基于多维级联知识抽取的核电建造经验反馈图谱研究

李池璐[1] 熊世权[1*] 雷玮剑[2] 刘疏桐[1] 李英敏[2] 易 茜[1] 易树平[1]
1.重庆大学机械与运载工程学院,重庆,400044
2.中广核工程有限公司,深圳,518124

摘要:针对核电装备建造过程中经验反馈时效性弱、各参与方对经验知识认知不统一、经验反馈知识复用率低等问题,将经验知识化,构建统一的经验反馈图谱。分析了经验反馈信息的特征,构建了经验反馈关键知识域结构与图谱框架;提出的多维信息感知模型和改进 PCNN 关系抽取模型在核电企业数据集上取得了良好效果,并基于 Neo4j 图数据库生成了经验反馈图谱。研究结果表明该方法可完成核电产品的隐性知识深层次检索,辅助核电专业技术与管理人员快速解析、传递经验反馈知识,实现核电建造经验反馈沉淀应用模式从以文档为中心向以知识为中心的转变。

关键词:核电装备;经验反馈;知识图谱;自然语言处理
中图分类号:TL372
DOI:10.3969/j.issn.1004-132X.2025.01.019 **开放科学(资源服务)标识码(OSID)**:

Research on Experience Feedback Graph of Nuclear Power Constructions Based on Multi-dimensional Cascading Knowledge Extraction

LI Chilu[1] XIONG Shiquan[1*] LEI Weijian[2] LIU Shutong[1] LI Yingmin[2]
YI Qian[1] YI Shuping[1]
1.College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing, 400044
2.China Nuclear Power Engineering Co., Ltd., Shenzhen, Guangdong, 518124

Abstract: Aiming at the problems of weak timeliness of experience feedback in the process of nuclear power equipment constructions, inconsistent perception of experience knowledge by each participant, and low reuse rate of experience feedback knowledge, the experience knowledge was intellectualized, and a unified experience feedback mapping was constructed. The characteristics of experience feedback information were analyzed, and the structure of key knowledge domains and mapping framework of experience feedback were constructed. The proposed multi-dimensional information perception model and improved PCNN relationship extraction model achieved good results on nuclear power enterprise dataset, and the experience feedback mapping was generated based on Neo4j graph database. The results show that the method may complete the deep-level retrieval of tacit knowledge of nuclear power products, assist the nuclear power professional technology and management personnels to quickly parse and transfer the experience feedback knowledge, and realize the transformation of the application mode of nuclear power construction experience feedback precipitation from document-centered to knowledge-centered.

Key words: nuclear power equipment; experience feedback; knowledge graph; natural language processing

0 引言

核电建造须严格遵守《中华人民共和国核安全法》《核电厂质量保证安全规定》(HAF003)等相关法律法规建立核安全经验反馈制度,即将质量活动的执行情况详细记录为文本,形成专业知识和经验并传承给各参与方。经验反馈作为核电建造过程中一种重要的知识管理方法和改进机制,能帮助核电建造团队识别有价值的问题和潜在风险,以便及时采取相应措施进行改进和优化,防止类似事件和问题重复发生,对提升建造质量、提高协同管控效率、优化资源配置等具有不可忽视的作用。为此,国内外核电企业建立了较为完善的经验反馈与运作信息化系统,以便和参与方共享经验反馈信息。

随着核电工程的发展和智能制造技术的提

收稿日期:2024-03-23
基金项目:国家重点研发计划(2020YFB1711700);中央高校基本科研业务费专项资金(2021CDJKYJH022)

高，核电企业开始研究如何进一步利用和传递经验反馈信息，从传统的以文档为中心向以知识为中心转变。赵栋等[1]开发的经验反馈管理系统为核电工程设计经验的归档、总结、发布及再利用提供了统一的平台。邹象等[2]提出一种基于AHP的核动力厂安全相关经验反馈信息重要性评价方法，通过对信息进行相对重要性评价得到反馈价值，以识别重要的经验反馈信息。上述方法传递的经验反馈信息颗粒度比较大，其中隐含的关键要素仍需通过专业人员作进一步分析处理，经验反馈信息难以被直接有效利用。核电领域迫切需要借助智能化技术来分析和提取经验反馈信息并获取重要内容，将建造过程产生的大量质量事件记录梳理整合为结构化的经验反馈知识，辅助相关人员快速分析关键要素，实现核电建造质量信息的跨企业高效流转和复用。

近年来，自然语言处理（natural language processing，NLP）利用计算机处理、理解、规范表达人类语言文本，能智能化解析经验反馈信息中的重要知识，提高可利用性。SINGHAL[3]提出的知识图谱（knowledge graph）可将经验反馈关键领域中的重要知识进行统一形式的可视化表达，有利于大规模经验反馈信息的高效利用和传递[4]。基于NLP技术的知识图谱能有效减轻核电经验反馈信息处理过程对人工先验知识的依赖，减少了人因失误，提高了数据要素赋能作用，同时满足了跨企业信息传递的需求。CHOI等[5]采用基于规则的信息提取方法在核电站的运行程序中提取语法和语义信息。邓志光等[6]通过改进长短时记忆神经网络，在核电厂反应堆出口温度现场数据中提取故障特征。吴庭伟等[7]采用预训练语言模型和门控循环神经网络挖掘出核电质量缺陷的相关要素。ZHAO等[8]采用字典与规则相结合的方法训练关系抽取模型，提取出核电质量事件中的因果关系。刘晓瑛[9]通过双向网络和改进的卷积神经网络模型从各类文档中抽取出核电设备相关知识，并用知识图谱进行储存。上述研究多依赖于手工与规则模版或是单一的算法，模型的泛化能力较弱、应用支持能力不足，针对核电建造经验反馈信息的多维度关联挖掘研究还存在一定空白。

知识图谱通过结构化和语义化的数据表示，提供了更加精确和丰富的信息组织方式，并且可进行逻辑推理和推广，有助于呈现与传递经验反馈文本中的知识。为此，笔者利用多维级联知识挖掘构建经验反馈图谱，根据经验反馈信息特点建立结构化图谱框架，利用自然语言处理方法搭建基于深度学习的知识抽取模型，通过图数据库统一规范化表达与关联，实现核电建造经验反馈的知识化。

1 核电建造经验反馈问题描述

1.1 价值链协同下的核电建造经验反馈内涵

核电建造[10]过程积累的大量质量相关知识、经验需要被传承，但由于核电装备的复杂性和特殊性，产业链涉及上千家企业，因此在多方协作参与建设的环境下，大规模的企业群显然无法通过专项会议等方式学习同期经验，更为直接有效的方式是各方将质量数据收集、记录形成的经验反馈知识与团队共享。

作为PDCA循环质量改进的重要体现和资源配置的有效手段，经验反馈能通过前期质量事件的发生和处理经验指导后续同类工作，避免在工程中犯同样的错误，实现持续改进。这既是进行质量管理的重要依据，又是“预防为主”质量方针的具体内容。同时，数据联动、业务合作、资源和技术优势互补等方式能帮助企业尽量避免和减少价值减少环节，扩大产业链整体价值，实现核电装备全生命周期整体价值创造和增值效益的最大化发展[11]。

1.2 经验反馈信息特征分析

核电建造中的经验反馈信息绝大部分是以质量事件文本形式记录的，包括事件描述、原因分析、改进行动，如表1所示，并具有以下特征。

1.2.1 经验反馈信息收集与传递有时效要求

一个典型成熟的2台百万kW级机组的压水堆核电站平均包括400余个工艺系统、48 000余台套设备与备品备件[12]，建造过程涉及数千家供应商，核电产品有多维度的质量指标、多层次的质量控制、长期的运营监测。核电站建造期间产生并记录的变更、意外事件、不符合项报告等质量文本共计十余万份。这些用作经验反馈的信息在各阶段活动会根据业务需求进行多个版本的迭代更新，并不断增长，需要在一定时间内及时反馈给相关单位开展纠正工作。

1.2.2 经验反馈信息内容复杂

核电装备作为典型的复杂产品，其建造是一个高度复杂的过程，涉及多个专业领域和大量技术要求。核电建造经验反馈信息记录了质量事件发生的全部过程，涵盖核电领域全生命周期中各活动的专有名词和缩略语，如主蒸汽过滤器法兰、常规岛高压管道支吊架、励磁涌流影响、氢脆等。这些信息相互之间产生的关联关系更复杂，需要借助专业人员及技术充分融合并挖掘利用。

表 1　经验反馈信息示例

Tab.1　Example of experience feedback information

序号	质量事件记录
1	2018 年 4 月 24 日，机组状态 RP 模式下压力容器参考水位压差 3RCP094/095MN 无法读取，ICCMSSU 工作站上监视软件故障无法启动。RICSU 单元作为 RIC 系统终端显示单元，主要作用包括系统运行状态监视、内部数据显示，但不参与系统逻辑运算及控制。大修期间该故障影响压力容器水位监视，oncall 维修仪控检查，软件与 ICCMS 机柜通信协议证书过期失效（EOMM 手册中给出的证书有效期为 15 年）。EOMM 手册提供的问题解决信息有限，目前现场暂无法处理，需等待外方厂家技术支持。目前已就该问题与法国厂家联系，初步分析原因为软件通信协议证书过期失效导致 SU 单元无法与系统建立通信。 2018 年 4 月 29 日依据厂家给出的方案完成软件通信协议证书激活，SU 单元工作恢复正常。
2	2016 年 3 月 24 日，某核电工程现场对 4 号机组 RPV54＃接管座上的 CRDMΩ 焊缝焊接后，通过 PT 发现接管座母材存在共三处 16 个圆形显示。现场依据现场验收文件第 10 条验收标准：不得有任何缺陷显示，现场开启 NCR。 2016 年 3 月 29 日至 31 日，供应商依据 NCR 答复意见，对 54＃接管座母材 PT 缺陷显示区域进行抛磨处理，部分位置已打磨到允许的极限位置。处理完成后进行 PT，此前的母材缺陷显示未完全消除，还剩余 6 个 ϕ0.1 mm、3 个 ϕ0.2 mm 圆形显示（三个位置打磨量约 0.08 mm、0.06 mm、0.08 mm），且在设计允许的抛磨余量范围内无法彻底消除缺陷显示。 事件后果：经过设计审查并进行了水压试验，同时向 GRO 汇报后，原样接收。

1.2.3　经验反馈信息来源离散

核电建造协同质量活动离不开各信息系统的参与，随着信息技术的发展，工作形式逐渐从现场办公转为互联网远程办公，产生大量经验反馈信息的见证签点工作也通过移动端在线上开展。多方协同环境下的工作方式变化给各个企业带来更多的自主性，导致质量事件往往由不同公司不同部门的质保人员在信息系统中进行记录和撰写，造成信息的孤立与分散。

1.2.4　经验反馈业务链条“以人为主”

据统计，单台压水堆核电机组在设计建造期间平均产生约 3 万项业务变更。受背景知识、变更或真实事实资料（过程、原因、趋势）掌握完整度、深度等差异因素影响，业务变更与经验反馈链条上各参与方对同一数据的理解、观点存在偏差，导致在文字组织、结论选择上存在“取舍”。另外，仅依靠人工记忆难以避免相关人员在部分经验反馈的范围、层级等维度上存在信息盲区。

1.2.5　经验反馈历史数据与增量数据具有强关联性

随着增量数据的不断累积，在新核电项目开工前，需对历史项目经验反馈数据开展分析，导致后续项目的数据累积增加，而人工判定会导致数据利用的广度与深度受限，很难发现和处理经验反馈数据之间的关联性，造成部分高价值信息难以有效反馈。

1.3　核电建造经验反馈现状及不足

核电建造业务活动包括设计、采购、施工、调试阶段，涉及三维建模、核岛布置、现场安全、辐射与屏蔽、环保等 29 个专业，各个环节都会产生大量、实时的经验反馈信息[13]。

针对发生的质量事件，一般的经验反馈流程如图 1 所示。首先，由现场人员记录业务变更情况并进行初步分析，经所在部门主管生产的领导审核后交生产处的经验反馈部门。生产处的经验反馈工程师根据质量事件对来自不同渠道的信息进行甄别分类，并将经验反馈报告下发给对应部门的负责人继续跟踪。各参与方通过收集与整理

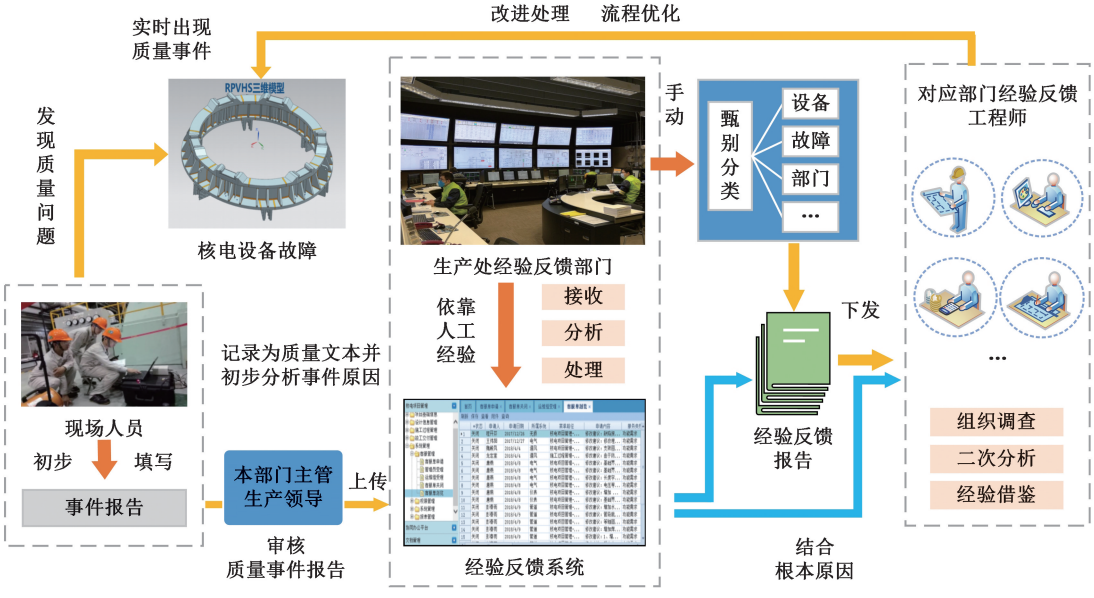


图 1　核电建造期间经验反馈流程

Fig.1　Experience feedback process in nuclear power construction

问题和事件记录、总结经验教训，为后续的核电建造提供有益的借鉴和参考，避免类似错误的重复发生，优化资源分配，持续提高核电质量水平。

实际执行过程中的经验反馈存在如下问题：

1)企业间经验反馈时效性弱。核电建造过程中，来自不同企业、部门的质量数据日益增长，需要各部门的经验反馈工程师花费大量的时间和精力来收集和整合。信息过载的情况下，缺乏有效的筛选机制和工具会导致单纯依靠人工难以从海量数据中快速准确地获得有效的质量信息并传递给其他企业，这使得经验反馈信息供给不及时、不充分，缺陷失去纠正窗口或短时期内二次发生，造成物料资源、数据资源的浪费和建造进度滞后。

2)各参与方对经验知识认知不统一。核电建造中产生的经验反馈信息通常是由现场人员在质量事件发生时记录，并以非结构化文本形式存在，包含丰富的核电领域知识和专业术语，信息之间的层次关系复杂。由于缺乏统一的格式和标准，“以人为主”的经验反馈方式使得个人的记录习惯与方式不同，导致信息提取和分析变得困难，一些重要的隐性知识和经验需要专业经验反馈工程师的理解和解读才能获取。文本形式的信息容易受当事人表达能力和主观意见影响，可能存在表述模糊、错误或主观偏见的问题。接收信息的工程师理解视角与评判尺度等个人因素会导致其对质量事件最终定级、核心属性字段的认定结果不同，从而导致重要的经验反馈知识丢失或误解，无法确保信息的可靠性和可信度。由于没有统一的知识表达规范，知识在企业间的传递和共享变得困难，各参与方对经验反馈信息认知不统一容易导致经验反馈失效。

3)经验反馈知识复用率低。经验反馈信息蕴含丰富重要的数据，但核电建造过程中碎片化的经验反馈信息分散在各个系统，没有被集中组织和高效管理，形成特定知识库，这导致集成度差、知识间缺少有机关联，简单依靠各方人员在对应系统中检索，甚至存在相同或相似的问题重复录入、在不同项目中重复出现的情况。企业难以充分挖掘和进一步利用已有的经验反馈信息，关联性的隐性知识大量流失，降低经验反馈的效果。

2 核电建造经验反馈知识化

2.1 经验反馈知识化的概念

核电建造经验反馈的知识化是指将从实践中获得的经验反馈文本信息转化为结构化、可管理、可共享、可复用的知识的过程。这个过程旨在将核电建造经验反馈信息中的关键内容、特征、洞察和最佳实践等关键信息内容系统化地提炼出来，并整合到一个核电建造领域知识库，以便指导各参与企业未来的行动、决策和问题解决。这有利于价值链协同模式下核电建造经验反馈信息的融会贯通和多企业合作下的有序协同，从而实现全生命周期的价值创造最大化。

2.2 经验反馈知识化的流程

经验反馈知识化流程是将经验反馈文本数据转化为知识的系统化过程。

针对经验反馈时效性弱的问题，需对经验反馈信息涉及的知识化范围进行界定，通过梳理核电质量文本数据的来源去除重复、不相关或低质量的内容，筛选有反馈价值的信息，减少投入在不必要信息处理上的精力，提高经验反馈效率。

针对建造中各参与方对经验知识认知不统一的问题，需将质量数据进行整理和筛选，通过分类、归纳、总结等方法建立统一的规则表示知识，以便融合智能化技术提取知识，确保上下游经验反馈知识理解的一致。

针对经验反馈知识复用率低的问题，需提高经验反馈信息的集中性和可检索性，通过建立领域知识库并结合核电内部业务逻辑映射知识间的关联关系，构建隐性知识索引解析与呈现表达体系，给经验反馈信息价值赋能。

具体而言，实践中获得的经验和教训转化为系统化、结构化的知识需经过范围界定、分类与组织、知识库建立等步骤。

2.2.1 经验反馈知识化范围界定

核电建造的经验反馈信息来源广泛，主要包括内部与外部。外部经验主要来自世界核电业主联合会(WANO)的重大事件报告(SER)、重大运行事件报告(SOER)。对本电站和世界同行电站已发生的问题与事件进行根本原因分析，可防止类似问题与事件的重复发生，例如福岛核事故后，核电企业纷纷汲取经验教训，将安全改进纳入设计标准，对严重事故的预防提出了更加明确的要求，并加大了对核电建造的概率安全评价(PSA)技术体系的研究[14]。核电企业还吸收世界同行电站的良好实践经验，进行内部比较检查，提出改进行动计划并跟踪落实，以降低重大事故的二次发生率。

相比于外部反馈，核电建造过程中更为及时有效的是来自内部的经验反馈。核电建造过程产生且累积了大量设计变更通知、意外事件单、不符合项报告、现场变更申请等记录，内容涵盖核电工

程项目设计、采购制造、施工安装、调试运维等全生命周期各个阶段各个方面的质量信息，如设计图纸不符合要求、现场质量见证不通过等。其中，升级为质量事件的信息包含详细事件过程、原因、具体整改措施和改进方案等内容，信息的记录具有参考性和反馈价值，需要及时处理与传递。

因此，为提高经验反馈效率，本文研究的经验反馈知识化的范围主要为核电建造过程中内部产生的质量事件记录。

2.2.2　经验反馈知识的分类与组织

核电建造过程中内部产生的质量事件记录包含的内容及类别众多，但并非所有的信息都具有传递价值，为实现经验反馈知识化，需从中筛选和界定出具有反馈价值的要素，形成统一规范的知识域，确保信息的一致性和可解释性。

经验反馈的关键知识域是经验反馈信息中关键知识的集合，是对建造过程质量事件的层次化描述，其内容主要包含关键领域实体和领域关系。前者指核电建造领域中具有重要意义和作用的实体，包括事件、对象、原因、行动、人员等；后者描述了关键领域实体之间的联系、依赖或影响关系，如因果关系、依赖关系、层次关系、时间关系等。

统计分析核电企业中记录在档的经验反馈信息具体内容可知，关于故障问题的描述性信息出现频率较高，是各参与方在建造过程中需要掌握的重点数据要素，也是影响核电建造效率和质量的关键。基于这一分析结果，结合相关领域专家意见，以“故障问题”这类实体为核心，通过实体聚类形成故障表现层；进一步分析评判与故障表现紧密相连的各类关键信息，其中，设备故障信息有利于筛选重点关注对象，并在后续阶段持续跟踪与对比项目其他机组，防止同类型设备再次发生质量问题；提取故障产生的原因能帮助企业或部门更有效地处理问题，从根源上预防问题，减少质量事件；提取故障对应措施的信息便于相关人员在同样问题发生后能快速查询解决，提高解决问题的效率。此外，时间、阶段、设备、人员、企业部门及解决效果等均为具有传递价值的必要信息，如表 2 所示。

梳理不同维度上的关键知识要素可形成经验反馈关键知识域的故障设备层、故障原因层、对应措施层等知识层。同时通过层间交互关系实现关联耦合，将离散的知识要素转化为具有语义关联和逻辑推理能力的空间知识结构，集成核电建造经验反馈资源。经验反馈关键知识域的结构如图 2 所示。

表 2　经验反馈知识类别及作用

Tab.2　Knowledge categories and their roles

知识类别	知识含义	知识作用
故障	质量事件具体表现	质量事件核心内容
设备	故障发生对象(含子设备)	便于后续跟踪监视
原因	质量事件发生原因	预测、防范质量事件发生
措施	质量事件的解决措施	调用方法处理问题
阶段	质量事件发生的阶段	了解质量事件的影响
机组	故障设备所属机组	横向对比其他机组
时间	质量事件被记录的时间	便于查询与调用
人员	质量事件的涉及人员	质量事件责任落实
企业	相关人员所属企业/部门	针对性传递经验反馈信息
效果	质量事件的解决效果	为选择措施提供参考

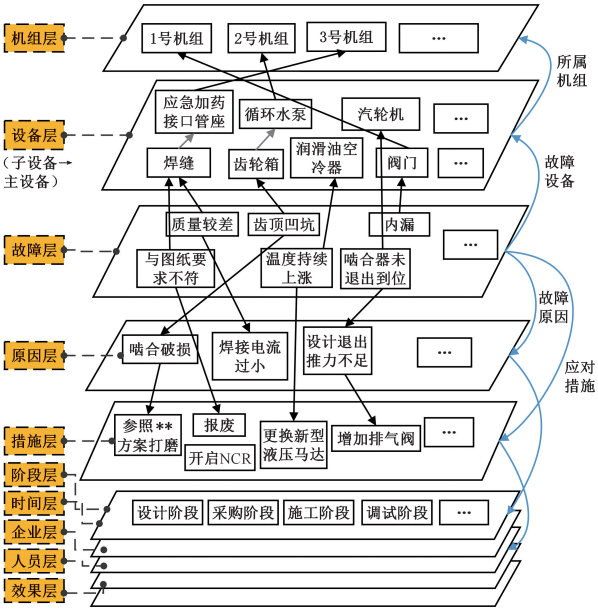


图 2　经验反馈关键知识域的结构

Fig.2　Thestructure of experience feedback key knowledge

2.2.3　经验反馈知识库建立

经验反馈知识化的目的是使经验反馈的关键信息在实际核电建造中能被传递、检索和共享。

利用知识库将核电建造经验反馈关键知识域集中管理，能改变关键知识独立分散的性质，确保经验反馈知识能统一、清晰地呈现。在此基础上，利用交互关系将经验反馈信息转化为具有语义关联和逻辑推理能力的知识结构，实现知识之间的耦合交互，即根据特定条件触发相应的推理过程，使关联知识能被获取，图 3 所示为经验反馈知识化应用的框架。

结合机器学习和数据挖掘技术，通过分析大量的经验反馈数据，自动发现潜在的规律和模式，辅助建立更智能和高效的推理规则。利用知识图谱或专家系统建立推理引擎，将推理规则转化为计算机可理解和执行的形式，提高知识的可解释性，实现核电建造经验反馈知识的自动推理和应用。

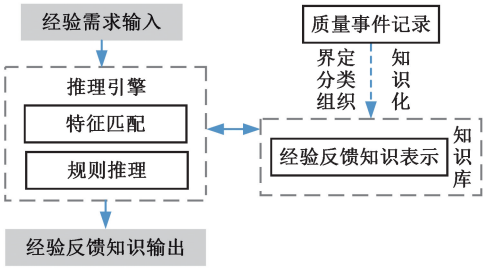


图 3　经验反馈知识化应用的框架

Fig.3　Framework for applying knowledge-based feedback on experience

3　核电建造经验反馈图谱构建

为满足跨企业环境下核电建造经验反馈知识化的需求，需借助智能化技术挖掘和组织经验反馈知识。

区别于传统的文本记录方式，本文利用知识图谱技术呈现经验反馈知识化的结果。该方法主要基于经验反馈关键知识域，利用自然语言处理技术来智能地抽取经验知识，构建统一数据格式的经验反馈图谱，通过图数据库存储结构化知识，并为相关人员提供知识检索工具，实现经验反馈知识化。

3.1　经验反馈图谱构建的总体框架及组织流程

知识图谱的构建方式通常分为自顶向下(top-down)和自底向上(bottom-down)两种[4]，前者先为知识图谱定义数据模式，从最顶层的概念开始构建，逐步向下细化形成本体模型，再将数据添加到知识库；后者通过对实体的归纳组织形成底层概念，再逐步向上抽象成上层概念。

本文采用自顶向下的构造方法，先梳理与重整核电企业中产生的大量繁杂、冗长的质量信息，将经验反馈关键知识域作为知识图谱的模式层；再根据核电经验反馈文本的特点，利用自然语言处理技术，针对性地设计命名实体识别方法和关系抽取方法，从文本中抽取出结构化形式的关键知识组并映射到图谱的本体模型，如图 4 所示。

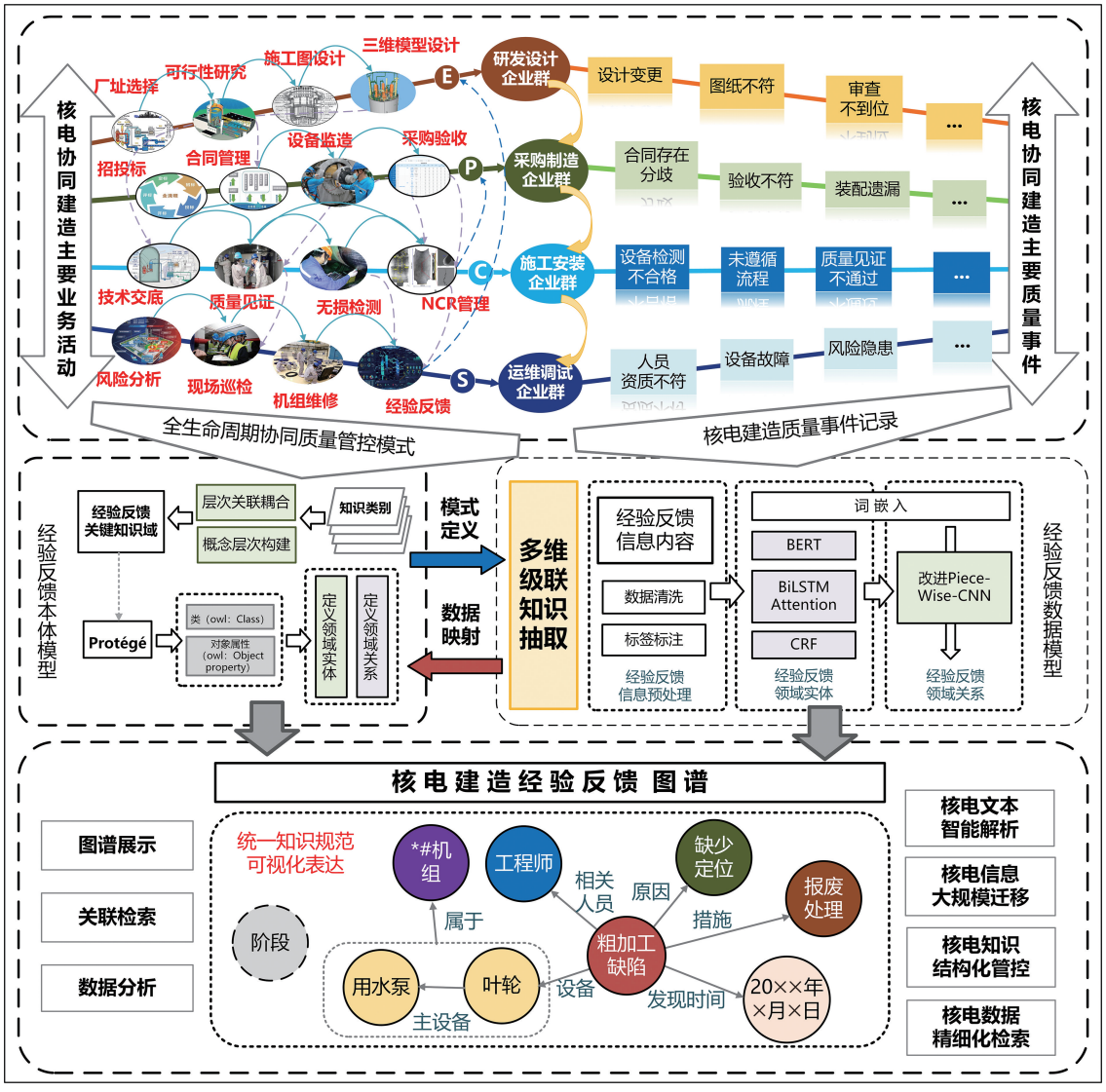


图 4　基于多维级联知识抽取的核电建造经验反馈方法

Fig.4　The feedback method for nuclear power construction experience based on multidimensional cascading knowledge extraction

3.2　经验反馈图谱本体建模

基于前文对经验反馈知识的分类与关联组织，将经验反馈关键知识域作为经验反馈图谱模式层，通过本体建模方法显性化表示，为文本知识抽取提供框架。利用 Protégé 软件将经验反馈关键知识域中不同层别的领域知识层构建为类(owl:Class)，各层之间的关联关系构建为对象属性(owl:Object property)，以建立节点类型及概念或实例之间的语义关系，获得〈实体-关系-实体〉结构的知识形式，并将经验反馈关键知识域显性化表征，形成经验反馈图谱本体模型，如图 5 所示。

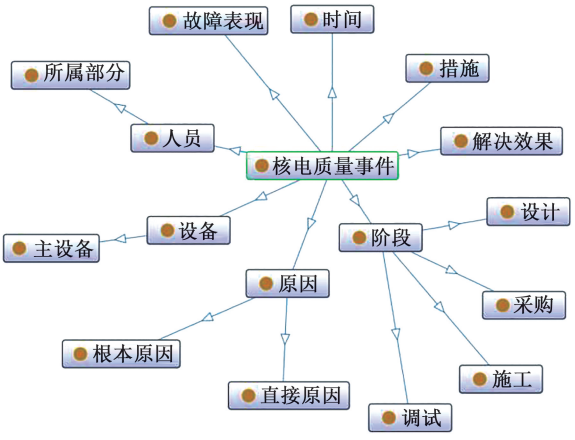


图 5　经验反馈图谱本体模型

Fig.5　Ontologica lmodeling of empirical feedback graph

通过以上方式定义的经验反馈关键知识域统一数据规范能明确知识挖掘目标，为经验反馈图谱的构建提供模式支撑。

3.3　基于多维级联知识抽取的经验反馈信息挖掘

为构建完整的经验反馈图谱，在经验反馈关键知识域框架的指导下，利用知识抽取的方法分析和挖掘核电领域质量信息的重要语义特征，向经验反馈图谱本体模型填充数据。

核电建造经验反馈信息文本内容复杂、关系类别繁多，单一的网络模型无法较好地识别和学习关键特征，经验反馈知识不易被挖掘、组织和表征。为提高在核电领域的挖掘效果，本文提出多维级联知识抽取方法，即搭建多层神经网络模型对核电领域的经验反馈信息进行多维度特征挖掘，高效抽取出关键领域知识。

3.3.1　基于多维信息感知的知识识别方法

核电质量事件所含的关键知识元素需要通过命名实体识别(named entity recognition，NER)方法从语料文本中自动识别并抽取。基于深度学习的方法可实现自动化的实体抽取，将文本以向量的形式嵌入到神经网络，实现端到端的命名实体识别，该方法可在短时间内处理大量的经验反馈文本，节约挖掘成本和时间。

结合核电领域经验反馈信息内容复杂性的特点，本文构建了包含以 Transformer 为核心的词嵌入层、以双向长短时记忆(bi-directional long short-term memory，BiLSTM)和注意力机制(attention)为核心的知识感知层、以条件随机场(conditional random fields，CRF)为核心的知识预测层的三层功能网络，从多个维度感知和抓取核电质量事件中的关键要素，如图 6 所示。在知识感知层融合注意力机制，加强对核电领域重点词的关注，有效解决核电领域知识识别效果不佳的问题。

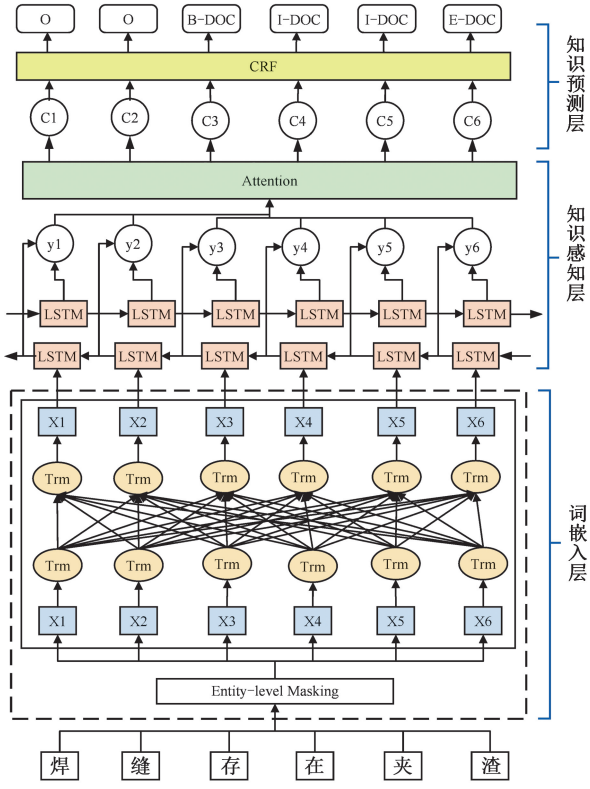


图 6　基于多维信息感知的知识识别模型

Fig.6　Knowledge recognition model based on multi-dimensional information perception

1)词嵌入层。为让计算机模型能识别核电领域文本，设计基于 Transformer 的双向编码器表示(bidirectional encoder representations from transformers，BERT)[15]的词嵌入层，将词汇映射到实数向量。BERT 是预训练网络模型，具备较强的泛化能力。该模型的输入具有三个不同的嵌入层，分别表示输入质量事件文本的词语信息(token embedding)、句子段落信息(segment embedding)和句子位置信息(position embedding)，将三层信息向量拼接作为 BERT 的最终输入，经过多层 Transformer 得到含有多个维度语义特征

的输出向量，为后续模型提供数值型数据的输入。

2）知识感知层。核电建造中产生的质量事件文本通常比较长，模型在学习识别过程中容易遗忘过去的信息，因此采用 BiLSTM[16] 双向挖掘质量事件文本中的历史信息和将来信息，将长时间的特征全面联系起来。为进一步加强文本对核电相关要素的敏感度，本文采用融合注意力机制[17]，计算核电经验反馈信息中词与词之间特征的关联程度，使模型在处理信息时将注意力集中在核电有关的区域，获得更有效的语义信息。

3）知识预测层。知识感知层的输出结果只是对核电质量事件文本中的每个字进行了学习，并没有考虑标签之间的约束关系，如“压力容器”是核电领域中的固定专业词语，字与字之间不能调整顺序，因此需对标签的连接顺序有所约束。引入条件随机场对知识感知层的输出结果进行优化。该模型在预测时结合发射矩阵和转移矩阵，通过上层输出的状态分数和转移分数来获得句子标签对应得分最高的标注序列，从而确保最后预测的经验反馈知识合理有效。知识预测层中，文本对应的标签概率 $P(y \mid \boldsymbol{x})$ 的计算公式为

$$P(y \mid \boldsymbol{x}) = \frac{1}{z(\boldsymbol{x})}\exp\left(\sum_k \boldsymbol{w}_k f_k(y, \boldsymbol{x})\right) \tag{1}$$

$$z(\boldsymbol{x}) = \sum_{y=1}^{N}\exp\left(\sum_k w_k f_k(y, \boldsymbol{x})\right) \tag{2}$$

式中：$\boldsymbol{x}$ 为经过知识感知层输出的语义特征向量；y 为一个状态序列概率；$\boldsymbol{w}_k$ 为参数向量；$f_k(y, \boldsymbol{x})$ 为特征函数；$z(\boldsymbol{x})$ 为归一化因子。

最后利用维特比（Viterbi）算法得到最佳预测标签：

$$Y^* = \arg\max_y P(y \mid \boldsymbol{x}) \tag{3}$$

3.3.2 基于改进 PCNN 的知识关系抽取

核电建造经验反馈信息关系复杂、难以梳理，需利用关系抽取（relation extraction，RE）技术提取核电质量文本关系特征。表达质量事件的核电经验反馈信息较长，使用基于词法、句法特征抽取关系的传统方法正确率不高，同时为避免引入噪声数据，采用基于多实例学习的分段卷积神经网络（piece-wise-CNN，PCNN）[18] 方法完成关系抽取。在 CNN 模型的基础上调整特征提取方案，将关系抽取视为多实例问题，仅将每个实例中最接近真实类别的一个示例作为输出。同时为减少误差的传播和积累，在句子的向量表达中引入位置向量，根据核电经验反馈信息中关键信息的位置将句子分成三段，每段单独作最大池化操作，提取出更细粒度的关系特征，最后将分段池化层的输出结果送入 softmax 层，获得核电经验反馈信息关系的分类。PCNN 网络模型的具体结构如图 7 所示。

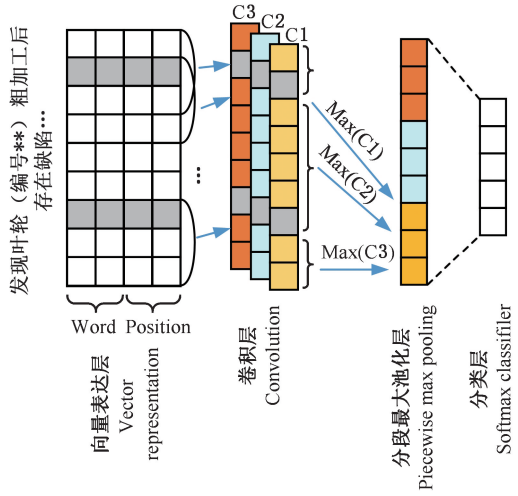


图 7 PCNN 网络模型结构

Fig.7 Structure of the PCNN network

由于核电经验反馈信息具有内容专业性强的特征，所以对 PCNN 关系抽取模型进行改进，引入正则化的思想，将具有更平滑非线性特征的高斯误差线性单元 GELU[19] 作为激活函数，提高经验反馈信息关系抽取模型的泛化能力并避免模型学习过程中的梯度消失问题。GELU 函数的具体表达式为

$$GELU(x) = \frac{x}{2}\left(1 + \tanh\left(\sqrt{\frac{2}{\pi}}x\left(1 + 0.047715x^2\right)\right)\right) \tag{4}$$

式中：x 为输入。

3.4 经验反馈图谱知识的映射与生成

3.4.1 数据来源与预处理

核电装备涉及到数据保密问题，当前相关研究和语料库都较为匮乏，核电建造经验反馈信息领域暂无可用于训练的公开数据集。为验证本文构建的多维级联知识抽取模型的有效性，笔者将项目合作核电企业信息系统中的质量事件记录作为原始训练数据。通过前期规则化采集、筛选、清洗训练数据后，根据文本的特点规范数据长度，并进行预实验，再按照预实验验证结果，利用自然语言处理任务中常用的标注策略进行标签标注和关系标注，完成数据的预处理，为文本挖掘工作提供数据基础。通过上述处理建立的数据集由 350 条质量事件记录组成，将数据按照 7∶1.5∶1.5 的比例分割为训练集、测试集、验证集，以便计算模型的具体评价指标。

3.4.2 模型参数设置

通过文献检索和预实验结果，设置知识挖掘模型中主要的超参数，如表 3 所示，运行环境为 Python3.7，使用 PyTorch1.11.0 框架，将预处理的经验反馈数据集投入模型进行训练。

表 3　模型超参数设置

Tab.3　Model super parameter setting

模型	模型参数	含义	参数值
经验反馈实体识别	lr	学习率	3×10^{-5}
	batch_size	批量大小	16
	dropout	随机失活率	0.3
	epoch	迭代次数	100
	max_seq_len	句子最大长度	150
经验反馈关系抽取	lr	学习率	10^{-3}
	batch_size	批量大小	256
	dropout	随机失活率	0.3
	epoch	迭代次数	100
	kernel_size	卷积核窗口	3

3.4.3　知识抽取效果

为验证多维信息感知模型和改进 PCNN 关系抽取模型对经验反馈信息中的语义特征提取效果，在同一数据集下，选取几种知识挖掘模型进行实验，结果如表 4 所示。在经验反馈信息的实体识别任务中，本文使用的融合注意力机制的 BERT-BiLSTM-CRF 模型的准确率、召回率和 F1 值分别为 91.49%、93.14%、92.31%，模型识别效果较好。

根据表 5 可知，本文提出的改进 PCNN 关系抽取模型的准确率、召回率和F1值均比PCNN

表 4　多维信息感知模型性能对比结果

Tab.4　Performance comparison results of multidimensional information perception models　%

多维信息感知模型	准确率 P	召回率 R	F1 值
BiLSTM-CRF	83.65	83.02	83.33
BERT-CRF	90.67	91.78	91.18
BERT-BiLSTM-Attention-CRF	91.49	93.14	92.31

表 5　关系抽取模型性能对比结果

Tab.5　Performance comparison results of relation extraction model　%

关系抽取模型	准确率 P	召回率 R	F1 值
PCNN	92.37	92.54	91.84
改进 PCNN	93.18	93.36	92.87

的大，能更好地识别核电质量文本的关系特征，为构建完整的经验反馈图谱提供数据支撑。

3.4.4　经验反馈图谱生成

通过多维级联知识抽取模型提取经验反馈信息后，原始的大量非结构化质量事件文本转化成了结构化的经验反馈关键域。在此基础上使用 Neo4j 图数据库构建知识图谱中的节点和边，完成本体模型映射和知识规范存储，实现经验反馈知识的可视化集成，形成核电建造经验反馈图谱(图 8)。

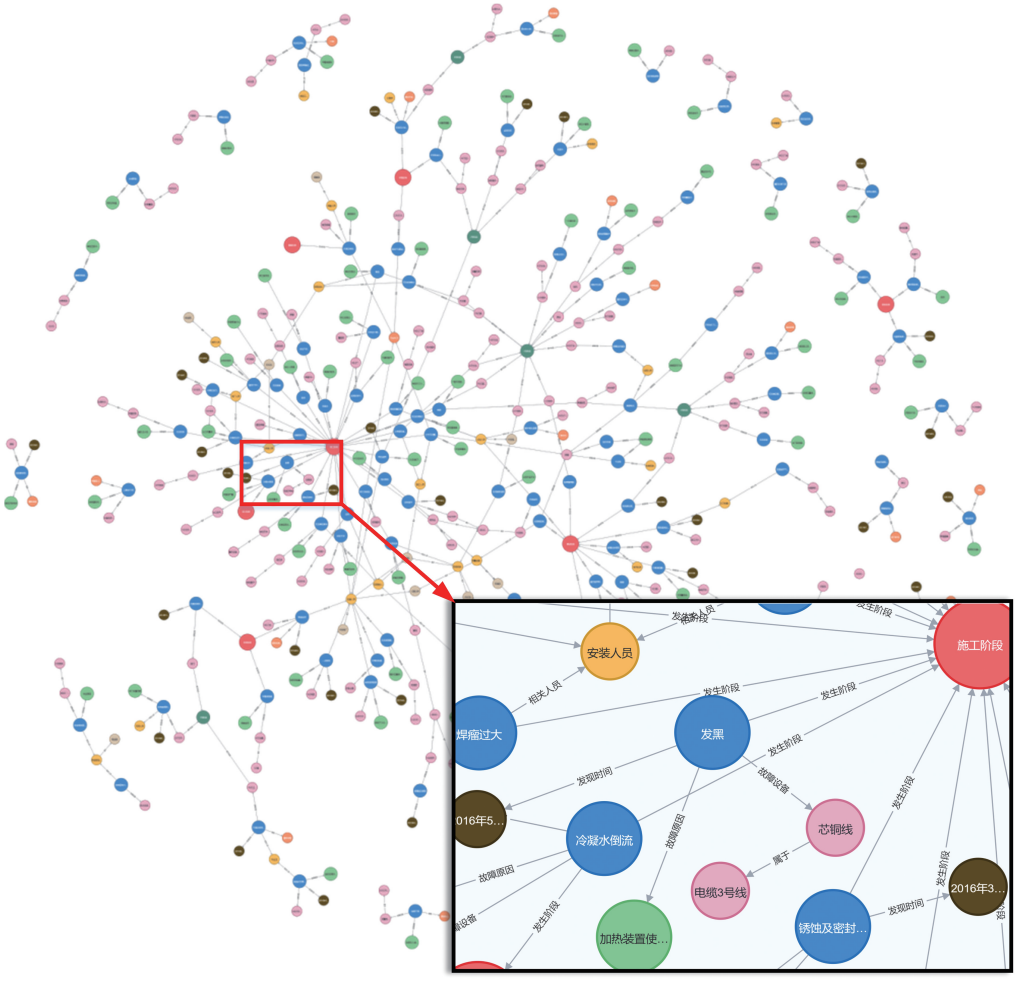


图 8　核电建造经验反馈图谱

Fig.8　The experience feedback graph of nuclear power construction

3.5 基于经验反馈图谱的深度知识挖掘

基于 Neo4j 图数据库生成的经验反馈图谱使用 Cypher 语言指令进行知识深度挖掘与应用。

3.5.1 质量问题处理措施的查询与复用

通过经验反馈图谱检索特定知识库，能辅助相关人员针对性查询具体质量问题的处理措施，如在核电建造过程中发生齿轮碰摩的质量事件，则可在经验反馈图谱中通过指令

```
MATCH (n)-[r]-()
WHERE n.gz CONTAINS '碰摩'
RETURN n,r
```

检索过去的经验知识，获得有关“碰摩”的故障节点，展开节点可获得对应的经验反馈图谱和节点的统计信息，如图 9 所示。

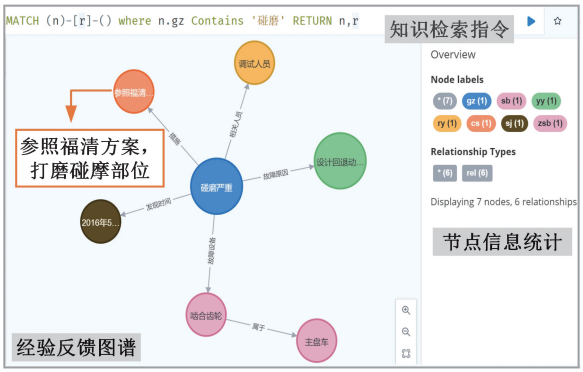


图 9 质量问题关联示例

Fig.9 Example of quality issue correlation

与文字记录质量事件的传统方式相比，统一规范的图结构更利于管理人员组织和管理经验反馈知识，所有的质量事件记录都按相同的结构和规范进行组织和存储，能确保各类复杂信息都以一致的方式进行表达。非专业人员无需阅读大量的文字文档也能通过浏览知识节点和连接关系，快速理解经验反馈信息中的重点内容，获取质量问题的处理措施及其他相关知识。经验反馈图谱还能避免不同人或团队对不同行文风格的经验反馈信息产生认知偏差，提高经验反馈信息的可理解性和可复用性，进而提高经验反馈的效率和质量。

3.5.2 质量问题关联实体的查询与推送

核电建造过程中，经验反馈工程师并非档案专业人员，信息检索能力有限，往往以关键词查询资源，查全率和查准率低、颗粒度大，这导致经验反馈信息的隐性知识难以发掘、传递滞后。基于经验反馈图谱的关联知识检索突破了传统的关键词检索方式，管理人员可根据质量问题节点和关联关系等条件快速查询特定质量问题相关的各种实体。例如，设备所属机组和故障发生阶段之间原本没有直接关系，但利用知识图谱多跳检索方式可查询到实体间隐含的路径关联，输入指令

```
MATCH q=(a)-[*1..3]-n
WHRER n.jd CONTAINS '施工阶段'
AND a.jz CONTAINS '3 号机组'
RETURN q
```

在图数据库中进行多级跳跃查询到“3 号机组”节点与“施工阶段”节点的间接关系，结果如图 10 所示。

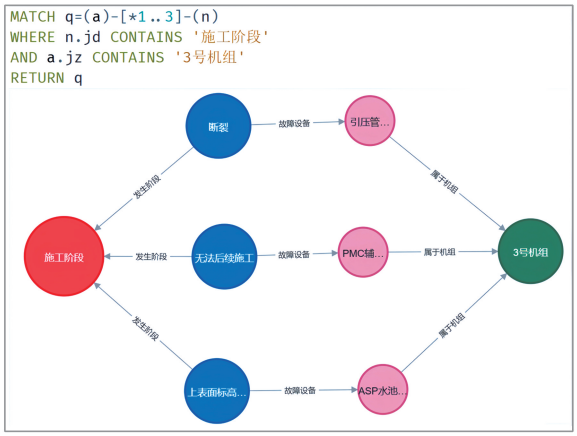


图 10 实体关联知识

Fig.10 Entity-associated knowledge

由图 10 易知，3 号机组在施工阶段有 3 台设备发生过不同类型的质量问题，在 Neo4j 图数据库中进一步展开设备类型的节点可获得对应的故障原因、人员等信息，能辅助管理人员明确企业间相关质量业务的接口和责任，从而向对应的部门反馈。相关人员在调试阶段需要重点关注这些设备的质量状态，降低相同设备相同质量问题二次发生的可能性。将查询结果推送到报告生成工具、数据分析平台等进行进一步的数据分析，有助于了解质量问题的影响范围和相关实体的状况，支持决策制定和改进措施的实施，使协同质量管控更精准、全面与高效。

3.5.3 经验反馈谱系知识统计与预防

通过聚类经验反馈图谱中记录的同一标签类，可得到用于表征特定类型知识的经验反馈知识谱系。基于经验反馈知识谱系，各类经验知识可被快速访问和检索，经验反馈工程师无需花费大量时间和精力去查找和收集。例如，输入指令

```
MATCH (n:gz) RETURN n
```

可得故障知识谱系，如图 11 所示。该指令能统计项目中质量事件发生次数，分析发生次数最多的同类型故障，有利于识别潜在的风险和预警信号，并采取相应的预防措施，减少质量问题的发生。

基于经验反馈谱系的关键知识管理与统计，可为用户提供强大的搜索和过滤功能，有助于发



图 11　故障知识谱系

Fig.11　The fault knowledge spectrum

现核电建造过程中存在的问题和数量趋势，帮助管理者识别常见的质量事件、故障模式和风险因素，辅助决策者和团队更好地了解和应对问题。

4　案例研究

基于上述研究，将经验反馈图谱和深度知识挖掘功能集成到核电企业的经验反馈系统。该系统具有原始文本记录、图谱展示、关联检索和数据分析等功能，辅助管理人员进行核电建造决策，将以人工经验为主的决策转变为知识驱动的智能决策，有效提高核电建造过程中经验反馈信息的可解释性。

反应堆压力容器（RPV）是核电主设备之一，零部件较多，其质量对核电站安全至关重要。安装或调试过程中，RPV 某零部件发生质量问题时，可通过现场人员记录至项目经验反馈系统，通过多维级联知识抽取方式智能解析文本库，在“图谱展示”功能模块中实时生成可视化的经验反馈图谱。基于统一规范的图谱，相关管理人员可通过推理引擎查询指令，在“关联检索”功能模块中针对反应堆压力容器这一设备进行深度信息挖掘，得到 RPV 故障发生的关联知识，如图 12 所示。



图 12　RPV 关联知识检索

Fig.12　Retrieval of related knowledge about RPV

管理者可从 RPV 经验反馈图谱快速了解结构化的关联知识，即在该项目历史记录的质量事件中，属于 3 号机组和 4 号机组的保温支承体、顶盖组件、接管座等 RPV 的关联设备（含子设备）发生过 5 种不同类型的故障，进一步可获得故障的根本原因和对应措施。通过“数据分析”功能模块可得到故障发生阶段的统计信息、常发生质量问题的零部件类别、过去质量事件的应对措施及涉及人员，以此生成统计报告并反馈给相关企业或部门，达到经验反馈信息精准传递的目的，促进高效、科学、综合地规划和制定核电装备建设方案，确保核电装备协同质量管控的高效与安全，具体应用流程见图 13。

5　结论与展望

本文根据核电装备的协同质量管控需求，针对企业间经验反馈时效性弱、各方经验反馈知识认知不统一、经验反馈信息复用率低等问题，在质量事件文本数据的基础上构建了统一数据模式的经验反馈图谱。分析了核电建造过程中经验反馈信息的特点、人工处理为主方式面临的困境，提出将经验反馈信息知识化处理的思路，详细论述了经验反馈关键知识域的构建方案。针对经验反馈图谱的数据挖掘任务，设计了多维信息感知模型，

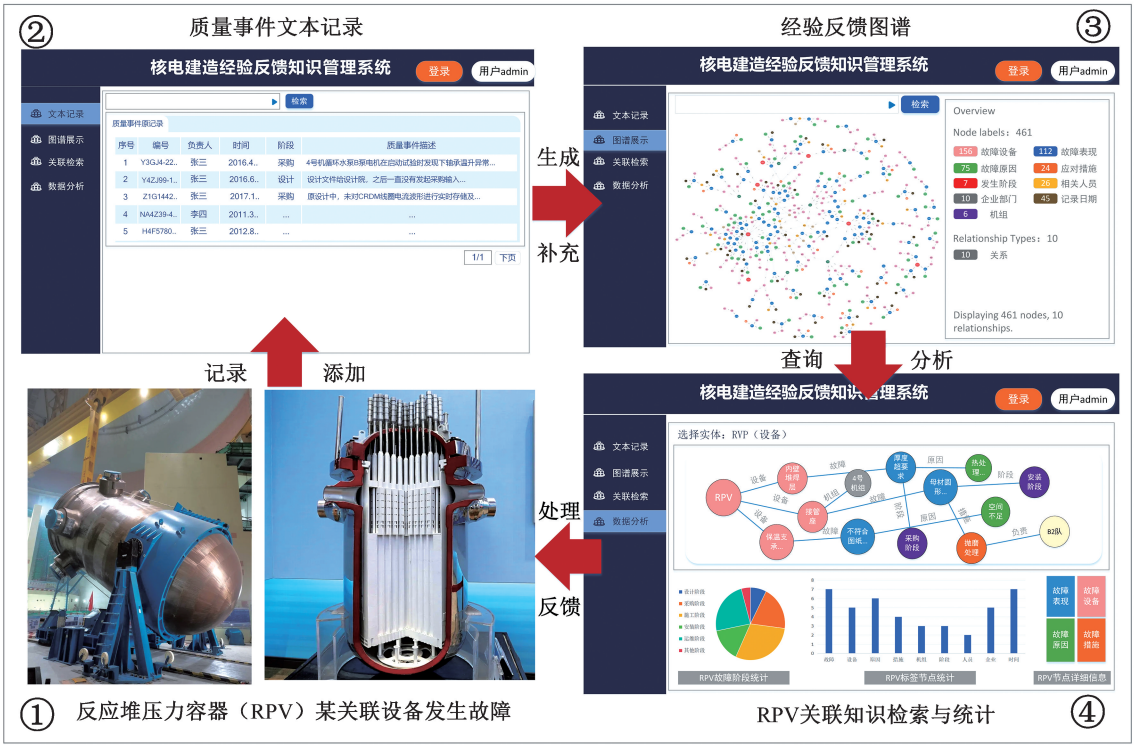


图 13　经验反馈图谱应用流程

Fig.13　Application process of experience feedback graph

利用上下文和重点语义特征提高了命名实体识别和关系抽取技术在核电领域的知识挖掘效果。该技术在某核电建造企业实际产生的经验反馈信息上得到验证，形成了可视化图谱，并利用图谱进行了以反应堆压力容器为对象的深度知识挖掘，为相关质保人员提供了智能化、精细化的质量知识管理工具并能够辅助管理决策，有助于核电企业深度复用质量数据资源，打破异地分布式企业之间的信息隔阂，提高核电建造的质量和效率。

未来将继续优化经验反馈图谱构建过程，扩充语料数据并动态更新知识图谱，开发基于经验反馈图谱的智能问答系统，优化知识检索方式。

参考文献：

[1] 赵栋，景晓冬. 工程经验反馈推荐系统在核电设备设计中的应用研究[J]. 制造业自动化，2017，39(10)：123-127.
ZHAO Dong，JING Xiaodong. Research of the Engineering Experience Feedback Recommendation System in the Equipment Design of Nuclear Power [J]. Manufacturing Automation，2017，39(10)：123-127.

[2] 邹象，张浩，雷蕾，等. 基于层次分析法的核动力厂安全相关经验反馈信息重要性评价方法研究与应用[J]. 核动力工程，2023，44(3)：189-195.
ZOU Xiang，ZHANG Hao，LEI Lei，et al. Research and Application of Significance Determination for Safety-related Feedback Information Based on Analytic Hierarchy Process in Nuclear Power Plants [J]. Nuclear Power Engineering，2023，44(3)：189-195.

[3] SINGHAL A. Introducing the Knowledge Graph：Things，not Strings[EB/OL].[2012-05-16]. https：// blog. google/products/search/introducing-knowledge-graph-things-not/.

[4] 吴信东. 知识图谱[M]. 北京：科学出版社，2022.
WU Xindong. Knowledge Graph[M]. Beijing：Science Press，2022.

[5] CHOI Y，NGUYEN M D，KERR T N Jr. Syntactic and Semantic Information Extraction from NPP Procedures Utilizing Natural Language Processing Integrated with Rules[J]. Nuclear Engineering and Technology，2021，53(3)：866-878.

[6] 邓志光，吴茜，朱加良，等. 基于改进 LSTM 的核电厂传感器故障诊断研究[J]. 自动化仪表，2023，44(6)：115-120.
DENG Zhiguang，WU Qian，ZHU Jialiang，et al. Research on Sensor Fault Diagnosis in Nuclear Power Plants Based on Improved LSTM[J]. Process Automation Instrumentation，2023，44(6)：115-120.

[7] 吴庭伟，王梦灵，易树平，等. 多尺度核电质量文本故障信息语义抽取方法[J]. 中国机械工程，2023，34(8)：976-981.

WU Tingwei, WANG Mengling, YI Shuping, et al. Semantic Extraction Method of Multi-scale Nuclear Power Quality Text Fault Information[J]. China Mechanical Engineering, 2023, 34(8):976-981.

[8] ZHAO Yunfei, DIAO Xiaoxu, HUANG J, et al. Automated Identification of Causal Relationships in Nuclear Power Plant Event Reports[J]. Nuclear Technology, 2019, 205(8):1021-1034.

[9] 刘晓瑛.基于知识图谱的核电设备健康管理系统设计与实现[D].杭州:浙江大学,2022.
LIU Xiaoying. Design and Implementation of a Nuclear Power Equipment Health Management System Based on Knowledge Graph[D]. Hangzhou: Zhejiang University, 2022.

[10] 冯毅雄，赵泽田，胡炳涛，等. 核电装备全生命周期价值链协同平台研究[J]. 计算机集成制造系统，2022，28(9):2633-2649.
FENG Yixiong, ZHAO Zetian, HU Bingtao, et al. Collaborative Platform of Lifecycle Value Chain for Nuclear Power Equipment[J]. Computer Integrated Manufacturing Systems, 2022, 28(9):2633-2649.

[11] 胡炳涛，冯毅雄，密尚华，等. 面向核电装备的全生命周期价值链协同模式研究[J]. 机械工程学报，2022，58(13):213-227.
HU Bingtao, FENG Yixiong, MI Shanghua, et al. Novel Collaborative Mode of Lifecycle Value Chain for Nuclear Power Equipment[J]. Journal of Mechanical Engineering, 2022, 58(13):213-227.]

[12] 雷玮剑，单玉忠，王理，等. 核电站建安调试文本语料抽取与语义图谱构建技术研究[J]. 信息技术与信息化，2024(1):6-12.
LEI Weijian, SHAN Yuzhong, WANG Li, et al. Research on Corpus Extraction and Semantic Map Construction Technology of Nuclear Power Plant Jian'an Debugging Text[J]. Information Technology and Informatization, 2024(1):6-12.

[13] 咎云龙.质量管理[M].北京:原子能出版社,2002.
ZAN Yunlong. Quality Management[M]. Beijing: Atomic Energy Press,2002.

[14] 汪南. 基于风险管理的 H 核电项目调试进度控制方案研究[D]. 大连:大连理工大学，2020.
WANG Nan. Research on Control Scheme of Commissioning Schedule of H Nuclear Power Project Based on Risk Management[D]. Dalian: Dalian University of Technology, 2020.

[15] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[J]. arXiv, 2018: 1810.04805.

[16] HOCHREITER S, SCHMIDHUBER J. Long Short-term Memory[J]. Neural Computation, 1997, 9(8):1735-1780.

[17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need[C]// Proceedings of the 31st Conference on Neural Information Processing Systems. Long Beach, 2017:6000-6010.

[18] ZENG D, LIU KCHEN Y, et al. Distant Supervision for Relation Extraction via Piecewise Convolutona Neural NetworksiCW/ACL[C]// Procedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, 2015: 1753-1762.

[19] HENDRYCKS D, GIMPEL K. Gaussian Error Linear Units (GELUs)[J]. ArXiv, 2016: 1606.08415.

（编辑　张　洋）

作者简介：李池璐，女，1999 年生，硕士研究生。研究方向为知识图谱、质量管理。E-mail:chichi@stu.cqu.edu.cn。**熊世权***（通信作者），男，1975 年生，博士、讲师。研究方向为质量管理、智能制造、人因工程。发表论文 20 余篇。E-mail:xiongshiquan@cqu.edu.cn。

本文引用格式：
李池璐，熊世权，雷玮剑，等.基于多维级联知识抽取的核电建造经验反馈图谱研究[J]. 中国机械工程，2025，36(1):177-189.
LI Chilu, XIONG Shiquan, LEI Weijian, et al. Research on Experience Feedback Graph of Nuclear Power Construction Based on Multi-dimensional Cascading Knowledge Extraction[J]. China Mechanical Engineering, 2025, 36(1):177-189.