

王慕华, 王天岳, 李雁鹏, 等. 基于通用事件表示模型(SEM)的暴雨预警事件图谱研究[J]. 灾害学, 2021, 36(4): 74-78. [WANG Muhua, WANG Tianyue, LI Yanpeng, et al. Research on Knowledge Graph Model About Rainstorm Disaster Based on Simple Event Model[J]. Journal of Catastrophology, 2021, 36(4): 74-78. doi: 10.3969/j.issn.1000-811X.2021.04.013.]

基于通用事件表示模型(SEM)的暴雨预警事件图谱研究^{*}

王慕华, 王天岳, 李雁鹏, 崔磊

(中国气象局公共气象服务中心, 北京 100081)

摘要: 根据2018—2020年汛期的气象灾害预警信息发布统计分析, 暴雨预警事件占到全年气象灾害预警事件的21.3%。摸清事件关系, 构建事件图谱, 分析事件动态变化, 实现事件影响的智能推理, 可提高气象防灾减灾的效益。传统的事件模型大多以事件之间关系为切入点, 难以全面的表示暴雨灾害的事件要素及其内部联系。该文以通用事件表示模型(SEM)为基础, 结合暴雨的特点改造“5W1H”事件分析法为“4W2H”, 并结合知识图谱设计了结构层、要素层、数据层三层结构的图谱模型, 从事件梳理分析和事件要素关系构建的角度对暴雨预警事件进行建模, 构建了暴雨灾情-降雨实况-暴雨预警发布相结合的暴雨预警事件图谱模型。以暴雨实况、灾情和暴雨预警信号数据为驱动, 实现了在暴雨预警事件服务中的推理应用。

关键词: 暴雨灾害; 预警; 知识图谱; 事件图谱; 通用事件表示模型(SEM)

中图分类号: P642; X43; X915.5 **文献标志码:** A **文章编号:** 1000-811X(2021)04-0074-05

doi: 10.3969/j.issn.1000-811X.2021.04.013

根据2018—2020年汛期(5—9月)气象灾害预警信息发布统计分析, 暴雨预警占到总气象灾害预警的21.3%, 构建暴雨预警事件图谱, 实现暴雨预警事件从气象监测、预警信息发布、灾害发生的智能推理服务, 是有效减少灾害发生概率、减少灾害中人员伤亡、降低灾害损失的重要举措之一。暴雨预警事件图谱是以事件为核心概念的领域事件知识图谱(Domain-Specific Event Knowledge Graph, DEKG), 不同于知识图谱中的概念实体, 事件能够直接刻画实体及其行为, 具有更丰富的语义, 比如, 事件间的时态、因果以及顺承等关系能够更好地刻画实例层知识^[1]。

在我国, 每年雨季都有不同地区因暴雨导致不同程度地受灾, 各地学者也对暴雨灾害事件展开了多方面的研究^[2-3]。早期的GIS模型主要是基于灾害事件发生的地点信息, 并结合其在时间序列上的变化, 来描述事件的动态时空特征^[4-8]。但是在对多源异构数据进行综合分析的方面, GIS模型需要对每种关系分别构建数据集与关系集, 不仅存在大量的数据冗余, 还难以表达数据之间的关系。DBPedia与YAGO引入了知识图谱构建模式来构建事件模型, 很好的解决了数据冗余的问题, 但是其采用的模式层和数据层两层图结构模型难以将事件的每种要素信息结合起来^[9-10]。通用事件表示模型SEM(Simple Event Model)基于知识图

谱, 通过事件的时间地点要素描述事件的时空信息, 在新闻, 医疗, 历史事件等研究领域得到了广泛运用^[11]。

综上, 现有的基于知识图谱的事件模型可以描述事件多维度的特征, 为网络舆情、突发事件等领域事件的分析研究提供依据。但应用于气象暴雨预警事件分析中仍存在以下问题亟需解决: 在大数据时代, 信息的来源多种多样, 导致信息呈现方式具有数据多源且描述多样的特点, 现有模型不能完善表达事件的各级概念层次及其之间的联系。因此, 本文利用人工专家经验总结分析暴雨预警事件的过程及要素间的关系, 以SEM模型为基础, 结合暴雨预警事件的特征, 对“5W1H”分析法进行优化, 构建针对于暴雨预警事件的模型“4W2H”分析方法, 设计了以知识图谱为基础的暴雨预警事件图谱分析模型, 并以2020年7月5日湖北英山县暴雨为例, 展示分析暴雨预警事件的多个要素及各要素之间的联系, 以气象监测、预警发布、灾情数据为驱动, 实现了湖北英山县暴雨预警事件智能推理服务。

1 暴雨预警事件要素及特点

暴雨预警事件是一种气象防灾减灾事件, 其

^{*} 收稿日期: 2021-02-02

修回日期: 2021-04-25

基金项目: 国家重点研发计划项目(2018YFF0300105; 2018YFC1507805)

第一作者简介: 王慕华(1982-), 女, 汉族, 山东蓬莱人, 高级工程师, 主要从事GIS及遥感技术在公共气象服务领域内的应用、防灾减灾服务技术研究。E-mail: wangmuhua@cma.gov.cn

本身具有时间与空间特征和发展规律^[12]。对暴雨预警事件的分析从出现暴雨的气象实况监测, 到针对事件全过程的预警信息发布四级联动(国家、省、市、县四级预警信息发布), 以及预警发布后的应急响应联动和暴雨造成的灾害影响等方面。因此本文在传统的“5W1H”(何事(What)、何人(Who)、何时(When)、何地(Where)、何故(Why)和如何(How)六要素)分析法^[13]的基础上, 删去“何人(Who)”要素, 添加“程度(How Much)”要素, 构建适合于暴雨预警事件分析的“4W2H”分析法。

“4W2H”的具体含义如下。“What”指事件的本体, 表示对暴雨事件发生全过程的描述, 其与事件一一对应, 可以来自灾情数据库或对网络舆情信息的抽取。“When”指事件发生的时间或时间段, 包含事件的开始时刻, 结束时刻, 变化时刻以及持续时间, 能够反映出暴雨事件在时间维度上的影响范围。“Where”指事件发生的空间范围, 内容为以坐标点描述的精确地理坐标以及所属的行政区划, 反映暴雨事件在空间维度上的影响范围。“Why”是指暴雨事件的诱因, 通常为瞬时降雨量的激增或累计降雨量大等原因, 需要体现降雨实况与事件结果之间的因果逻辑。在暴雨预警事件中“Why”要素主要为实况降水量超警或持续累计降水量超警, 需要综合多个气象站点的监测数据进行表示。且结合地理位置的特点, 警戒指标线的划定会有很大区别, 需要具体融合人工专家知识进行区别分析。“How”指应对措施及手段, 主要为对于预警信息发布的及时性与准确性等方面的评估, 也包括事件发生后的应急响应联动, 是对气象防灾减灾业务进行评价的重要依据。“How Much”表示暴雨事件导致的影响, 包括经济损失、人口伤亡、农作物受灾、交通运输等方面的影响。其中“When”和“Where”要素是模型将事件与专业气象数据库中的结构化数据进行联系的重要纽带, 也是模型描述对象同一性的有效保障。“Why”和“How”是在暴雨预警事件分析过程中需要进行彼此对照、分析运算的重要指标, 是对灾害应对过程的重要评价依据, 也是发现预警发布存在的问题的有利线索, 与此同时也对“Why”要素和“How”要素提取的准确性提出了很高的要求。

暴雨预警事件“4W2H”六要素之间的关系错综复杂。比如, “Why”和“How”要素之间, 即降雨实况和预警信息发布的关系中, 既要考虑降雨的时间强度, 也要考虑降雨的空间范围, 可能是由多个气象站点的降雨量超过一定量级导致多地预警, 也有可能是单个气象站点的降雨强度过大, 实况超警事件引发某地的多次预警。因此, 暴雨预警事件的六要素之间存在着一对多和多对多关系, 需要以多个三元组共同表示的方式来进行知识描述。现有的研究定义了多种事件要素之间的关系, 如因果关系、顺承关系、并发关系、同步异步关系等^[14-15]。本文在此基础上加以扩展、细化, 构造出组成、预告、诱发、跟随、并

发关系, 且通过构建中间量的方式对无法直接表示的事件要素关系进行间接表示, 进而分析得出其之间的关系。

2 暴雨预警事件图谱的构建

2.1 建模思路与方法

知识图谱以节点为核心表示事件及事件的要素, 以边表示各种实体之间的联系, 基于< 实体, 关系, 实体>三元组和< 实体, 属性, 属性值>三元组, 将静态知识以图结构展示^[16]。以实体为中心的知识图谱聚焦于静态概念, 忽略了知识的动态性和过程性, 而暴雨预警事件图谱是以事件为核心概念的领域事件图谱, 考虑事件本身, 以及事件随时间、空间的动态发展。因此, 暴雨预警事件图谱的构建以灾害数据库和专业气象数据库为数据来源, 提取暴雨预警事件的组成要素集和要素之间的关系集, 并且以专家知识作为依据对关系集进行约束。再通过知识图谱对事件要素和关系进行建模, 构建事件六元组以及事件要素之间关联的三元组(图1)。

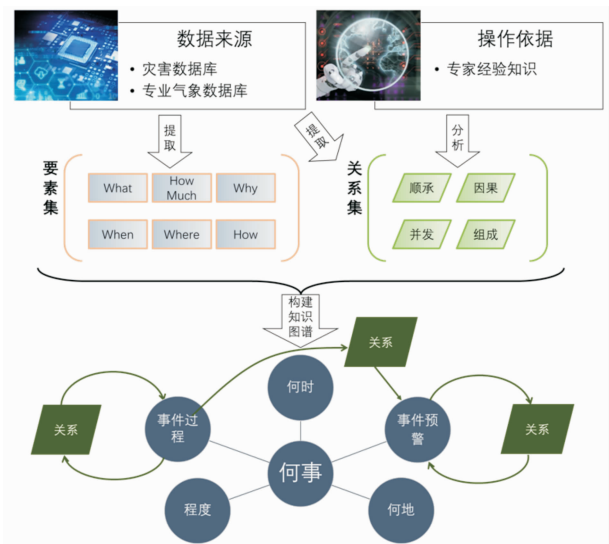


图1 暴雨预警事件图谱模型构建方法

暴雨预警事件图谱的关系模式以 SEM 通用事件表示模型为基础。该模型本身不需要以领域的特有词汇为依据, 而是以事件核心、事件分类以及属性约束描述不同领域的事件。其中事件核心(sem: Core)是事件的组成部分, 包含事件的几大基础要素, 如时间(sem: Time), 地点(sem: Place)等; 事件分类与事件核心一一对应, 将图谱中的事件组成部分进行分类, 但不在此基础上进行类别的扩展与细分; 属性约束构建知识图谱中结点的属性, 对事件各分类的关系进行约束, 实现事件的非层次关系表示。本文在 SEM 的基础上对暴雨预警事件的原因(暴雨实况过程)及应对措施(暴雨预警发布)两个方面进行扩展, 即 rdg: Process, rdg: Warning 与 rdg: ProcessType, rdg:

WarningType。并且在约束中加入专家经验的约束 (rdg: Expertise)。结合既有的时间关系, 空间关系对暴雨预警事件更为全面的过程描述与要素展示 (图2)。

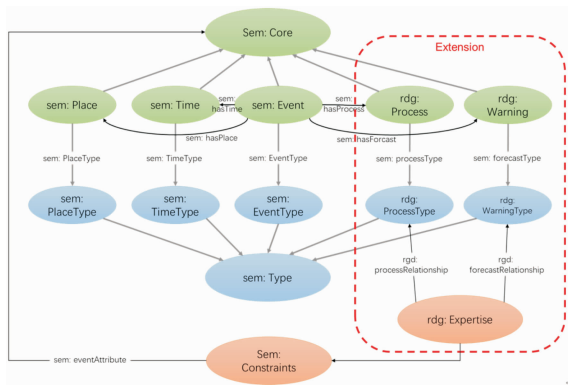


图2 基于SEM的暴雨预警事件模型

2.2 模型设计

本文结合暴雨预警事件的属性, 依据事件图谱的结构特性, 从实体和关系的表示角度出发, 将事件、要素和属性抽象为实体, 将事件和要素之间的关系抽象为边, 图谱中包含“要素-要素”关系, “事件-要素”关系, 并以此为依据构建面向事件结构、要素关系、具体数据的三层暴雨预警事件表示模型 (图3)。暴雨预警事件图谱 (G_{rd}) 包含三层结构: 结构层 (G_c)、要素层 (G_e)、数据层 (G_d), 以及数据层对应到要素层之间的关系 (R_d), 结构层与要素层的对应关系 (R_c), 所以暴雨预警事件图谱可以表示为以下五元组的结构:

$$G_{rd} = \langle G_c, G_e, G_d, R_d, R_c \rangle. \quad (1)$$

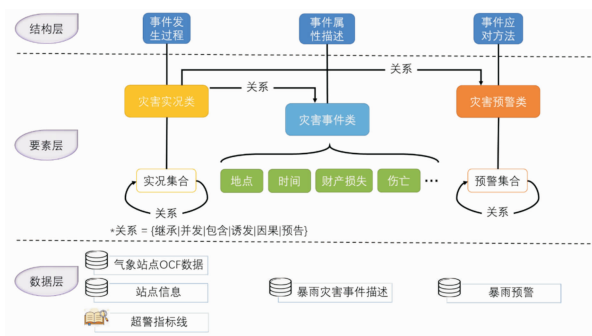


图3 暴雨预警事件模型层次

通过表示事件各种要素之间的关系, 形成要素之间的关系层次, 构建出结构层; 再以专家经验知识为约束, 根据不同要素的属性、属性值, 经过数据运算与知识判断, 得到事件要素之间的关系, 构建要素层; 最后再通过获取灾害数据库、专业气象数据库中的数据得到数据层, 并对要素层和结构层的节点进行数据扩充和关系补全。结构层 G_c 是表示一起暴雨预警事件组成部分的关系图谱, 对应基于 SEM 的暴雨预警事件模型中的 sem: Type 的所有要素; 要素层 G_e 表示暴雨预警事件要素组成及各要素之间的关系, 暴雨事件的

各要素之间关系分为组成、跟随、并发、因果四种关系, 因果关系具体体现为预警关系与诱发关系。关系的构建需要对要素之间进行关系判断, 分为专家经验知识判断和数据关系判断两步骤, 要素层对应 sem: Core 及其下属要素; 数据层 G_d 为结构层和要素层的各节点提供数据支持, 并为要素层提供数据作为数据关系判断的依据; 关系 R_d 为数据层到要素层的对应关系, 要素层通过数据层数据进行节点扩充, 并以要素层的专家经验约束与数据层提供的数据约束进行节点关系构建; 关系 R_c 为结构层与要素层元素的包含关系, 表示要素层对结构层的扩展和细化。

2.3 构建方法

(1) 结构层概念及对应要素。结构层是暴雨事件分析的核心, 对应暴雨分析的要素及其层次关系。暴雨预警事件包含如下几个方面: ①时间。暴雨发生的时间点和持续的时间段范围, 将事件在时间维度进行定位。②地点。包含事件发生地点的精确经纬度坐标以及所属的行政区划, 以平面坐标位置和行政区层次两个方式将事件在空间维度进行定位。③类型。参考中国暴雨的天气学研究进展中的分类标准^[17], 对暴雨事件的影响范围及时长等特征进行定性分析。④影响。灾后对灾害伤亡与经济损失的统计, 包含人员伤亡、农作物受灾、房屋损坏、直接经济损失、影响行业等多个方面。⑤应对方式。包含气象灾害预警发布过程以及应急响应过程。⑥事件过程。从事件开始到结束的过程中暴雨实况的变化过程。

(2) 数据层设计。数据层主要是为要素层和结构层提供数据支持。对于结构层和要素层而言, 数据层的主要作用为数据扩充和内容丰富; 对于要素层, 除上述作用, 数据层还起到作为连接判断依据的作用。数据层可以直接获取到的是结构化的数据, 需要进行数据处理, 如预警信息按照参考预警进行归组。

(3) 要素层构建方法。暴雨预警事件图谱构建的关键点在于构建事件要素之间的关系, 关系构建的全面性以及可靠性是决定图谱应用价值的关键因素。要素层为事件要素之间的关系构建, 构建过程如图4所示。

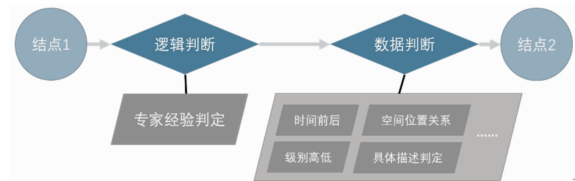


图4 节点关系构建过程

节点的关系判断需要经过逻辑判断与数据判断。逻辑判断是根据专家经验知识进行关系连通可行性的判断, 如暴雨实况的节点与暴雨灾害的节点之间因果关系的判断, 需要进行专家经验对数据判断进行约束; 数据判断则是根据节点中的数据数据进行具体的关系判断, 如上级发布单位发布

- Use of the Simple Event Model (SEM) [J]. *Journal of Web Semantics*, 2011, 9(2): 128–136.
- [2] 郭大梅, 潘留杰, 史月琴, 等. 西安地区一次罕见秋季冷锋后暴雨过程分析[J]. *高原气象*, 2020, 39(5): 986–996.
- [3] 张雅乐, 俞小鼎. 黄河气旋暴雨过程发展演变成因分析[J/OL]. *高原气象*, 2021, 40(1): 74–84.
- [4] 龚健雅. GIS 中面向对象时空数据模型[J]. *测绘学报*, 1997(4): 10–19.
- [5] 徐志红, 边馥苓, 陈江平. 基于事件语义的时态 GIS 模型[J]. *武汉大学学报(信息科学版)*, 2002(3): 311–315.
- [6] 吴长彬, 闫国年. 一种改进的基于事件-过程的时态模型研究[J]. *武汉大学学报(信息科学版)*, 2008(12): 1250–1253, 1278.
- [7] 龚健雅, 李小龙, 吴华意. 实时 GIS 时空数据模型[J]. *测绘学报*, 2014, 43(3): 226–232, 275.
- [8] 张军海, 李仁杰, 傅学庆, 等. *地理信息系统原理与实践* [M]. 北京: 科学出版社, 2009.
- [9] AUER S, BIZER C, KOBILAROV G, et al. Dbpedia: A Nucleus for a Web of Open Data [M]. *The Semantic Web*. Springer, Berlin, Heidelberg, 2007: 722–735.
- [10] HOFFART J, SUCHANEK F M, BERBERICH K, et al. YAGO2: A Spatially and Temporally Enhanced Knowledge Base from Wikipedia [J]. *Artificial Intelligence*, 2013, 194: 28–61.
- [11] 漆桂林, 高桓, 吴天星. 知识图谱 研究进展 [J]. *情报工程*, 2017, 3(1): 4–25.
- [12] 王艳东, 李昊, 王腾, 等. 基于社交媒体的突发事件应急信息挖掘与分析 [J]. *武汉大学学报(信息科学版)*, 2016, 41(3): 290–297.
- [13] 王伟, 赵东岩. 中文新闻事件本体建模与自动扩充 [J]. *计算机工程与科学*, 2012, 34(4): 171–176.
- [14] 杨雪蓉, 洪宇, 陈亚东, 等. 事件关系检测的任务体系概述 [J]. *中文信息学报*, 2015, 29(4): 25–32.
- [15] LI Z, ZHAO S, DING X, et al. EEG: Knowledge Base for Event Evolutionary Principles and Patterns [C]// *SOCIAL MEDIA PROCESSING*, SMP 2017, Book Series: Communications in Computer and Information Science, 2017: 40–52.
- [16] 陈浩, 李永强, 冯远静. 基于多关系循环事件的动态知识图谱推理 [J]. *模式识别与人工智能*, 2020, 33(4): 337–343.
- [17] 寿绍文. 中国暴雨的天气学研究进展 [J]. *暴雨灾害*, 2019, 38(5): 450–463.

Research on Knowledge Graph Model about Rainstorm Disaster Based on Simple Event Model

WANG Muhua, WANG Tianyue, LI Yanpeng and CUI Lei
(CMA Public Meteorological Service Center, Beijing 100081, China)

Abstract: According to the statistical analysis of the meteorological disaster early warning information issued during the flood season from 2018 to 2020, heavy rain early warning events accounted for 21.3% of the annual meteorological disaster early warning events. The benefits of meteorological disaster prevention and mitigation can be improved by understanding the relationship of events, constructing event map, analyzing the dynamic changes of events, and realizing intelligent reasoning of the impact of events. Most of the traditional event models take the relationship between events as the starting point, so it is difficult to comprehensively represent the event elements and their internal relations of the rainstorm disaster. In this paper, based on the Simple Event Model (SEM), combining with the characteristics of the heavy rain, change “5W1H” event analysis method for the “4W2H”. Based on the knowledge graph design a three layers structure graph model includes structure layer, element layer and data layer. The rainstorm early warning events model are modeled from the perspective of event carding analysis and relationship construction of event elements, includes the rainstorm disaster, rainfall reality and rainstorm warning release. Driven by rainstorm reality, disaster situation and rainstorm warning signal data, and the reasoning application in rainstorm warning event service is realized.

Key words: rainstorm disaster; early warning; knowledge graph; the event model; Simple Event Model (SEM)