

基于知识图谱的抽水蓄能施工质量问题库构建与智能辅助决策系统

兰 天

中国水利水电第十二工程局有限公司, 浙江 杭州 310030

[摘要]针对抽水蓄能电站施工质量管理过程中出现的知识分散、问题查找困难、决策反应迟缓等状况,提出基于知识图谱的施工质量问题库构建和智能辅助决策系统。经过多源数据采集和标准化处理、知识抽取与融合、图数据库存储等一系列步骤之后,得到渗漏偏大、变形偏大、结构破损、洪水破坏、设备故障这五种典型质量问题的领域知识图谱。在此基础上,创建包含质量问题关联发掘和根因剖析,施工进度质量风险预警并调控,人机协同智能决策互动等模块的辅助决策体系。以某抽水蓄能电站为例对系统的知识图谱覆盖度、决策推理速度、自适应更新能力、场景通用性进行评价。

[关键词]抽水蓄能电站; 施工质量; 知识图谱; 智能决策; 知识抽取

DOI: 10.33142/hst.v9i7.20250

中图分类号: TV737

文献标识码: A

Construction of a Quality Problem Database for Pumped Storage Construction Based on Knowledge Graph and Intelligent Decision Support System

LAN Tian

Sinohydro Bureau 12 Co., Ltd., Hangzhou, Zhejiang, 310030, China

Abstract: In response to the problems of knowledge dispersion, difficulty in finding problems, and slow decision-making response in the construction quality management process of pumped storage power stations, a construction quality problem library based on knowledge graph and an intelligent auxiliary decision-making system are proposed. After a series of steps including multi-source data collection and standardization processing, knowledge extraction and fusion, and graph database storage, domain knowledge graphs of five typical quality problems, namely leakage, deformation, structural damage, flood damage, and equipment failure, were obtained. On this basis, create an auxiliary decision-making system that includes modules such as quality problem association mining and root cause analysis, construction process quality risk warning and regulation, human-machine collaborative intelligent decision-making interaction, etc. Taking a pumped storage power station as an example, evaluate the knowledge graph coverage, decision reasoning speed, adaptive updating ability, and scenario universality of the system.

Keywords: pumped storage power station; construction quality; knowledge graph; intelligent decision-making; knowledge extraction

引言

抽水蓄能电站是电力系统重要的调节电源,近几年来进入了大规模建设阶段。由于地质条件复杂、工程技术多样化、参建单位能力不同等原因,渗漏偏大、变形偏大、结构损坏、洪水破坏、设备故障等问题时有发生,对工程的安全、进度和投资产生影响。我国抽水蓄能电站工程质量问题具有系统性、多发性,原因涉及勘察设计、施工工艺、材料质量、管理组织等许多方面。施工质量管理中问题记录大多以文本报告、表格台账的形式,知识分散在各个参与方,没有形成系统的组织。当新的问题出现的时候,管理者很难从大量的历史资料中迅速找到经验,决策大多依靠个人的经验。近年来,知识图谱技术被应用于水电、

铁路、桥梁等工程中,但是对于抽水蓄能施工质量管理方面的研究较少。

1 系统总体架构与关键技术路线

系统总体结构由四层组成。数据层收集设计图纸、施工记录、质量检查报告、验收文档、规范标准以及历史质量问题案例等各种数据。知识层对多源数据进行标准化处理,通过知识抽取和融合来创建施工质量问题的知识图谱,并将其存放在图数据库里。决策层基于知识图谱的基础,可以挖掘质量问题之间的联系,推理问题产生的原因,并给出相应的预警和处置建议。交互层通过可视化图谱展示,自然语言问答等形式进行人机协作的决策。

系统用自顶向下和自底向上的方法来建立知识图谱。

自顶向下方向按照水电工程施工质量管理领域的知识,预先确定实体类型体系和关系类型体系;自底向上方向从具体的施工文档、问题案例中提取实例,填充知识图谱的数据层。两种方法互相验证、互相补充,保证知识体系的完整性、准确性。将大语言模型应用于知识抽取和智能问答,以提高非结构化文本的处理速度和用户交互的自然程度。

2 抽水蓄能施工质量问题库构建

2.1 施工质量问题多元数据采集与标准化处理

施工质量问题数据来源有三类。结构化数据指的是质量验收评定表、检测试验数据、测量监控数据等带有一定字段定义和数值范围的数据。半结构化数据包含施工日志、质量检查记录、监理月报等,正文用自然语言描述。非结构化数据包含设计变更通知单、质量问题处理报告、专家评审意见、规范标准条文和工程案例经验总结等,知识密度高但是提取难度大。

数据采集以该抽水蓄能电站为研究对象,在施工质量控制方面从上水库、下水库、输水系统、地下厂房、机电安装等主要施工环节入手,增加与同类电站公开案例的资料来源以提高资料多样性。预处理包括文本清洗(去除无关信息)、分词词性标注、格式统一转换、实体边界初步

标注等环节,将半结构化和非结构化数据全部转换为纯文本形式。

2.2 质量问题知识抽取、融合与图谱存储

实体抽取采用 BiLSTM CRF 模型,对质量问题的类型、施工部位、材料、工艺、设备、责任主体等进行实体识别。关系抽取用依存句法分析和规则匹配相结合的方法,用大语言模型辅助抽取自由文本中隐藏的关系。通过知识融合解决实体冲突和指代消解问题,把“面板混凝土裂缝”和“防渗面板开裂”等统一映射。

建立的知识图谱模式层有六个主要实体类别,即质量状况、施工区域、处置办法、材料及责任方等五个关系类型。实例层保存具体的三元组。使用 Neo4j 图数据库进行存储,创建节点以及关系的索引^[1]。表 1 中的知识图谱有五个质量问题类型,每一个质量问题类型下面又有成因、处置措施、预防措施这三个部分。

图 1 为以渗漏偏大会导致的质量问题为中心的局部知识图谱,其中包含了各种各样的实体之间的关系。该节点向施工部位、成因、处置措施节点延伸,成因节点之间存在相互影响,给根因推理提供图数据基础。

表 1 施工质量问题知识图谱实体与实例统计

质量问题类型	实体数量	关系数量	覆盖的施工部位	典型成因节点	典型处置措施节点
渗漏偏大	86	142	上水库库盆、输水隧洞、大坝防渗体	勘察质量缺陷、防渗体系设计缺陷	补充灌浆、防渗面板修复
变形偏大	74	128	大坝坝体、边坡、地下洞室	填筑质量控制不严、支护设计滞后	加固支护、补充填筑、监测加密
结构破损	68	115	混凝土防渗面板、隧洞衬砌	温控防裂措施不周、配合比不当	裂缝修补、结构加固、温控优化
洪水破坏	42	83	土石坝施工期坝体、导流设施	施工期洪水估计不足、排水措施不力	排水系统改造、坝体修复
设备故障	55	97	闸门、启闭机、水泵水轮机	设备制造缺陷、安装质量控制不严	设备维修更换、安装工艺复核

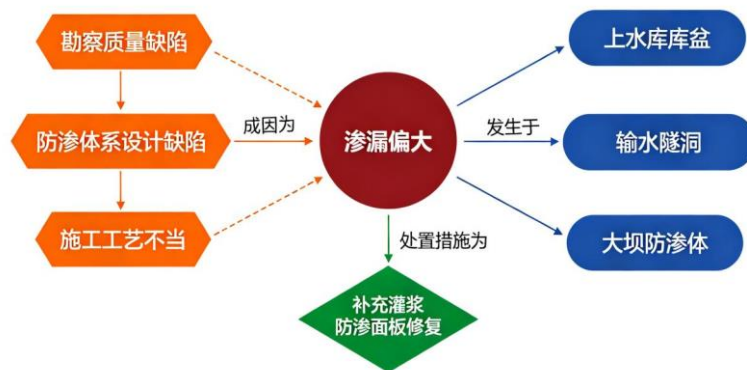


图 1 渗漏偏大质量问题局部知识图谱示意

3 基于知识图谱的智能辅助决策系统设计

3.1 质量问题关联挖掘与根因诊断推理

当施工现场出现新的质量缺陷的时候,系统会先对缺陷描述文本进行实体识别以及关系抽取,从而将其转换成图谱中的候选节点。接着用该候选节点为起点,在知识图谱上进行多跳路径遍历,找到和目前问题相似的历史案例。相似度计算使用的是语义相似度和结构相似度相结合的方式。语义相似度用 Sentence BERT 模型计算当前问题文本和历史问题文本的向量距离,结构相似度比较当前问题在局部子图中的模式匹配程度。根据两者的评分结果,给出当前问题最符合的案例以及该案例的根因分析结论和处理办法。根因诊断推理依靠图谱里的多步路径分析,自动推导出可能的成因链条。系统检索出隧洞衬砌裂缝作为当前问题节点,沿着“属于”到结构破损的关系边定位到问题类别,再沿着“成因”为“温控防裂措施不周全”、“成因”为“混凝土配合比不当”两条关系边获取候选成因,最后检索路径中隐含的原材料质量波动作为更深层次的因素一起呈现出来。

3.2 施工过程质量风险预警与动态调控

风险预警功能就是将知识图谱和实时监测数据结合在一起。系统在知识图谱中预先定义了各种质量问题类型和施工工序、施工参数之间的关联模式。大坝填筑时碾压遍数不够和沉降变形过大有很强的关联性,混凝土浇筑过程中入仓温度过高和结构裂缝出现之间存在可以量化的关联强度^[2]。系统在施工过程中不断采集实测数据,当某一工序的实际施工参数连续超过标准范围,并且和图谱上记录的某种质量问题前兆特征相符的时候,就会发出风险预警。预警信息采用分级方式呈现,包含预警等级、可能引发的质量问题种类、建议采取的预防措施和预计的影响范围。动态调控模块把预警信息同施工计划联系起来,给现场管理人员发出调整提议,比如改变下一层填筑的碾压参数,延长预沉降时间或者调节混凝土浇筑时段等,把质量把控从事后检查提前到过程中的主动干涉之中。

3.3 人机协同的智能决策交互与解释机制

智能辅助决策系统不是要取代人的判断,而是加强人

的决策能力。系统设计了两种交互方式。自然语言问答模式使管理人员可以使用自然语言提问,比如“输水隧洞出现渗水应该怎样处理”,系统会通过意图识别来理解用户的问题,然后在知识图谱中查找相关的实体以及关系,并用结构化的答案或者段落的形式给出处理方法。可视化图谱浏览模式用交互式的图谱展示当前知识库中实体之间的关系网络,用户可以沿着关系边一层层地展开,自己建立对于质量问题成因和处置逻辑的认识。为了提高决策的可信度和可复核性,系统每次推理的结果都会附带解释信息,即所依据的实体路径、匹配的历史案例来源以及相似度评分等,使得系统输出的决策可以被追溯、被验证。

3.4 系统模块集成与数据流设计

系统各个模块用同一个数据总线来实现集成。数据流的前向过程是从施工质量管理文档、实时监测数据出发,经过数据清洗模块、知识抽取模块,得到新的三元组实例,最后由图谱管理模块判断是否需要更新知识图谱。新知识被人工审核通过之后,就以增量的方式被合并到 Neo4j 数据库里,从而达到知识不断积累的目的。决策推理模块只用知识图谱作为唯一的数据源来完成查询、匹配、推理等工作,推理结果由交互模块展示出来。知识图谱是各个模块共有的核心资产,它的质量好坏会直接影响到所有的上层决策功能。

4 系统决策效能与可靠性评估

4.1 知识图谱覆盖度与准确性评估方法

覆盖度评价是从实体覆盖、关系覆盖两方面来完成的。实体覆盖度是以施工质量管理标准知识体系为依据,统计知识图谱中已经定义的实体类型和实体实例所占的比例。准确性评价用抽样核查的方法,从知识图谱中随机抽取 100 个三元组实例,由有施工质量管理经验的专家独立作出判断,对每一个三元组的事实层面是否正确进行标记^[3]。从表 2 可知,各类质量问题实体的覆盖度都在 80%以上,渗漏偏大、变形偏大两类实体的覆盖度均大于 85%,与这两类问题在该电站实际工程案例中发生频率高、积累的文档资料多相吻合。关系抽取准确率为 89.0%,比实体抽取高,说明关系模式定义比较有区分度。

表 2 知识图谱覆盖度与准确性评估结果

评估维度	渗漏偏大	变形偏大	结构破损	洪水破坏	设备故障	平均值
实体覆盖度 (%)	86.2	85.7	83.4	81.2	80.9	83.5
关系覆盖度 (%)	84.5	83.9	82.1	79.8	79.3	81.9
实体抽取准确率 (%)	87.0	86.0	85.0	83.0	84.0	85.0
关系抽取准确率 (%)	91.0	89.0	90.0	87.0	88.0	89.0

4.2 决策推理效率与实时性测试

部署服务器到某抽水蓄能电站之后对系统的决策推理效率进行了测试。测试数据集为该电站在施工期间所记录的 50 个质量缺陷描述文本，每个文本对应一个查询请求。系统记录从接收查询到返回推荐结果全过程所用的时间。从测试结果可知，简单的查询平均响应时间为 1.2s，复杂的查询平均响应时间为 3.8s。查询复杂度主要由知识图谱的检索深度决定，单跳路径匹配的简单查询响应很快，三跳以上路径遍历和多源信息融合的复杂查询响应时间会有所增加，但是仍然可以接受。实时监测数据预警触发时间小于 500ms，可以满足施工现场预警信息的及时响应。

4.3 系统自适应更新与鲁棒性分析

施工质量管理知识是不断发展的，随着工程的推进，新的质量问题类型以及处理方法层出不穷。系统设计了增量更新机制，每次新增施工文档入库之后就会自动执行知识抽取任务，对新产生的实体、关系进行提取并初步融合，最后提交给质量管理专家快速审核确认。审核通过的知识用增量的方式合并到 Neo4j 数据库里，不需要重建整个图谱。鲁棒性分析是从两个方面进行的^[4]。系统对于输入噪声的抗干扰能力，用向测试文本中随机加入不同比例无关词语的方式来评价，当无关词语比例小于 15% 的时候，系统的大部分推理功能不会受到太大的影响。局部知识缺失对于系统性能的影响用人为删除图谱中 5%、10%、20% 的实体节点来模拟，当缺失比例小于 10% 的时候，系统仍然可以基于剩余子图给出合理的推理结论，只是相似度推荐的置信度会降低。

4.4 面向不同施工场景的泛化能力验证

为了检验系统在不同的施工场景中是否适用，选取了某抽水蓄能电站上水库库盆施工、输水隧洞开挖支护、地下厂房混凝土浇筑这三个典型的施工场景进行系统的测试，测试系统在场景切换时知识匹配的效果。测试结果说明系统对于相似案例的推荐准确率分别是 87%、84%、86%，场景间存在差异，主要是因为各个场景下文档资料的丰富程度以及知识图谱对该场景的覆盖程度不同所导致的。上水库库盆施工在本电站已经形成了较为完整的质量问题记录，覆盖率、准确率均较高；输水隧洞开挖支护场景下，围岩地质条件变化比较复杂，一些非常规地质条

件下出现的工程质量问题没有被完全反映到现有的知识库当中，从而使得准确率有所降低。从整体上来说，在主流的施工场景中系统有较好的知识迁移能力，所用到的知识图谱构建方法是可推广的。

5 结束语

文章针对抽水蓄能电站施工质量管理中知识分散、问题追溯困难、决策依靠个人经验等现象，创建起以知识图谱为基础的施工质量问题库和智能辅助决策系统。选取某抽水蓄能电站作为数据来源和验证平台，从多源数据采集、知识抽取融合、图谱存储和智能决策等各方面进行全过程的设计。系统对工程质量关联性进行挖掘和分析，对施工过程中的各种风险做出预测并加以控制，给工程实践提供功能支持。通过评价可知知识图谱覆盖率 80% 以上，实体抽取准确率为 85.0%，关系抽取准确率为 89.0%，系统响应时间可以满足施工现场的实时性要求。该系统给抽水蓄能电站施工质量管理由经验驱动转向知识驱动提供了一条可行的技术途径，对同类水电工程的智能化质量管理有借鉴意义。未来需要继续扩大知识图谱的规模以及数据来源，加入更多的同类电站典型质量问题案例，研究物理模型、数值模拟结果等可以被知识图谱所包含的数据；探索将物理模型、数值模拟结果等加入知识图谱中，使知识图谱由数据驱动型向机理驱动型转变；改进人机交互方式，开发移动端应用提高现场使用便捷性。

[参考文献]

- [1]杨阳蕊,董方宁,王鹏斐,等.基于大语言模型的水利工程运行管理质量概念模型及知识图谱自动化构建[J].水利学报,2025,56(5):634-645.
- [2]向云飞,罗一鸣,宁泽宇,等.水电地下工程安全管理多模态知识图谱构建方法[J].清华大学学报(自然科学版),2025,65(3):433-445.
- [3]陈述,王典学,杨应柳,等.水电工程施工安全隐患语义匹配模型[J].中国安全科学学报,2024,34(12):40-47.
- [4]余端仁,史健勇.基于大语言模型的施工质量知识图谱构建[J].土木工程与管理学报,2025,42(2):92-100.

作者简介：兰天（1988—），男，高级工程师，毕业于中国石油大学（北京），现就职于中国水利水电第十二工程局有限公司，项目执行经理，主要从事大型水电站、抽水蓄能项目管理。