): 1041-1051.]
- [9] Yu M Z, Yang C W, Li Y. Big data in natural disaster management: A review. Geosciences, 2018, 8(5): 165. DOI: 10.3390/geosciences8050165.
- [10] Ji M, Liu L, Buchroithner M. Identifying collapsed buildings using post-earthquake satellite imagery and convolutional neural networks: A case study of the 2010 Haiti earthquake. Remote Sensing, 2018, 10(11): 1689. DOI: 10.3390/rs10111689.
- [11] Voigt S, Giulio F, Lyons J, et al. Global trends in satellite-based emergency mapping. Science, 2016, 353(6296): 247-252.
- [12] Tong X H, Hong Z H, Liu S J, et al. Building-damage detection using pre- and post-seismic high-resolution satellite stereo imagery: A case study of the May 2008 Wenchuan earthquake. ISPRS Journal of Photogrammetry & Remote Sensing, 2012, 68(3): 13-27.
- [13] Worthem A B, Madore G, Imahori J, et al. Emergency response imagery related to hurricanes Harvey, Irma, and Maria. American Geophysical Union, Fall Meeting 2017: NH23E-2842.
- [14] Cervone G, Schnebele E, Waters N, et al. Using social media and satellite data for damage assessment in urban areas during emergencies. Seeing Cities Through Big Data, 2017: 443-457.
- [15] Jiang J H, Qu J H, Xu D Z, et al. Application of geographical information system technology to epidemiological surveillance and prevention and cure decision-making for SARS. Journal of Medical Colleges of PLA, 2006(3): 186-

- 188.
- [16] Deng Jiqiu, Wu Qianhong, Liu Aizhong. Analysis about environmental factor of SARS epidemic based on GIS logical information method. *Chinese Journal of Health Statistics*, 2006, 23(4): 301-305. [邓吉秋, 吴堃虹, 刘爱忠. 基于 GIS 逻辑信息法的 SARS 流行环境因素分析. *中国卫生统计*, 2006, 23(4): 301-305.]
- [17] Yu X. Research and preliminary construction of GIS-based epidemiological monitoring platform [D]. Guizhou: Guizhou University of Finance and Economics, 2013.
- [18] Cheng Changxiu, Shi Peijun, Song Changqing, et al. Geographic big-data: A new opportunity for geography complexity study. *Acta Geographica Sinica*, 2018, 73(8): 1397-1406. [程昌秀, 史培军, 宋长青, 等. 地理大数据为地理复杂性研究提供新机遇. *地理学报*, 2018, 73(8): 1397-1406.]
- [19] Pei Tao, Liu Yaxi, Guo Sihui, et al. Principle of big geodata mining. *Acta Geographica Sinica*, 2019, 74(3): 586-598. [裴韬, 刘亚溪, 郭思慧, 等. 地理大数据挖掘的本质. *地理学报*, 2019, 74(3): 586-598.]
- [20] Zhou Chenghu. The value of spatial data in the era of big data: Review of "theory and application of spatial data mining. *Acta Geographica Sinica*, 2016, 71(7): 1281. [周成虎. 大数据时代的空间数据价值:《空间数据挖掘理论与应用》评介. *地理学报*, 2016, 71(7): 1281.]
- [21] Cole T, Wanik D, Molthan A, et al. Synergistic use of nighttime satellite data, electric utility infrastructure, and ambient population to improve power outage detections in urban areas. *Remote Sensing*, 2017, 9(3): 286. DOI: 10.3390/rs9030286.
- [22] Zhao X Z, Yu B L, Liu Y, et al. NPP-VIIRS DNB daily data in natural disaster assessment: Evidence from selected case studies. *Remote Sensing*, 2018, 10(10): 1526. DOI: 10.3390/rs10101526.
- [23] Román M, Stokes E, Shrestha R, et al. Satellite-based assessment of electricity restoration efforts in Puerto Rico after Hurricane Maria. *PLOS ONE*, 2019, 14(6): e0218883. DOI: 10.1371/journal.pone.0218883.
- [24] Wilson R, Erbach-Schoenberg E Z, Albert M, et al. Rapid and near real-time assessments of population displacement using mobile phone data following disasters: The 2015 Nepal Earthquake. *PLoS Currents*, 2016, 8. DOI: 10.1371/currents.dis.d073fbee328e4c39087bc086d694b5c.
- [25] Panteras G, Cervone G. Enhancing the temporal resolution of satellite-based flood extent generation using crowdsourced data for disaster monitoring. *International Journal of Remote Sensing*, 2018, 39: 1459-1474.
- [26] Balakrishnan S, Zhang Z M, Machemehl R, et al. Mapping resilience of Houston freeway network during Hurricane Harvey using extreme travel time metrics. *International Journal of Disaster Risk Reduction*, 2020, 47: 101565. DOI: 10.1016/j.ijdrr.2020.101565.
- [27] Mudigonda S, Ozbay K, Bartin B. Evaluating the resilience and recovery of public transit system using big data: Case study from New Jersey. *Journal of Transportation Safety & Security*, 2019, 11(5): 491-519.
- [28] Yuan M. Geographical information science for the United Nations' 2030 agenda for sustainable development. *International Journal of Geographical Information Science*, 2020, 35(1). DOI: 10.1080/13658816.2020.1766244.
- [29] Ma T, Pei T, Song C, et al. Understanding geographical patterns of a city' diurnal rhythm from aggregate data of location-aware services. *Transactions in GIS*, 2019, 23(1): 104-117.
- [30] Heckmann T, Schwanghart W, Phillips J D. Graph theory: Recent developments of its application in geomorphology. *Geomorphology*, 2015, 243: 130-146.
- [31] Siabato W, Claramunt C, Ilarri S, et al. A survey of modelling trends in temporal GIS. *ACM Computing Surveys*, 2018, 51(2): 1-41.
- [32] Liu W B, Li X G, Rahn D A. Storm event representation and analysis based on a directed spatiotemporal graph model. *International Journal of Geographical Information Science*, 2016, 30(5): 948-969.
- [33] Del Mondo G, Rodríguez M A, Claramunt C, et al. Modeling consistency of spatio-temporal graphs. *Data & Knowledge Engineering*, 2013, 84: 59-80.
- [34] Pokorný J. Graph Databases: Their power and limitations. *International Conference on Computer Information Systems and Industrial Management*, 2015. DOI:10.1007/978-3-319-24369-6_5.
- [35] Robinson I, Webber J, Eifrem E. Graph Database. 2nd ed. Sebastopol: O'Reilly Media Inc., 2015.
- [36] Zhu R, Guilbert E, Wong M S. Object-oriented tracking of the dynamic behavior of urban heat islands. *International Journal of Geographical Information Science*, 2017, 31(2): 405-424.
- [37] Xue C J, Wu C B, Liu J Y, et al. A novel process-oriented graph storage for dynamic geographic phenomena. *ISPRS International Journal of Geo-Information*, 2019, 8(2). DOI: 10.3390/ijgi8020100.
- [38] Stell J, Del Mondo G, Thibaud R, et al. Spatio-temporal evolution as bigraph dynamics. *International Conference on Spatial Information Theory*, 2011. DOI:10.1007/978-3-642-23196-4_9.
- [39] Liu Yuefeng, Lu Min, Yang Zhongzhi. Modeling dynamic networks in temporal GIS and realization methods. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2012, 48(3): 451-458. [刘岳峰, 陆恣, 杨忠智. 时态 GIS 中动态网络

建模及实现方法. 北京大学学报(自然科学版), 2012, 48(3): 451-458.]

- [40] Zhang Zhonggui, Wu Xincui. A general network spatiotemporal data model. *Computer Applications and Software*, 2015, 32(9): 84-87. [张忠贵, 吴信才. 一种通用的网络时空数据模型. *计算机应用与软件*, 2015, 32(9): 84-87.]
- [41] Thibaud R, Del Mondo G, Garlan T, et al. A spatio-temporal graph model for marine dune dynamics analysis and representation. *Transactions in GIS*, 2013, 17(5): 742-762.
- [42] Xie Jiong, Liu Renyi, Liu Nan. Study of land use multi-level change model and its transaction processing mechanism. *Journal of Zhejiang University(Science Edition)*, 2007(4): 469-475. [谢炯, 刘仁义, 刘南. 土地利用现状管理信息系统中土地利用多级变更模型的建立及其事务处理机制设计. *浙江大学学报(理学版)*, 2007(4): 469-475.]
- [43] Zheng Y, Zhang H C, Yu Y. Detecting collective anomalies from multiple spatio-temporal datasets across different domains. *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM Press, Bellevue, Washington, 2015: 2(1). DOI: 10.1145/2820783.2820813.
- [44] Xue C J, Song W J, Qin L J, et al. A spatiotemporal mining framework for abnormal association patterns in marine environments with a time series of remote sensing images. *International Journal of Applied Earth Observation & Geoinformation*, 2015, 38: 105-114.
- [45] Xu Z H, Bian F L. Spatio-temporal GIS data model based on event semantics. *Geo-spatial Information Science*, 2003, 6(3): 43-47.
- [46] Chen J, Jiang J. An event-based approach to spatio-temporal data modeling in land subdivision systems. *GeoInformatica*, 2000, 4(4): 387-402.
- [47] Lin G, Feng X, Wang L, et al. An event-centric object oriented spatio-temporal data model. *Acta Geographica et Cartographica Sinica*, 2002(1): 71-76.
- [48] Ikezaki M, Watanabe T. Event-based specification for managing change history of geographic information. *International Journal of Computer Science and Network Security*, 2006, 6(6): 106-110.
- [49] Yi J W, Du Y Y, Liang F Y, et al. A representation framework for studying spatiotemporal changes and interactions of dynamic geographic phenomena. *International Journal of Geographical Information Science*, 2014, 28(5): 1010-1027.
- [50] Luo W, Peuquet D J, Robinson A C, et al. A method for discovery and analysis of temporal patterns in complex event data. *International Journal of Geographical Information Science*, 2015, 29(9): 1588-1611.
- [51] Liu Yu. Revisiting several basic geographical concepts: A social sensing perspective. *Acta Geographica Sinica*, 2016, 71(4): 564-575. [刘瑜. 社会感知视角下的若干人文地理学基本问题再思考. *地理学报*, 2016, 71(4): 564-575.]
- [52] Liu Z, Du Y, Yi J, et al. Quantitative estimates of collective geo-tagged human activities in response to typhoon Hato using location-aware big data. *International Journal of Digital Earth*, 2020, 13(9): 1072-1092.
- [53] Wang N, Du Y Y, Liang F Y, et al. Spatiotemporal changes of urban rainstorm-related micro-blogging activities in response to rainstorms: A case study in Beijing, China. *Applied Sciences*, 2019, 9(21): 4629. DOI: 10.3390/app9214629.
- [54] Liu J Y, Xue C J, He Y W, et al. Dual-constraint spatiotemporal clustering approach for exploring marine anomaly patterns using remote sensing products. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, 11(11): 3963-3976.
- [55] Cheng T, Wicks T. Event detection using Twitter: A spatio-temporal approach. *PLOS ONE*, 2014, 9(6): e97807. DOI: 10.1371/journal.pone.0097807.
- [56] Shiode S, Shiode N. Network-based space-time search-window technique for hotspot detection of street-level crime incidents. *International Journal of Geographical Information Science*, 2013, 27(5): 866-882.
- [57] Nakaya T, Yano K. Visualising crime clusters in a space-time cube. *Transactions in GIS*, 2010, 14(3): 223-239.
- [58] Pang L X, Chawla S, Liu W, et al. On detection of emerging anomalous traffic patterns using GPS data. *Data & Knowledge Engineering*. 2013, 87(9): 357-373.
- [59] Shi Y, Deng M, Yang X X, et al. A framework for discovering evolving domain related spatio-temporal patterns in Twitter. *ISPRS International Journal of Geo-Information*, 2016, 5(10): 193. DOI: 10.3390/ijgi5100193.
- [60] Telang A, Deepak P, Joshi S, et al. Detecting localized homogeneous anomalies over spatio-temporal data. *Data Mining and Knowledge Discovery*, 2014, 28(5/6): 1480-1502.
- [61] Lee C H. Mining spatio-temporal information on microblogging streams using a density-based online clustering method. *Expert Systems with Applications*, 2012, 39(10): 9623-9641.
- [62] Feng W, Zhang C, Zhang W, et al. STREAMCUBE: Hierarchical spatio-temporal hashtag clustering for event exploration over the Twitter stream. *IEEE 31st International Conference on Data Engineering*, Seoul, South Korea, 2015: 1561-1572.
- [63] Liu Z, Du Y Y, Yi J W, et al. Quantitative association between nighttime lights and geo-tagged human activity dynamics during Typhoon Mangkhut. *Remote Sensing*, 2019, 11(18): 2091. DOI: 10.3390/rs11182091.
- [64] Yao Y, Wu D Q, Hong Y, et al. Analyzing the effects of rainfall on urban traffic-congestion bottlenecks. *IEEE Journal of*

- Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 504-512.
- [65] Robinson C, Hohman F, Dilkina B. A deep learning approach for population estimation from satellite imagery. Proceedings of the 1st ACM SIGSPATIAL Workshop on Geospatial Humanities, 2017: 47-54.
- [66] Liu Y J, Racah E, Prabhat, et al. Application of deep convolutional neural networks for detecting extreme weather in climate datasets. International Conference on Advances in Big Data Analytics and Data Processing, 2016.
- [67] Racah E, Beckham C, Maharaj T, et al. Extreme weather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017.
- [68] Yi Jiawei, Wang Nan, Qian Jiale, et al. Spatio-temporal responses of urban road traffic and human activities in an extreme rainfall event using big data. Acta Geographica Sinica, 2020, 75(3): 497-508. [易嘉伟, 王楠, 千家乐, 等. 基于大数据的极端暴雨事件下城市道路交通及人群活动时空响应. 地理学报, 2020, 75(3): 497-508.]
- [69] Yi J W, Du Y Y, Liang F Y, et al. Anomalies of dwellers' collective geotagged behaviors in response to rainstorms: A case study of eight cities in China using smartphone location data. Natural Hazards and Earth System Sciences, 2019, 19 (10): 2169-2182.

Modeling and analysis of geographic events supported by multi-source geographic big data

DU Yunyan^{1,2}, YI Jiawei^{1,2}, XUE Cunjin^{2,3}, QIAN Jiale^{1,2}, PEI Tao^{1,2}

(1. Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Aerospace Information Research Institute, CAS, Beijing 100101, China)

Abstract: Geographic events, as a basic construct in geographic process description, have become a core content of geographic information system (GIS). Due to the limitation of acquiring human activity data, GIS modeling and analysis of geographic events has long been focused on the event-induced changes of geospatial objects and the interaction between the objects. However, in recent years, with the explosive growth of location-based service data and the rapid development of quantitative depiction of human activities, the impact of geographic events on human activities and online social participation in geographic events have aroused wide concern in many fields, which poses great challenges to the space-time cognition, modeling methods and analysis framework of geographic events. In this regard, this study discussed the conceptualization and categorization of geographic events in the context of big data, and then introduced the space-time semantics and graph-based data model for geographic events. The "node-edge" graph data structure is used to establish event ontology, the secondary or cascading events, the evolution process, and the "cause-effect" interaction. The spatiotemporal data mining approaches for geographical events were also summarized, which are limited to conventional event detection in "physical space". Integrating "virtual space" event discovery and propagation simulation ideas into data mining approaches is essential for recognizing multi-scale spatiotemporal responses and understanding regional difference of human activities under diverse geographic events. Finally, the study used urban rainstorm events as an example to examine the conceptualization and modeling method of geographic events. Social responses to urban rainstorms and regional differences were examined at inter-urban and intra-urban scales. The case study proved the concept and verified the feasibility and practicability of the proposed framework.

Keywords: geographic event; graph model; geographic big data; human activity response