



水利学报
Journal of Hydraulic Engineering
ISSN 0559-9350, CN 11-1882/TV

《水利学报》网络首发论文

题目：知识驱动的水利工程应急方案智能生成方法——以南水北调中线工程为例
作者：刘雪梅，卢汉康，李海瑞，槐先锋，陈晓璐
DOI：10.13243/j.cnki.slxb.20220994
收稿日期：2022-12-12
网络首发日期：2023-06-07
引用格式：刘雪梅，卢汉康，李海瑞，槐先锋，陈晓璐. 知识驱动的水利工程应急方案智能生成方法——以南水北调中线工程为例[J/OL]. 水利学报.
<https://doi.org/10.13243/j.cnki.slxb.20220994>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

文章编号：0559-9350(2023)00-0000-11

知识驱动的水利工程应急方案智能生成方法 —以南水北调中线工程为例

刘雪梅^{1,2}, 卢汉康¹, 李海瑞³, 槐先锋⁴, 陈晓璐⁴

(1. 华北水利水电大学 信息工程学院, 河南 郑州 450000;

2. 黄河流域水资源高效利用省部共建协同创新中心, 河南 郑州 450000;

3. 华北水利水电大学 管理与经济学院, 河南 郑州 450000;

4. 中国南水北调集团中线有限公司, 北京 100038)

摘要：水利工程传统应急方案存在数字化程度低、内容关联性差、智能辅助决策不足等问题。本文利用知识图谱和深度学习技术，创建一种水利工程应急方案智能生成模式。首先基于风险防控手册及险情抢险应急方案等文本，提出应急方案知识图谱本体模型，构建应急方案知识图谱，实现应急方案文本中非结构信息的结构化表达。其次，基于水利工程巡检文本，利用 BERT(Bi-directional Encoder Representation from Transformers)+BiLSTM+CRF(Bi-directional Long Short Term Memory with Conditional Random Fields)实体识别模型，智能识别巡检文本中的风险事件、工程等实体。最后，设计应急方案智能生成模板，通过多特征融合的实体对齐技术、知识检索与推理技术，实现应急方案的智能生成与推送。通过模型准确性分析以及“渠道渗漏”等实例验证，本文方法识别准确率高($F1$ 值为 96.21%)，生成的应急方案可靠，可推广到水利工程应急抢险以及应急预案智能生成等应急管理工作。

关键词：知识图谱；水利工程；应急方案；南水北调

中图分类号：X93；TV513

文献标识码：A

doi：10.13243/j.cnki.slxb.20220994

1 研究背景

南水北调中线工程线路长，沿途环境复杂多变，工程运行过程中存在诸多风险。针对各类突发险情，敏捷有效的应急响应颇为重要。目前，南水北调应急方案大多以纸质文本、电子文档形式存储，存在查询检索效率低、内容关联弱、智能辅助决策不足等问题。因此，探索科学高效的应急方案智能生成方法十分必要。

目前应急方案智能生成主要采用案例推理技术，通过匹配目标案例与历史案例的相似性，复用或修改相似度最高的历史案例应急方案来解决突发险情^[1]，封超^[2]、王兴鹏等^[3]考虑案例特征属性，通过计算案例相似度生成应急方案；姚鑫^[4]、Zheng^[5]、Yu^[6]、Zhang 等^[7]分别采用剪枝算法、融合主观偏好与客观信息的动态检索算法、设计差分进化算法等来优化案例检索，提高应急方案的准确性与实效性。上述研究主要依赖于历史案例库规模以及案例属性的筛选与计算，生成的应急方案以文档呈现，数字化程度低，知识关联性弱。

知识图谱^[8]于 2012 年诞生，最初被应用于语义搜索^[9]、智能问答^[10]、智能推荐^[11]等，实现海量知识的快速获取、合理组织、科学利用。近年来，知识图谱亦被应用于应急管理领域，Li 等^[12]利用

收稿日期：2022-12-12

基金项目：河南省科学院科技开放合作项目(220901008)，国家自然科学基金项目(72271091)

作者简介：刘雪梅(1965-)，教授，主要从事水利信息化研究。E-mail: 1354879559@qq.com

混合知识表示方法,集成非结构化、半结构化、结构化知识,开发应急知识系统,提供自然灾害应急辅助决策;Liu 等^[13]提出地质灾害应急响应的知识概念模型,在突发事件发生时能够智能获取并提供所需要的应急知识,实现快速应急响应;Ni 等^[14]构建应急预案知识系统,为应急响应提供可靠信息;Yang 等^[15]构建知识共现网络,分析了突发公共卫生事件应急管理的重要性。知识图谱作为一种语义网络拥有极强的知识组织与表达能力,目前主要应用于自然灾害、地质灾害、公共卫生等特定应急管理领域,虽然在水利领域已有初步应用^[16-17],但在水利应急管理中尚未应用。

为提高水利工程应急处置与管理的智能化水平,本文创建一种水利工程应急方案智能生成模式。以南水北调中线工程为例,提出应急方案知识图谱本体模型,设计应急方案模板,结合不同水利工程文本特点,采用与之适应的知识图谱和机器学习技术,构建南水北调中线工程应急方案领域知识图谱,结合图谱检索与推理技术实现应急方案的智能生成。

2 相关技术方法

2.1 知识图谱构建过程 知识图谱构建过程一般包括知识建模^[18]、抽取^[19-20]、融合^[21]、存储^[22-23]四个阶段。知识建模即建立知识图谱的概念模式,是知识图谱的核心,主要有自顶向下和自底向上两种构建方式。知识抽取是从数据源中抽取实体、关系、属性。知识融合是整合不同数据源中的信息,得到规范统一的描述。知识存储将融合的实体、关系、属性存入图数据库,便于下游应用。

2.2 基于 BERT+BiLSTM+CRF 的实体识别模型 实体识别方法主要包含基于词典规则的方法^[24]、传统机器学习方法^[25]以及基于深度学习的方法^[26]。基于词典规则的实体识别模型识别速度快、准确率高,但受限于词典规则、通用性不强,适用于小规模并且形式单一的数据,并且无法解决多词同义问题。传统的单一机器学习模型准确率较低,难以在水利工程领域实际应用。本文基于 2019 年 Devlin 提出的深度学习模型 BERT^[27],BERT 模型预训练于 Google 的大规模通用中文语料库,通用语义信息获取能力强,但水利领域语义信息获取效果差,而 BiLSTM+CRF 模型只需在较小量级的水利标注数据内训练,即可获得水利领域语义信息。本文将两者结合,构建 BERT+BiLSTM+CRF 模型识别险情文本实体,通过微调训练,既能有效避免大量标注数据,又能获得较高的实体识别准确率。

2.3 实体对齐技术 实体对齐一般通过计算两个实体的相似度来完成。传统实体相似度计算主要基于 Jaccard 算法^[28],准确率高但召回率低,并且词语缺乏语义信息。目前常用的实体相似度计算方法有三种:其一是基于词典或某种分类体系,常见词典有 Hownet、Wordnet 和同义词词林,这三种词典的构造方法互不相同,但都局限于通用领域,难以解决水利巡检文本面临的专业领域性强以及大量同义词现象;其二为基于上下文向量空间的统计方法,以 Google 的 Word2Vec 词语向量化工具为代表,将词语映射到空间向量,通过计算向量距离来衡量相似度;最后一种是基于深度学习的方法,但其所需语料库庞大,计算量与成本过高。本文设计基于 Word2Vec 模型^[29]的 Jaccard 实体相似度算法,融合巡检文本中的多类实体特征,实现实体对齐。

2.4 图谱推理技术 图谱推理中的规则推理亦称为本体逻辑的演绎推理,包含建立规则与做出推理两部分,其可解释性强并且推理精度高。例如建立二跳推理规则〈风险事件,对应控制措施,控制措施〉〈控制措施,包含设备,设备〉推出为〈风险事件,所需抢险设备,设备〉并写入 cql 查询语句中,即可实现由图谱中存在的两个三元组〈管涌,对应控制措施,在涌水口采用……〉〈在涌水口采用……,包含设备,编织袋砂石等……〉,推出一个未知三元组〈管涌,所需抢险设备,编织袋砂石等……〉。

3 南水北调中线工程应急方案知识图谱构建

南水北调应急方案均为非结构化和半结构化文本,知识关联性弱,查询检索效率低,难以智能化应用。为解决这些问题,需要构建南水北调中线工程应急方案领域知识图谱。领域知识图谱构建流程如图 1。

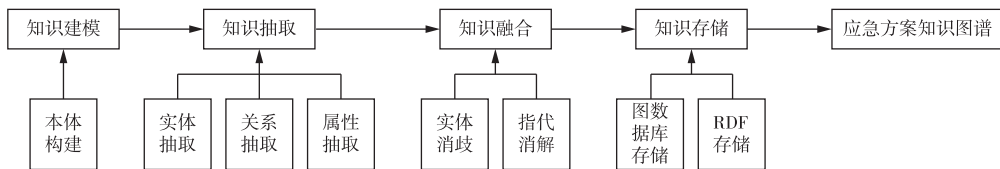


图 1 应急预案知识图谱构建流程

3.1 知识建模 知识建模即为本体建模，本体模型是知识图谱的骨架与核心。通过南水北调中线工程应急方案本体建模，可以将应急方案以结构化知识的方式进行组织与管理，为应急方案知识图谱构建提供底层数据模式，同时也是应急方案智能生成的基础。本文以南水北调中线工程 47 个管理处的风险防控手册以及一期工程总干渠工程险情抢险体系总体方案设计报告为数据源，结合自顶向下与自底向上定义知识图谱本体以及本体间的关系，构建南水北调中线工程应急方案知识图谱本体模型，如图 2。

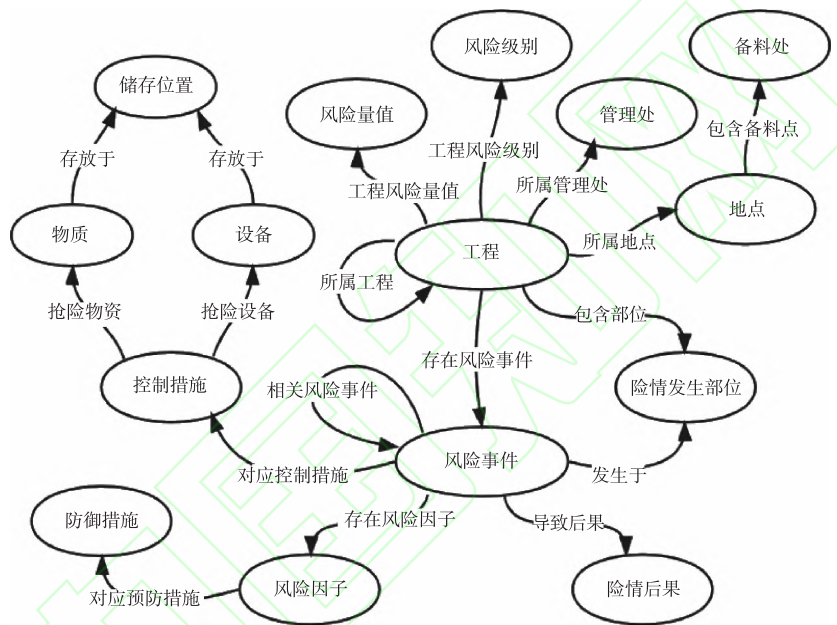


图 2 南水北调中线工程应急方案知识图谱本体模型

3.2 知识抽取 根据领域资料不同特点，采取不同的方式实现知识抽取。考虑到应急方案领域资料大多以半结构化表格为主，非结构化数据包含的实体关系稀少且分散，难于构建有效的学习样本，而基于统计模型以及基于深度学习的方法需要大量标注数据，因此本文主要采用基于规则的方法完成知识抽取。

3.2.1 风险防控管理手册知识抽取 由于风险防控管理手册数据主要由半结构化数据构成，非结构化数据包含的实例数据稀少，因此选用简单高效的人工提取结合正则模板方式实现知识抽取。部分实例如表 1。

表 1 风险防控管理手册部分实体实例

实体类别	实体实例
工程	易县高填方渠道、釜山隧洞、荆轲山分水口、东楼山西沟排水倒虹吸
风险事件	管身冲刷失稳、过流能力减小、洪水浸泡渠坡、上部渠基破坏
地点	邓州市、陶岔县、南阳市、辉县
管理处	大宁管理处、西四环管理处、卫辉管理处、镇平管理处

3.2.2 险情抢险体系总体方案知识抽取 险情抢险体系总体方案主要为非结构化数据，但其具备清晰的文本结构，且大多为短文本数据。因此利用“关键词+短文本”的方式抽取实体关系，部分实例如表 2。

表 2 险情抢险体系总体方案部分实体实例

实体类别(关键词)	实体实例(短文本)
设备	反铲、人工辅助作业工具；螺旋钻机、注浆机、人工辅助作业工具
物资	塑料防水膜、水泥、天然砂石料；塑料防水膜、钢筋、水泥、天然砂石料
存储位置	塑料防水膜存放在管理处抢险物资仓库，水泥、天然砂石料当地采购
险情后果	严重变形形成滑坡隐患；小规模滑坡形成溃口隐患；大规模滑坡导致渠道堵塞
险情发生部位	渠坡坡顶；渠坡中部；一级马道附近；建筑物外侧填土坡

3.3 知识融合 应急方案领域资料主要存在实例共指现象，在图谱构建过程中需要进行共指消解，从而得到规范统一的描述。考虑到融合的实例大多以词语形式出现，并且在抽取的知识中基本包含所需映射的实例，因此本文基于链接的方式实现实例的共指消解。南水北调中线工程应急方案知识图谱属于垂直领域知识图谱，数据源稀缺，本图谱仅考虑对工程、风险事件、风险因子三种实体进行实体融合。以“刁河渡槽地基渗漏失稳”为例，抽取所得风险事件实体“渗漏”，经过上下文实体指代扩充成“地基渗漏”，通过候选实体生成得到一组实体“集中渗漏、渠道渗漏、管涌等”，最后通过候选实体排序，融合等价于图谱中的风险事件实体“集中渗漏”。

3.4 知识存储 考虑到南水北调中线工程应急方案知识图谱数据规模较小，下游应急方案生成任务对关系表达要求较高，并且图数据库对比关系型数据库具有遍历效率较高、关系表达能力强等特点，因此本研究采用实体关系查询效率高的 Neo4j 原生图数据库来实现知识存储。将知识抽取、知识融合得到的实体关系整理成三元组格式〈实体 1，关系，实体 2〉并存储，如〈管涌，存在风险因子，蚂蚁鼠洞危害〉〈张庄大桥，风险量值，4.0〉等。

4 南水北调中线工程应急方案智能生成

应急方案本质为一种知识的组织形式，其生成离不开知识图谱的结构化知识。基于南水北调中线工程应急方案知识图谱，提出面向工程巡检文本的应急方案智能生成方法，主要流程如图 3。首先对巡检文本进行实体识别，识别出险情相关的关键实体，如工程、风险事件、险情发生部位等；其次进行实体对齐，将识别的风险事件实体与知识图谱中的风险事件进行匹配；最后基于设计的应急方案模板，利用图谱检索推理技术^[30-31]智能生成应急方案。

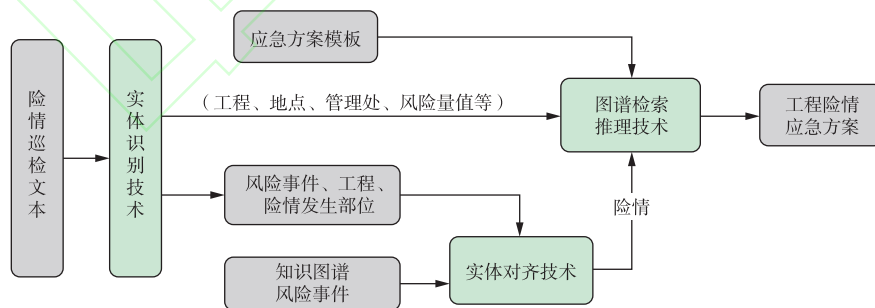


图 3 南水北调中线工程应急方案智能生成流程

4.1 险情相关实体识别 由于巡检人员大多为非专业人士，缺乏水利领域专业知识，导致每个巡检人员提交的巡检文本存在较大差异，对同一个风险事件也会有不同描述，因此，巡检文本中会存在大量多词同义现象。因此首先要对巡检文本进行语义分析，识别文本中的风险事件、工程等关键险情相关实体。考虑到水利专业领域词典稀缺以及应急方案标注数据不足，基于深度学习方法，构建 BERT+BiLSTM+CRF 模型识别险情文本实体，其具体结构如图 4。以输入险情信息“邓州高填方渠道衬砌板裂开”为例，模型最初通过编码层将文本转化为机器可识别的数字编码列表；传入 BERT 层，抽取巡

检文本中的通用语义特征并生成字向量；通过 BiLSTM 层使字向量携带巡检文本上下文语义信息；利用全连接层将字向量转换成水利实体标签；最后利用条件随机场对水利标签进行修正，输出最优的水利实体标签序列。

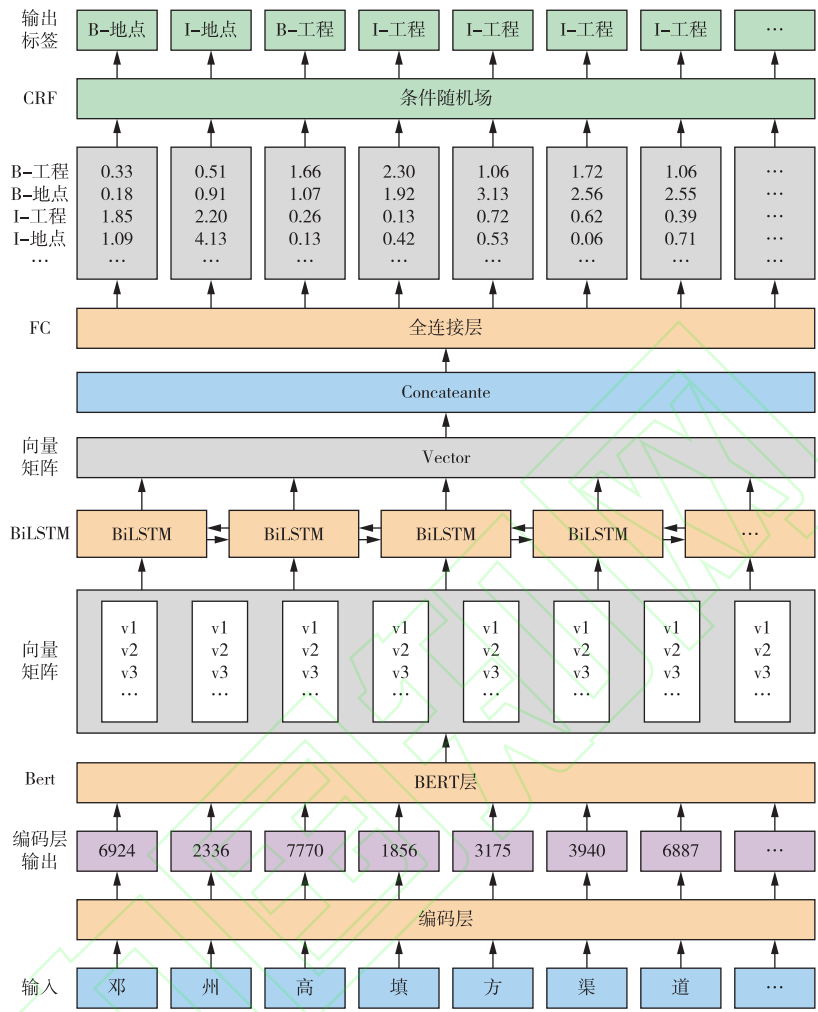


图 4 结合险情信息的 BERT+BiLSTM+CRF 模型架构

4.2 多特征融合的实体对齐 从水利工程险情巡检文本中抽取的风险事件往往不完全相同于图谱中的风险事件，存在缩略词、同义词等现象，例如“恐怖袭击”对应于缩略词“恐袭”，“排水孔淤堵”对应于同义词“排水孔堵塞”等。考虑到不同工程、不同险情发生部位的风险事件对应控制措施的不一致性，设计基于 Word2Vec 模型的 Jaccard 实体相似度算法，融合巡检文本中工程、险情发生部位、风险事件三类实体特征，实现巡检文本风险事件对齐于图谱风险事件。其中部分风险事件实体对齐情况如表 3，通过输入巡检文本中的风险事件，对齐图谱中匹配度最高的 5 个风险事件。

表 3 部分图谱风险事件实体匹配情况

风险事件	匹配风险事件 1	匹配风险事件 2	匹配风险事件 3	匹配风险事件 4	匹配风险事件 5
恐袭	恐怖袭击	无人机袭击	恐怖劫持事件	控制系统故障	排水渡槽漫溢
泥石流淤塞渠道	渠道堵塞	灌渠淤积	管道淤堵	下穿管道淤堵	阀井淤泥堵塞
大规模滑坡堵塞	挖方渠道堵塞	渠道边坡塌滑	渠道边坡开裂	边坡开裂滑塌钢管渗漏	堤坡塌滑失稳
渠坡防渗体系破坏	下穿管道渗漏	渠堤一级边坡渗漏	裹头边坡渗漏	钢管渗漏	渠坡冲刷破坏
坡面保护措施损坏	护坡裂缝破坏	运维养护施工不当	渠坡冲刷破坏	渠堤渗流破坏	附属设施破坏

从表 3 得知，输入风险事件实体为“恐袭”时，该算法匹配正确的风险事件实体为“恐怖袭击”，并且返回的其他风险事件实体与恐怖袭击相关，例如“无人机袭击”、“恐怖劫持事件”（恐怖袭击的表现形式）、“控制系统故障”（可能是控制系统被恐怖袭击破坏的结果）。

4.3 应急方案模板 为智能生成应急方案，需要设计应急方案模板。本文设计的应急方案模板分为工程概况、险情分析、抢险方案、备料点四部分，其中工程概况包括所属工程、所在地点、风险量值、风险级别；险情分析包括诱发因子、相关风险事件、导致后果、发生部位；抢险方案包含设备、物资、存储位置、抢险措施；备料点包含备料点编号、总干渠桩号、长度、面积等，具体应急方案模板如图 5。

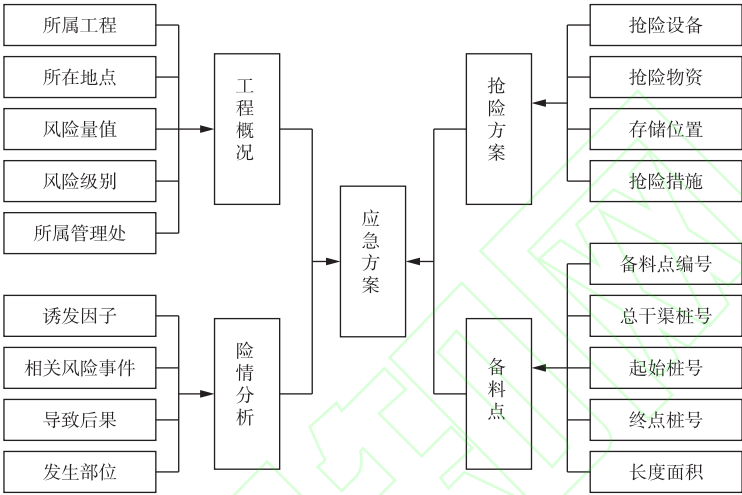


图 5 应急方案模板

5 实验结果分析

5.1 实验环境 本文实验运行环境为 CPU (Intel (R) Core (TM) i7-8700)、GPU (NVIDIA GeForce GT 710)、python 版本 3.6、操作平台 pycharm2020.3、神经网络框架 Keras，其中 BERT+BiLSTM+CRF 模型的部分参数如表 4。

5.2 图谱构建结果分析

5.2.1 图谱实体关系抽取结果 以 47 本南水北调中线工程风险防控手册以及抢险体系总体方案为数据源，抽取实体类型包含工程、地点、管理机构、风险量值、风险事件等共 15 种（具体实体统计见表 5），抽取关系类型包含工程-风险事件（存在风险事件）、工程-风险量值（工程风险量值）、工程-风险级别（工程风险级别）等共 17 种，具体关系统计见表 6。

表 4 BERT+BiLSTM+CRF 模型的部分参数

实验参数	含义	训练参数值
Epoch	训练轮数	50
Max_len	输入句子最大长度	128
Batch_size	每一次训练句子数	16
Lstm_units	BiLSTM 单层神经元数	128
Drop_rate	随机丢失率	0.1
Learning_rate	训练初始学习率	5×10^{-5}
Optimizer	优化器	Adam

表 5 南水北调中线工程抽取实体统计

实体类型	实体数量	实体类型	实体数量
工程	3353	管理处	47
风险事件	411	设备	35
风险因子	222	物资	25
风险量值	95	存储位置	6
预防措施	640	险情后果	101
控制措施	680	险情发生部位	102
地点	47	备料点	188
风险级别	4		

5.2.2 图谱结果展示 由于南水北调中线工程应急方案知识图谱实体关系数量众多，本文以部分实体关系进行图谱展示，如图 6。图中包含信息：刁河渡槽地处邓州市，其所属工程为渠系建筑物，所属管理处为邓州管理处；该工程包含风险事件地基渗漏失稳，地基渗流失稳属于建筑物地基失稳，其又包含建筑物地基承载能力不足以及集中渗漏导致地基水土流失两类子风险事件；各类子风险事件对应各类风险因子，风险事件、风险因子分别对应控制措施、预防措施；其中控制措施中涉及抢险设备、抢险物资两大类实体，不同风险事件与不同险情发生部位对应不同控制措施。由此将各类工程、风险事件、控制措施、抢险设备、抢险物资等实体相关联。

5.3 实体识别模型结果分析 该实体识别模型主要识别巡查文本中的工程、桩号、地点、管理处、风险量值、风险事件、风险因子、险情发生部位共 8 类实体。通过人工标注数据，切分成独立不同的训练集、验证集、测试集，基于模板利用数据增强方法扩充训练集和验证集，进而输入 BERT+BiLSTM+CRF 模型进行训练，模型训练结果如表 7。

表 6 南水北调中线工程抽取关系统计

实体类型 1	关系类型	实体类型 2	关系数量
工程	存在风险事件	风险事件	13162
工程	工程风险量值	风险量值	3567
工程	工程风险级别	风险等别	3567
工程	所在地点	地点	3599
地点	包含备料点	备料点	188
工程	所属工程	工程	3600
工程	所属管理处	管理处	3599
风险事件	存在风险因子	风险因子	8718
风险事件	导致后果	险情后果	411
风险事件	相关风险事件	风险事件	564
风险事件	发生于	险情发生部位	102
工程	包含部位	险情发生部位	102
风险事件	对应控制措施	控制措施	2918
控制措施	抢险设备	设备	92
控制措施	抢险物资	物资	92
设备	存放于	存储位置	92
物资	存放于	存储位置	92
风险因子	对应预防措施	预防措施	8487

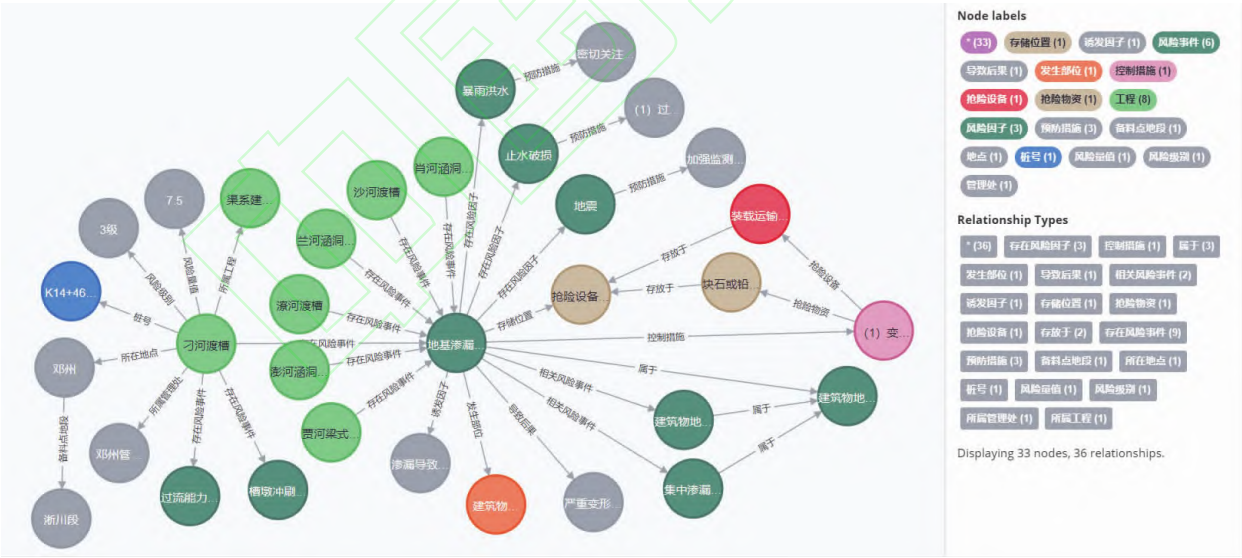


图 6 部分实体关系知识图谱展示

结果表明，该模型的 $F1$ 值达 96.21%，能够准确识别风险事件及其相关实体，并且该模型在以后图谱更新阶段，可以用于知识抽取，具有较高价值。

5.4 实体相似度算法结果分析 本文设计基于 Word2Vec 模型的 Jaccard 实体相似度算法，手动构建风险事件相似度算法测试集，对比了单一 Word2vec 模型与单一 Jaccard 算法在此测试集上的效果。三组算法评测结果见表 8。

实验结果表明，融合 Word2Vec 词向量的 Jaccard 实体相似度算法结果要优于单一 Word2vec 模型以及单一 Jaccard 算法。

5.5 应急方案智能生成结果分析 基于 4.2 节匹配的图谱风险事件实体和 4.1 节识别的工程、地点、险情发生部位等实体，结合应急方案中的各个要素，利用图谱推理技术中的规则推理以及图谱检索技术，实现应急方案智能生成。

以“邓州管理处高填方渠道桩号 K15+125—K16+140 处渠堤堤身发生渗漏现象”险情巡检文本为例，输入实体识别模型，返回标签以及险情相关实体：管理处（邓州管理处）、工程（高填方渠道）、桩号（K15+125—K16+140）、险情发生部位（渠堤堤身），以及风险事件（渗漏现象）。根据 Jaccard 算法粗粒度排序生成 20 个图谱风险事件候选实体，如表 9。

结合 Jaccard 算法和 Word2Vec 模型融合工程、风险事件、险情发生部位 3 个实体特征，从图谱中匹配相似度最高的前 5 个风险事件实体。利用图谱检索技术，返回风险事件匹配实体以及相应处置措施，如表 10。

由上述可知识别的风险事件实体“渗漏现象”匹配图谱风险事件实体为“渠道渗漏”。结合识别的实体列表以及应急处置方案模板，利用图谱推理技术与图谱检索技术智能生成险情应急处置方案，返回方案如图 7。该“渠道渗漏”应急方案包括险情相关信息识别结果、险情相关演练及处置事件、图谱风险事件匹配列表、风险事件预防控制手段、图谱概览、工程概况、抢险方案、险情分析、备料点 9 大板块：图谱概览展示风险事件“渠道渗漏”相关的知识；工程概况展示工程实体“邓州高填方渠道”的相关信息；抢险方案包含抢险需要的物资与设备；险情分析阐述“渠道渗漏”的成因、导致后果、相关风险事件以及险情发生部位等内容。

表 7 基于巡查文本的 BERT+BiLSTM+CRF 模型训练结果

标签	准确率	召回率	F1 值
工程	97.68%	97.56%	97.62%
风险事件	99.46%	99.56%	99.51%
风险因子	93.53%	93.53%	93.21%
险情发生部位	92.78%	92.96%	93.67%
桩号	93.84%	93.25%	93.54%
管理处	97.63%	97.88%	97.75%
地点	94.49%	91.90%	93.18%
风险量值	100%	100%	100%
平均值	96.26%	96.15%	96.21%

表 8 三组算法评测结果

算法	准确率	召回率	F1 值
Word2Vec	86.67%	69.53%	77.16%
Jaccard	88.67%	73.43%	80.33%
Word2Vec+Jaccard	90.33%	77.34%	83.32%

表 9 风险事件候选实体列表

涵身渗漏	结构渗漏	控制闸渗漏水	调压塔渗漏水
箱涵渗漏	集中渗漏	调压井渗漏水	地基渗漏失稳
接缝渗漏	箱涵渗漏水	隧洞衬砌渗漏	管桥渗漏
止水渗漏	结构缝渗漏	下穿箱涵渗漏	穿渠建筑物渗漏
渠道渗漏	明槽渗漏水	保水堰渗漏水	集中渗漏、流土

表 10 风险事件匹配实体以及相应处置措施

匹配实体	匹配概率	处置措施
渠道渗漏	0.7708	(1)对填方渠道，在渗漏出口设置压浸平台，防止水土流失；(2)必要时采用小型围堰进行水下浇筑...
集中渗漏	0.7618	(1)在集中渗漏出口设置压浸平台，防止水土流失；(2)迅速查明渗漏通道；(3)靠近渗漏通道入口...
接缝渗漏	0.7450	(1)在渠坡渗漏出口设置压浸平台，防止水土流失；(2)必要时对结构缝进行临时灌浆处理；(3)在...
止水渗漏	0.7418	在冬季无水情况下，或水量不大时进口临时封堵，采用左岸截流沟将水流疏导至其他排水建筑物...
集中渗漏、流土	0.7269	(1)检查结构缝止水，如止水破损，可采取更换止水，或者对结构缝缝面采用柔性材料进行加固，或...

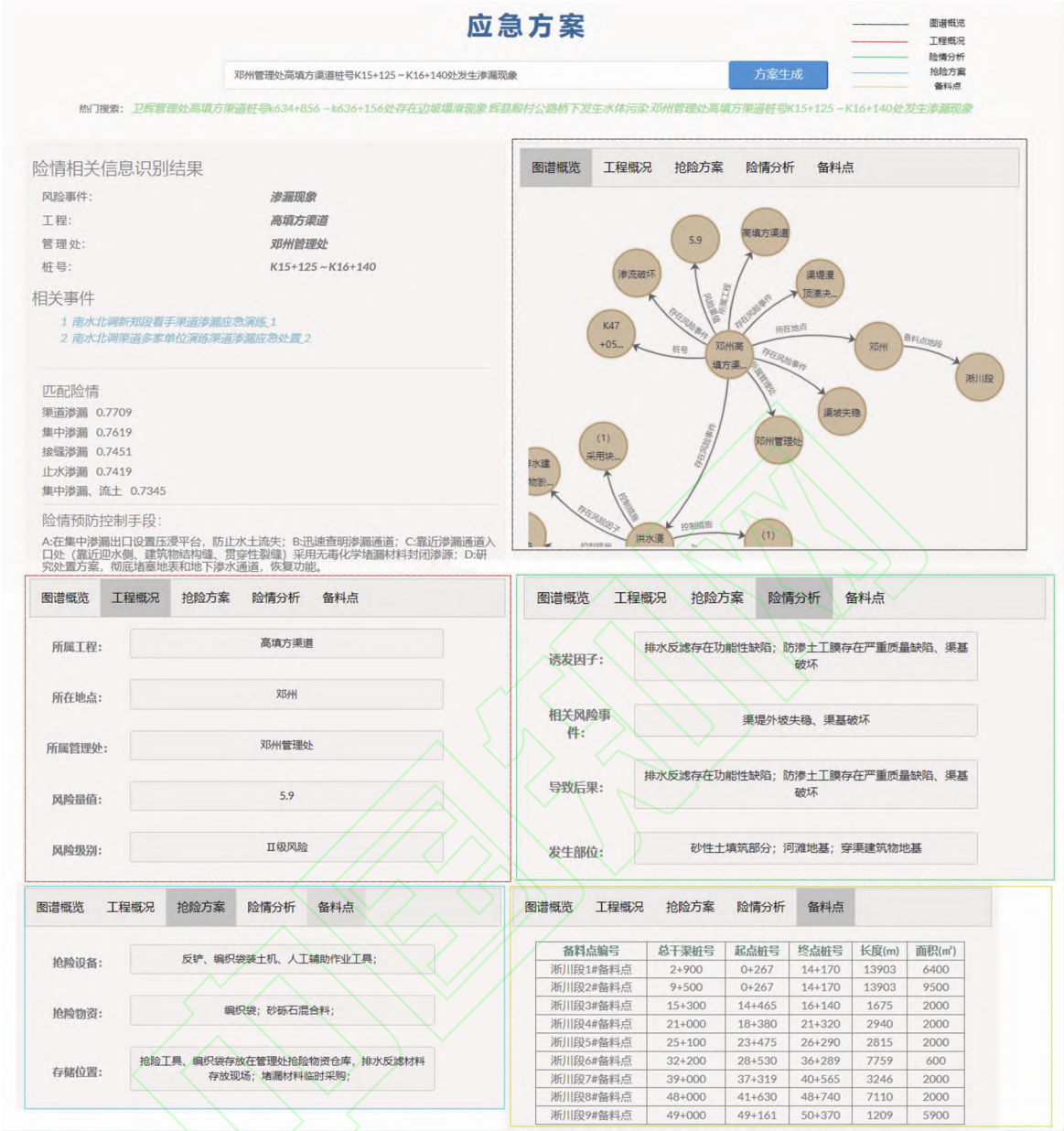


图 7 智能生成应急方案

6 结论与展望

本文提出基于知识图谱的应急方案智能生成方法，总结如下：

(1) 基于 47 本南水北调中线工程风险防控手册以及 1 本抢险体系总体方案，对风险事件、工程等知识组成及关系进行梳理分析，提出了应急方案知识图谱本体模型，合理组织与管理应急方案各类知识。利用知识抽取、融合、存储技术，构建了应急方案知识图谱，为应急方案智能生成奠定了基础，同时提供了险情相关知识查询与检索服务，解决了水利工程传统应急方案知识关联性差、查询检索不便问题。

(2) 设计了基于 BERT+BiLSTM+CRF 的实体识别模型，实现了基于小数据集的险情相关实体的有效识别。提出了基于 Word2Vec 模型的 Jaccard 实体相似度算法，实现了风险事件实体的智能对齐，解决了传统检索分词方法在水利工程文本专有名词识别与对齐上存在的识别准确率低、同义词无法对齐等问题。

(3) 研究设计了应急方案智能生成模板，利用图谱推理技术与检索技术，实现了应急方案的智能生成，有效辅助专业人员进行应急处置决策。

(4) 本文成果为南水北调中线工程智能化管理提供了一种有效手段, 并可推广至其他水利工程。

本研究实现了南水北调中线工程险情应急方案的智能生成, 但在知识融合、更新方面, 还存有许多值得探索的问题, 例如长文本融合、图谱动态更新等。在后续工作中, 需要进一步完善实体识别模型、优化实体对齐算法。需要进一步研究知识融合、知识更新技术, 扩大知识图谱广度与深度, 进一步服务于南水北调中线工程, 提升智慧化管理水平。

参 考 文 献:

- [1] GILBOA I, SCHMEIDLER D. Case-based decision theory[J]. Quarterly Journal of Economics, 1995, 110(3): 605-639.
- [2] 封超, 杨乃定, 桂维民, 等. 基于案例推理的突发事件应急方案生成方法[J]. 控制与决策, 2016(8): 1526-1530.
- [3] 王兴鹏, 雷书华, 刘文梅. 基于 CBR 的地铁工程事故应急决策方法研究[J]. 铁道工程学报, 2018(9): 104-109.
- [4] 姚鑫, 郭海湘, 顾明赞, 等. 基于案例推理的滑坡灾害应急相似案例智能生成研究[J]. 系统工程理论与实践, 2021(6): 1570-1584.
- [5] ZHENG J, WANG Y M, CHEN S Q. Dynamic case retrieval method with subjective preferences and objective information for emergency decision making[J]. IEEE/CAA Journal of Automatica Sinica, 2018, 5(3): 749-757.
- [6] YU X B, LI C L, ZHAO W X, et al. A novel case adaptation method based on differential evolution algorithm for disaster emergency[J]. Applied Soft Computing, 2020, 92: 106306.
- [7] ZHANG K K, LUO N X, LI Y B. STGA - CBR: A case - based reasoning method based on spatiotemporal trajectory similarity assessment[J]. IEEE Access, 2020, 8: 22378-22385.
- [8] 徐增林, 盛泳潘, 贺丽荣, 等. 知识图谱技术综述[J]. 电子科技大学学报, 2016(4): 589-606.
- [9] DONG X, GABRILOVICH E, HEITZ G, et al. Knowledge vault: a Web-scale approach to probabilistic knowledge fusion[C]// Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014.
- [10] HAO Y C, ZHANG Y Z, LIU K, et al. An end-to-end model for question answering over knowledge base with cross-attention combining global knowledge[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017.
- [11] GONG F, WANG M, WANG H F, et al. SMR: medical knowledge graph embedding for safe medicine recommendation[J]. Big Data Research, 2021, 23: 100174.
- [12] LI B, LI T T, JIANG Q, et al. A Knowledge-based system for disaster emergency relief[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2020, 34(11): 2059038.
- [13] LIU X H, LIU Z, TIAN J K, et al. Integration of a geo-ontology-based knowledge model and spatial analysis into emergency response for geologic hazards[J]. Natural Hazards, 2021, 108(2): 1489-1514.
- [14] NI Z J, RONG L L, WANG N, et al. Knowledge model for emergency response based on contingency planning system of China[J]. International Journal of Information Management, 2019, 46: 10-22.
- [15] YANG R, DU G M, MIAO X, et al. Knowledge system analysis on emergency management of public health emergencies[J]. Sustainability, 2020, 12(11): 1-18.
- [16] 段浩, 韩昆, 赵红莉, 等. 水利综合知识图谱构建研究[J]. 水利学报, 2021, 52(8): 948-958.
- [17] 毛文山, 赵红莉, 蒋云钟, 等. 基于文献计量学的国内水生态环境研究知识图谱构建与应用[J]. 水利学报, 2019, 50(11): 1400-1416.
- [18] AYACHI R, GUILLON D, ALDANONDO M, et al. Risk knowledge modeling for offer definition in customer-supplier relationships in Engineer-To-Order situations[J]. Computers in Industry, 2022, 138: 103608.
- [19] Al-MOSLMI T, OCANA M G, OPDAHL A, et al. Named entity extraction for knowledge graphs: a literature overview[J]. IEEE Access, 2020, 8: 32862-32881.
- [20] SMIRNOVA A, CUDRE-MAUROUX P. Relation extraction using distant supervision: a survey[J]. ACM Computing Surveys, 2019, 51(5): 1-35.

- [21] ZHAO X J, JIA Y, LI A P, et al. Multi-source knowledge fusion: a survey[J]. World Wide Web-Internet and Web Information Systems, 2020, 23(4): 2567-2592.
- [22] ZOU L, ZSU M T. Graph-based RDF data management[J]. Data Science and Engineering, 2017, 2: 56-70.
- [23] WYLOT M, HAUSWIRTH M, CUDRE-MAUROUX P, et al. RDF data storage and Query Processing Schemes: A survey[J]. ACM Computing Surveys, 2019, 51(4): 1-36.
- [24] RAU L F. Extracting company names from text[C]// Proceedings of the 7th IEEE Conference on Artificial Intelligence Applications, 1991.
- [25] LIU X, ZHANG S, WEI F, et al. Recognizing named entities in tweets[C]// Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, 2011.
- [26] HUANG Z H, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging[EB/OL]. (2015-08-09) [2022-10-16]. <https://arxiv.org/pdf/1508.01991v1.pdf>.
- [27] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pretraining of deep bidirectional transformers for language understanding[EB/OL]. (2019-05-24) [2022-10-17]. <https://arxiv.org/pdf/1810.04805.pdf>.
- [28] 李晓庆, 唐昊, 司加胜, 等. 面向混合属性数据集的改进半监督 FCM 聚类方法[J]. 自动化学报, 2018, 44(12): 2259-2268.
- [29] MIKOLOV T, CHEN K, DEAN J, et al. Efficient estimation of word representations in vector space[EB/OL]. (2013-09-27) [2022-10-17]. <https://arxiv.org/pdf/1301.3781v3.pdf>.
- [30] 官赛萍, 靳小龙, 贾岩涛, 等. 面向知识图谱的知识推理研究进展[J]. 软件学报, 2018(10): 2966-2994.
- [31] CHEN Y, GOLDBERG S, WANG D Z, et al. Ontological pathfinding: mining first-order knowledge from large knowledge bases[C]// Proceedings of 2016 International Conference on Management of Data, 2016.

A knowledge-driven approach for intelligent generation of hydraulic engineering contingency plans: A case study of the Middle Route of South-to-North Water Diversion Project

LIU Xuemei^{1,2}, LU Hankang¹, LI Hairui³, HUAI Xianfeng⁴, CHEN Xiaolu⁴

(1. School of Information Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450000, China;

2. Collaborative Innovation Center for Efficient Utilization of Water Resources, Zhengzhou 450000, China;

3. School of Management and Economics, North China University of Water Resources and Electric Power, Zhengzhou 450000, China;

4. China South-to-North Water Diversion Group Middle Line Co., Ltd, Beijing 100038, China)

Abstract: Traditional emergency plans for hydraulic engineering projects have problems such as low digitisation, poor content relevance and insufficient intelligent aid for decision-making. In this paper, we use the knowledge graph and deep learning technology to create an intelligent generation model of emergency plans for hydraulic engineering projects. Firstly, based on the text of risk prevention and control manual and the emergency plan, we propose a knowledge graph ontology model of emergency plan, construct a knowledge graph of emergency plan, and realise the structured expression of non-structural information in the text of emergency plan. Secondly, based on the water resources engineering inspection text, we use BERT (Bi-directional Encoder Representation from Transformers) + BiLSTM + CRF (Bi-directional Long Short Term Memory with Conditional Random Fields) entity recognition model to intelligently identify risk events, projects and other entities in the inspection text. Finally, an intelligent generation template for emergency solutions is designed, and through multi-feature fusion of entity alignment technology, knowledge retrieval and inference technology, the intelligent generation and pushing of emergency solutions is realised. Through the model accuracy analysis and the validation of examples such as “channel leakage”, this paper shows a high recognition accuracy ($F1$ value of 96.21%) and a reliable emergency plan generation, which can be extended to emergency management such as emergency rescue of hydraulic engineering projects and intelligent generation of emergency plans.

Keywords: knowledge graph; hydraulic engineering; emergency plan; South-to-North Water Diversion

(责任编辑: 于福亮)