

基于知识图谱的油田集输与注水系统能耗异常智能辅助决策方法

王文君^{1,2}, 陈由旺³, 朱英如³, 贺思宸¹, 刘珈铨¹, 张鑫儒¹, 王敏聪¹, 侯磊^{1*}, 王伟⁴

1 中国石油大学(北京)机械与储运工程学院, 北京 102249

2 国家石油天然气管网集团有限公司油气调控中心, 北京 100013

3 中国石油天然气股份有限公司规划总院, 北京 100083

4 中海油能源发展股份有限公司工程技术分公司, 天津 300450

* 通信作者, houleicup@126.com

收稿日期: 2024-07-02; 修回日期: 2024-10-21

中国石油天然气股份有限公司科技项目“油气田能量系统优化与能源管控研究”(2121DJ67)资助

摘要 随着油田能源系统的复杂性不断增加, 传统的监控、分析和优化的方法往往难以应对来自不同来源的大量数据, 导致在识别和解决能耗异常方面效率低下, 难以实现能源利用的最佳效果。为克服这些局限性, 实现油田集输与注水系统能源管控的智能决策, 本研究针对海量多源异构数据, 提出一种基于知识图谱的能耗异常智能辅助决策方法。以能耗异常台账及操作手册等文本资料为主要数据源, 建立能源管控知识内容体系框架, 作为组织和整合多源数据的基础, 确保数据的高效利用。采用 BiGRU-CRF 对文本资料进行实体抽取, 识别设备、参数、异常等关键概念; 采用 BiGRU-ATT 进行实体间关系抽取, 捕捉油田集输与注水系统中复杂的相互依赖性; 通过 Neo4j 图数据库对提取到的能耗知识进行存储和可视化展示, 其结构化的知识表示形式为后续数据的高效利用奠定了基础。根据所构建的知识图谱开发能源管控可视化平台, 提供用户友好的界面, 使操作人员能够以直观的方式探索能耗知识。该平台从数据和知识层面提供可行的措施推荐, 以辅助指导能耗控制。油田现场的应用结果表明, 采用所提出的基于知识图谱的油田集输与注水系统能耗异常智能辅助决策方法, 能够融合多源异构数据, 为工艺流程中发生的能耗异常事件提供及时性、整体性、智能性的辅助决策推荐, 指导操作人员进行快速有效的能耗控制, 显著减少了决策所需的时间。本研究为油田能源管控建设提供了新的思路, 对其它油田的能耗控制管理具有指导意义。

关键词 集输系统; 注水系统; 能耗异常; 知识图谱; 辅助决策

中图分类号: TE866; TE863

Intelligent assisted decision-making method for abnormal energy consumption of oilfield gathering and water injection system based on knowledge graph

WANG Wenjun^{1,2}, CHEN Youwang³, ZHU Yingru³, HE Sichen¹, LIU Jiaquan¹, ZHANG Xinru¹, WANG Mincong¹, HOU Lei¹, WANG Wei⁴

引用格式: 王文君, 陈由旺, 朱英如, 贺思宸, 刘珈铨, 张鑫儒, 王敏聪, 侯磊, 王伟. 基于知识图谱的油田集输与注水系统能耗异常智能辅助决策方法. 石油科学通报, 2025, 10(03): 620–632

WANG Wenjun, CHEN Youwang, ZHU Yingru, HE Sichen, LIU Jiaquan, ZHANG Xinru, WANG Mincong, HOU Lei, WANG Wei. Intelligent assisted decision-making method for abnormal energy consumption of oilfield gathering and water injection system based on knowledge graph. Petroleum Science Bulletin, 2025, 10(03): 620-632. DOI: 10.3969/j.issn.2096-1693.2025.02.007

1 College of Mechanical and Transportation Engineering, China University of Petroleum, Beijing 102249, China

2 PipeChina Oil & Gas Control Center, Beijing 100013, China

3 PetroChina Planning & Engineering Institute, Beijing 100083, China

4 CNOOC EnerTech-Drilling & Production Co., Tianjin 300450, China

Revived: 2024-07-02; Revised: 2024-10-21

Abstract The increasing complexity of energy systems in oilfields necessitates advanced approaches to monitor, analyze, and optimize energy usage. Traditional methods are often inadequate for processing the vast amounts of data generated from diverse sources, leading to inefficiencies in identifying and resolving energy consumption anomalies and making it difficult to achieve optimal energy utilization. To overcome these limitations and achieve the intelligent decision-making for energy management and control in oilfield gathering and water injection systems, an intelligent assisted decision-making method for abnormal energy consumption was proposed based on knowledge graph, addressing the challenges posed by massive multi-source heterogeneous data. Specifically, the abnormal energy consumption records and operation manuals were utilized as the primary data source, and the comprehensive knowledge framework for energy management and control was established. This framework serves as the foundation for organizing and integrating multi-source data, ensuring systematic and efficient data utilization. Additionally, the BiGRU-CRF (Bidirectional Gated Recurrent Unit-Conditional Random Field) model was applied to extract entities from the textual data, identifying key concepts such as equipment, parameters, and anomalies. And the BiGRU-ATT (Bidirectional Gated Recurrent Unit-Attention) model was adopted to extract relationships between entities, capturing the complex interdependencies within the oilfield gathering and injection systems. The extracted energy consumption knowledge is stored and visualized using the Neo4j graph database, providing a robust platform for data querying and analysis. Its structured representation lays the foundation for the efficient utilization of data in subsequent stages. Finally, based on the constructed knowledge graph, an energy management and control visualization platform was developed, providing a user-friendly interface that enables operators to explore energy consumption data and knowledge in an intuitive manner, significantly enhancing the usability of the operational system. The platform provides actionable recommendations at both the data and knowledge levels, supporting energy consumption control effectively. The field application results in oilfields demonstrate that the proposed intelligent decision-making method, based on knowledge graphs, effectively integrates multi-source heterogeneous data for abnormal energy consumption detection in oilfield gathering and injection systems. Timely, comprehensive, and intelligent decision-making recommendations are provided for energy consumption anomaly events in the gathering and injection processes, guiding operators in achieving rapid and effective energy consumption control. The time required for decision-making is significantly reduced through this method. This study offers a novel and impactful approach for the construction of energy management and control systems in oilfields, which provides valuable guidance for the management of abnormal energy consumption in other oilfields.

Keywords gathering and transportation system; water injection system; abnormal energy consumption; knowledge graph; assisted decision-making

DOI: 10.3969/j.issn.2096-1693.2025.02.007

0 引言

油田集输与注水系统包含多个工艺环节,系统庞大且分散。随着油田开发的不断推进,单井产量持续递减,含水率不断升高,泵机组等设备长年使用易发生腐蚀、老化现象,集输与注水系统存在能耗高、运行效率低、能源浪费大的问题。油田能源管控能够为此问题的解决提供新思路。能源管控是一种应用于能源生产、输配、消耗环节的集中化、扁平化、全局化的能源生产运行模式,旨在强化能源利用的有效管理与控制,实现能源利用的最佳效果。

当前油田集输与注水系统能源管控模式大多通过

对积累的现场运行数据进行挖掘,借助理论优化算法或工艺软件,辅以对现有管理文件的人工分析,进行能源优化管理。近年来,数据采集传感技术、大数据分析等新兴信息技术迅猛发展,油田能源管理体系持续优化,积累了海量的运行数据和繁杂的管理文件。从庞大、繁杂的管理文件中手动提取所需知识,协同运行数据完成决策需要耗费大量时间,难以满足油田能源实时优化控制的智能决策需求。由人工智能技术发展而来的知识图谱(Knowledge Graph, KG)是在大数据时代背景下产生的一种高效的知识表达形式,能够融合多源异构数据,为领域内研究提供整体性、关系链的参考^[1-3]。

作为一种知识表示的新方法和知识管理的新思路,知识图谱将研究领域多源异构的知识统一结构化并构建知识间的关联,解决数据多样复杂导致的孤岛化。本质上,知识图谱是一种由节点和边组成的语义网络(Semantic Network)^[4],含有丰富的语义知识。节点表示实体或概念,如某个设备或某个异常事件,节点间的连线表示不同实体或概念之间的关系,如包含、监测、推荐措施等。除了整合表达知识以外,知识图谱还具有强大的知识表示和推理能力,能够更好地组织、管理和理解海量信息。其特殊的图形化展示和推理检索方式,帮助使用者从人工过滤寻找答案的模式中解脱出来。目前,知识图谱已被广泛应用于医疗^[5-6]、交通^[7-8]、电力^[9-10]等领域,在石油与天然气行业也获得初步应用^[11-15],但大多集中于油田勘探开发以及油气管道风险评估,在油田地面集输与注水系统能源管控方面的研究仍是空白。

本研究针对某油田集输与注水系统能源管控问题,通过深度学习方法识别领域中的实体,抽取实体间关系,以抽取结果为基础,构建油田集输与注水系统能源管控知识图谱,形成基于知识图谱的能耗异常

智能辅助决策方法,为解决油田集输与注水系统能源管控智能决策面临的海量多源异构数据难题提供指导。

1 能源管控数据资料概述

本研究使用数据来源于某油田集输与注水系统近年来的能耗管理资料。该油田经过数年的油田数字化管理建设,目前已完成联合站、中转站、集油注配间和油水井的数字化,初步实现规模化数字化油田建设,在油田生产过程中采用数字化管理,已有较全面的数据基础。所用资料涵盖管网拓扑结构、设备资产数据等结构化数据以及能耗异常台账、运维手册等文本资料。

1.1 结构化数据

结构化数据包括管网拓扑结构以及设备资产情况,以表格形式存储。管网拓扑结构信息(表 1)包括管道名称、管道编号、管道代码、投用日期、起点站库类型、起点名称、终点站库类型、终点名称等。设备资

表 1 管道数据示例
Table 1 Pipeline data examples

管道信息	1-8#集油阀组间至卫 1-29-13 掺水管道	卫 1-29-13 至卫 1-28-14 集油管道
管道编码	1-8-4-1	1-8-4-2
管道代码	1 0015000100010100000 000000	0 10015000100010122000 263648
起点站库类型	阀组(间)	采油井
起点名称	1-8#集油阀组间	卫 1-29-13
终点站库类型	采油井	采油井
终点名称	卫 1-29-13	卫 1-28-14
所属环名称	1-8-4#环	1-8-4#环
投用日期	2009-11-01	2009-11-01
设计最大输量/ $\text{m}^3\cdot\text{d}^{-1}$	150	150
设计最小输量/ $\text{m}^3\cdot\text{d}^{-1}$	24	24
设计最大工作压力/MPa	3	2.5
长度/m	468.6	335.22
管径/mm	76	76
壁厚/mm	4.5	4.5
输送介质	含油污水	含水油
管道类型	掺水管道	出油管道
...

产包括站(间)内外的管道、掺水泵、掺水炉、注水泵等,例如,注水泵基础信息包括所属站名称、泵名称、站内编号、泵类型、投产日期、泵生产厂家、额定流量、额定转速、额定工作压力等。

1.2 非结构化数据

非结构化数据包括能耗异常台账、运维手册等。能耗异常台账是工作人员整理的自然文本,包含能耗异常事件发生信息及处理过程,通常由参数异常、事件定位、异常原因、解决措施4部分内容组成(表2)。参数异常指集输与注水系统中能耗参数异常事件;事件定位是参数异常事件发生的位置节点;异常原因包括设备本体的异常因素以及上下游节点的影响;解决措施指为缓解异常事件所需采取的针对性处置措施。运维手册介绍了生产系统总体架构和异常事件处理流程(表2)。

2 能源管控知识图谱的构建

2.1 构建框架

油田集输与注水系统能源管控知识图谱构建框架(图1)。知识图谱在逻辑上可分为模式层与数据层两个层次。模式层是知识图谱的知识组织架构,通过分析集输与注水系统能源管控领域文本内容,提炼实体

类型、关系类型等重要概念,形成领域知识体系。利用知识抽取方法,在模式层知识架构的指导下对实体、实体间关系以及实体属性进行抽取,得到能源管控知识图谱数据层。模式层提供规则指导数据层的构建。反之,数据层是模式层的实例化表示。二者相互结合得到油田集输与注水系统能源管控知识图谱。

2.2 知识图谱模式层

综合考虑油田集输与注水系统能源管控领域的数据特点和能耗异常情况下的操作人员的业务逻辑,将油田集输与注水系统能源管控知识图谱分为概念图谱和实体图谱,对图谱模式层进行设计(图2)。

概念图谱是能源管控知识图谱的核心内容,由能耗管控台账等非结构化文本以及部分结构化数据如设备资产数据构建得到。图谱将油田集输与注水系统能耗异常分析过程“参数异常—事件定位—异常原因—解决措施”以结构化的形式展示,主要包括参数异常等7种实体及实体间的关系,在节点间设置双向关系,即在“问题—定位—原因—措施”的推理基础上增加了反向推理机制,便于后续知识推理。

实体图谱表征实际油田集输与注水系统的管网拓扑结构,为概念图谱提供实体的查询和匹配。图谱以联合脱水站、转油站、站内设备、集油阀组间、注配间、采油井、注水井以及各类连接管道作为实体节点,设施间的连接情况用关系表示,井口和管道所属的集油环、设备的具体参数等信息以实体属性的形式存储其中。

2.3 知识图谱数据层

针对结构化和非结构化两种数据类型,需要采用

表2 文本示例

Table 2 Text examples

能耗异常 台账示例	系统监测发现卫1联脱水站来液汇管的来液压力持续升高。查阅运维手册,可能原因包括该站所辖集油阀组间的总回油压力过高、产液量增大、阀组间所辖集油环的环掺水量过大或管道憋压。经作业人员检查,发现管输流量不变,且来液压力已高于掺水压力,确定管道存在憋压的情况。按照操作规程,对管线进行冲洗除垢。
运维手册 示例	环的状态除了环本身以外,还受环所管辖的油井的状态所影响。一旦推送环告警信息,需要先排除环本身故障,再进行环所管辖油井的故障排除。环、井关系获取流程需要从三个数据表获得数据:环基础信息表、井基础信息表和集油间/站库基础信息表。三个数据表的数据经过一些列过滤、合并、汇总等关联操作之后,环信息将与其下辖的井信息进行一对多的对应,最后环井关系作为输出,在环告警流程中使用。

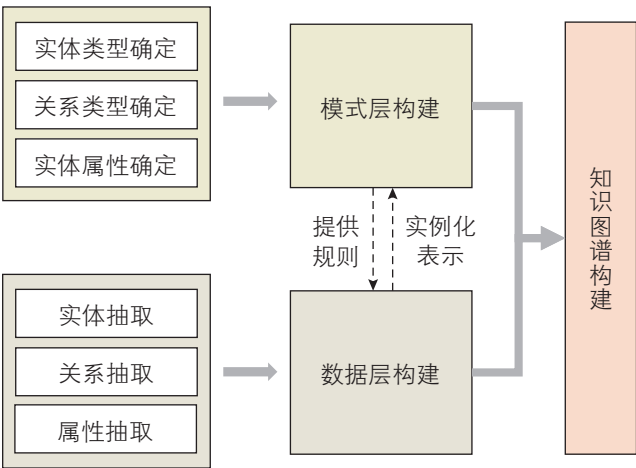


图1 知识图谱构建框架

Fig. 1 Construction framework of knowledge graph

不同的数据处理方式。对于油田集输与注水系统能源管控相关资料中的管网拓扑结构等结构化数据，能直接抽取其中的实体关系三元组和实体属性。抽取得到的管网实体包括联合/接转站、站(间)内位置等 15 个类型(表 3)。

对于能耗异常台账等文本类的非结构数据，需要通过命名实体识别和关系抽取将其转换为结构化数据。命名实体识别(Named Entity Recognition, NER)和关系抽取(Relation Extraction, RE)是知识抽取任务中的重要部分^[16]，旨在从半结构或非结构化的集输与注水领域文本数据中抽取实体关系三元组。

油田集输与注水系统能源管控领域文本中的专有名词多、知识结构复杂，并且实体多由位置、设施、参数变化情况等多种名词嵌套构成，实体界限较模糊。目前常用的中文分词工具适用于通用领域，应用在集输与注水领域会造成分词结果误差增加，影响后续实体间关系的抽取。本研究选择引入基于字向量表示的双向门控循环单元—条件随机场(BiGRU-CRF)模型对集输与注水领域中的实体进行识别。在实体识别的基础上，采用基于注意力机制的双向门控循环单元(BiGRU-ATT)模型对实体间关系进行抽取。将 BiGRU 能够考虑字符的上下文语义信息的优势，分别与 CRF

保证输出结果合理性的特点和 Attention 机制能够提高模型关注度的特点相结合，提高对油气集输与注水系统能源管控领域文本知识的抽取效果。

(1)基于 BiGRU-CRF 的命名实体识别

命名实体识别本质上是一个序列标注问题，即对给定文本中的每一个字符指定标签^[17-18]。本研究选择采用 BIO(B-begin, I-inside, O-outside)序列标注方法，BIO 标注规则规定实体一定是 B 开头、I 结尾，O 表示不属于任何实体。

针对所研究的油田集输与注水系统能源管控领域实体识别任务的实体界限模糊、分词困难的问题，在利用 BiGRU-CRF 模型(图 3)进行实体抽取时，采用单个字符作为模型的输入方式，避免分词误差对模型的影响。通过字转换向量模型技术将中文字符处理成向量表示形式，把字向量输入到 BiGRU 模型进行深层次的特征提取，输出得到初步的标签序列。BiGRU 模型能同时对预测词的上下文信息进行处理，提高模型对文本特征的保留程度，达到增强预测结果准确性的效果，且计算代价更小，收敛性更好。初步得到的输出标签需通过 CRF 层^[19]进一步修正。CRF 层通过定义特征函数对输出的实体标签加以约束，解决了由于 BiGRU 未考虑输出标签之间相互约束性可能输出无效标签的问题，提高输出结果的准确性。

(2)基于 BiGRU-ATT 的关系抽取

实体间关系的抽取以命名实体识别为基础的，高度依赖上下文信息，其目的在于从非结构化文本中抽取得到实体关系三元组。考虑到句子中不同成分信息对关系分类结果影响不同，本研究在 BiGRU 模型的基础上，将注意力机制(Attention)引入关系抽取任务中，充分学习文本上下文信息。

注意力机制是深度学习中模仿人脑的注意力模型，通过给予文本中部分关键词更多关注达到提高特征

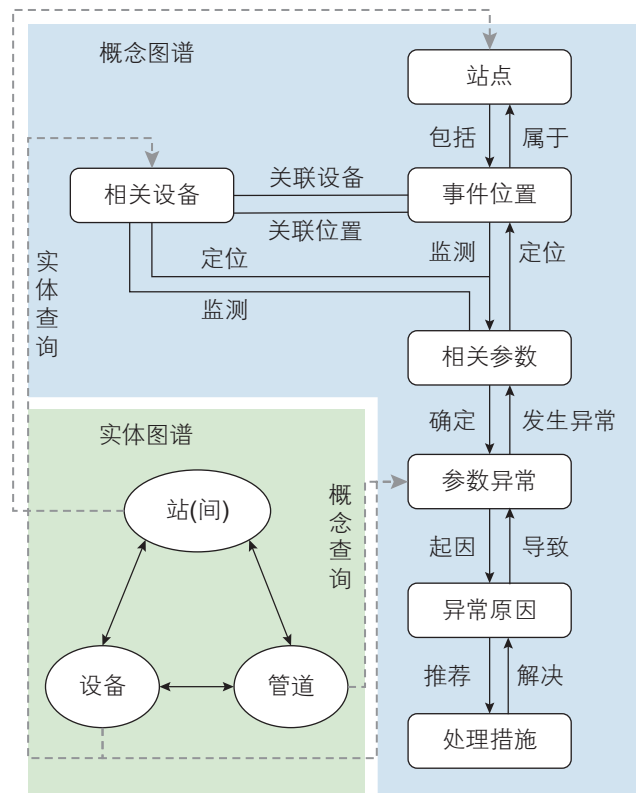


图 2 知识图谱模式层设计
Fig. 2 Scheme layer of knowledge graph

表 3 管网实体分类情况

Table 3 Classification of pipeline network entities			
实体类型	实体数量	实体类型	实体数量
联合/接转站	3	集油管道	652
站(间)内位置	22	采油井	607
站(间)内设备	40	注配间	22
输油管道	4	供水管道	45
站间掺水管线	26	注水管道	332
站间回油管线	23	注水井	332
集油间	22	其它	85
掺水管道	207		

提取质量的目的^[20-21]。BiGRU-ATT模型(图4)以集输与注水系统能耗相关文本和命名实体识别结果为基础,利用BiGRU模型能够同时正向和反向传播的优点获取文本中的上下文信息,结合注意力机制对信息输

出加以优化。注意力层的输入矩阵对应为BiGRU输出矩阵,根据向量特征间的相对距离确定向量间的相似性,计算得到每个向量的注意力权重。赋予句子中相似性较高、对关系分类影响较大的部分信息较高的权

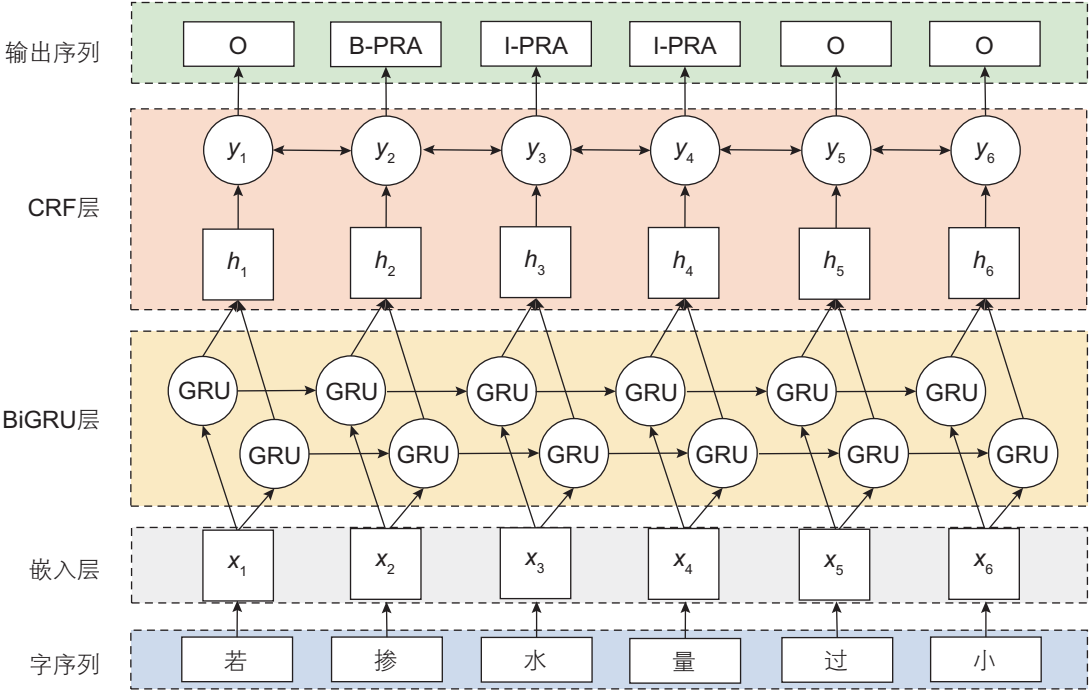


图3 BiGRU-CRF命名实体识别模型
Fig. 3 BiGRU-CRF named entity recognition model

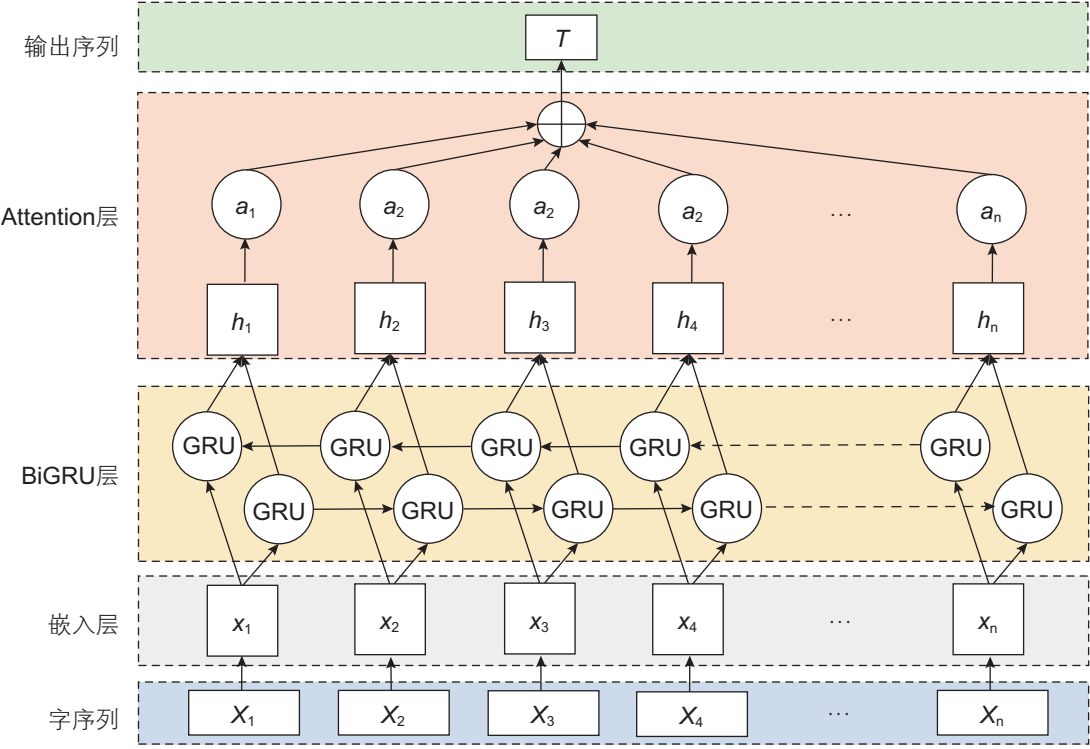


图4 BiGRU-ATT关系抽取模型
Fig. 4 BiGRU-ATT relation extraction model

重，加权求和生成用于分类整体特征以作为模型分类的额外输入，提高集输与注水系统能耗异常事件因果关系抽取效果。根据输出结果判断实体间是否存在预定义的关系，实现实体间关系的抽取，得到关系三元组(表 4)。

3 能源管控知识图谱的可视化与应用

3.1 知识图谱可视化

油田集输与注水系统能源管控领域文本数据经知识抽取后，共形成 3263 个节点和 5034 条关系，需要进行存储管理。典型的知识图谱存储管理方式分为基于关系数据库的存储和基于原生图的存储^[22]。本研究的知识图谱使用 Neo4j 图数据库实现。Neo4j 采用的基于原生图的知识图谱存储管理方式，是一个开源的基于 java 的高性能的 NoSQL 图数据库^[23-24]。与传统的关系型数据库不同，图数据库将结构化数据以“节点—关系—节点”的三元组形式可视化存储在网络中。特殊的存储结构使其能够高效处理复杂的知识图谱查询，更适合表征集输与注水领域文本这类具有复杂关联的知识。

鉴于涉及的数据量较大，在知识存储和图谱构建时选择采用 Py2neo+Neo4j 的方式批量创建节点、关系及属性。Py2neo 是 Python 中面向 Neo4j 的一个库，用于从 Python 应用程序和命令行中调用 Neo4j 图数据库完成知识的存储与查询，提高了 Neo4j 使用的便捷程

度。构建得到油田集输与注水系统能源管控知识实体图谱和概念图谱(图 5-6)，图谱中针对不同类别节点采用不同颜色加以区分。

图 5 是与 2-2# 集油间相关的部分实体图谱，展示有该间所辖的 10 个集油环和 34 口油井，以及所属站场的相关信息。图 6 是与集油间相关的部分概念图谱，共有 station、location、equipment、parameter、variation、reason、measurement 7 种实体类型(表 5)。绿色实体表示集输与注水系统中的站场类型(station)，如卫 2 转油站、集油阀组间等。深蓝色实体表示事件位置(location)，即可能发生异常事件的位置节点，如集油阀组间集油汇管、集油阀组间掺水汇管等。浅蓝色实体表示本体设备(equipment)，即工艺流程中的主要工艺设备，如掺水泵、掺水炉等。紫色实体表示相关参数(parameter)，即事件位置或本体设备所监测的工艺参数，如总回油压力、掺水出站温度等。红色实体表示参数异常(variation)，即监测能耗参数可能发生的异常情况，如总掺水压力降低、环回油温度低等。浅棕色实体表示异常原因(reason)，即可能导致某个参数发生异常的原因，如总回油压力升高的原因可能是产液量增大、集油环掺水量过大、站间回油管线缩径或站间回油管线保温效果差。黄色实体表示解决措施(measurement)，即为解决参数异常，根据异常原因需采取针对性处置措施，如缓解站间管线缩径的措施有对管线冲洗、除垢或更换。

表 4 抽取得到的关系三元组(部分)
Table 4 Extracted relation triples(part)

实体 1	关系	实体 2	原句
集油阀组间	包括	集油阀组间集油汇管	集油阀组间容易发生异常事件的位置包括集油阀组间集油汇管和集油阀组件掺水汇管
集油阀组间	包括	集油阀组间掺水汇管	集油阀组间容易发生异常事件的位置包括集油阀组间集油汇管和集油阀组件掺水汇管
掺水量增大	导致	掺水出站温度降低	掺水量增大、进气量变小以及各单台掺水炉出口温度差异大都可能导致掺水出站温度升高
进气量变小	导致	掺水出站温度降低	掺水量增大、进气量变小以及各单台掺水炉出口温度差异大都可能导致掺水出站温度升高
各单台掺水炉出口温度差异大	导致	掺水出站温度降低	掺水量增大、进气量变小以及各单台掺水炉出口温度差异大都可能导致掺水出站温度升高
来水压力	定位	来水汇管	通过来水压力定位到来水汇管这一事件位置上
注水泵	监测	注水泵进口汇管压力	针对注水泵主要监测注水泵进口汇管压力和注水泵出口汇管压力两个参数
注水泵	监测	注水泵出口汇管压力	针对注水泵主要监测注水泵进口汇管压力和注水泵出口汇管压力两个参数

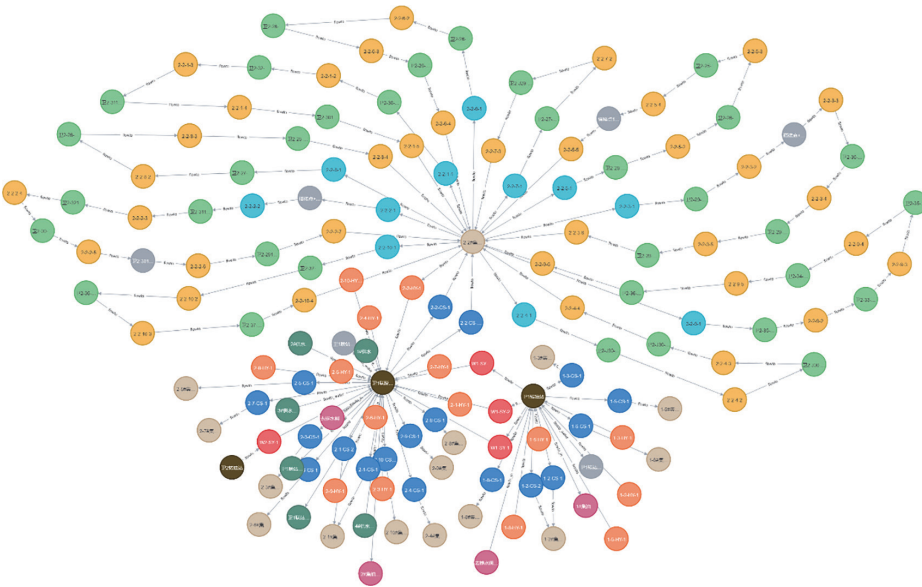


图 5 实体图谱(部分)
Fig. 5 Graph for pipe network(part)

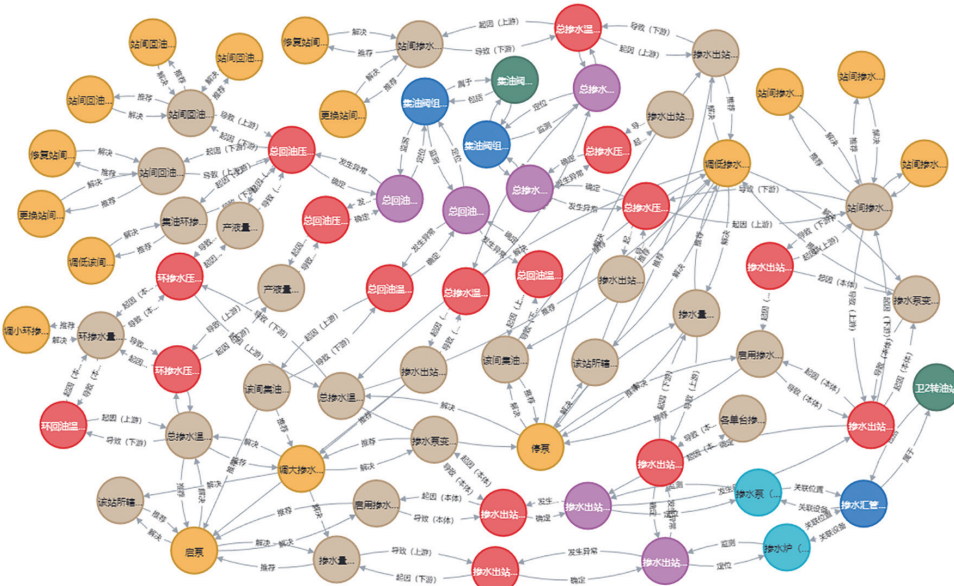


图 6 概念图谱(部分)
Fig. 6 Graph for concepts(part)

表 5 概念图谱实体情况
Table 5 Entity situation of concept graph

实体类型	实体含义	实体数量
station	站场	6
location	事件位置	32
equipment	本体设备	40
parameter	相关参数	67
variation	参数异常	134
reason	异常原因	283
measurement	解决措施	279

3.2 知识图谱应用

考虑油田能源管控任务实际场景,应用知识图谱进行能耗异常诊断(图 7)。针对领域文本数据进行知识抽取,将抽取到的三元组结果存入Neo4j图数据库,构建油田集输与注水系统能源管控知识图谱,基于知识图谱搭建能源管控可视化平台并接入现场生产管理系统。当油田集输系统发生能耗异常时,由生产管理系统返回异常告警信息。解析告警信息,根据告警位置在实体图谱中查找相应实体节点,结合概念图谱返

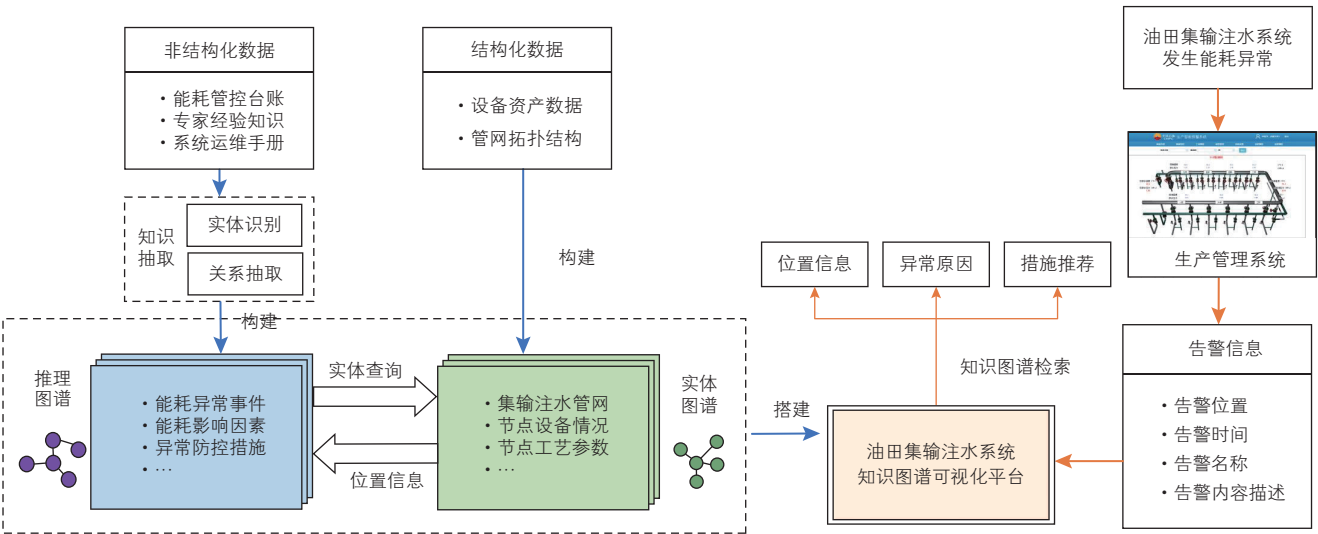


图 7 能源管控知识图谱应用流程

Fig. 7 Application process of energy management and control knowledge graph

回该节点监测参数发生异常的原因及相应的处置措施。

与从运行数据层面分析的传统管控手段相比，本研究提出的基于知识图谱的油田集输与注水系统能耗异常智能辅助决策方法，在知识表达与查询方面更加高效，能够提供数据和知识层面全面、准确的能耗控制指导，在优化运行工艺流程的同时提高管控效率和管控质量。

以业务需求为导向，设计油田集输与注水系统能源管控知识图谱可视化平台，在 windows10 下的 Python3.6 开发环境中完成搭建，包括前端和后端两个方面的实现。前端主要采用 Bootstrap 框架以及 HTML、CSS、JavaScript 等技术进行前端编码，完成对于可视化平台界面的编码及展示。后端采用 FLASK 框架实现知识检索等业务逻辑的构建和前后端的交互。在 FLASK 对象的 route 方法中，通过 GET 或 POST 方法向服务器发送请求。利用 request.args.get 获取前端查询请求 (GET 请求) 数据后，通过 Py2neo 接口访问 Neo4j 图数据库实时获取请求相关的节点及关系；或利用事先从图数据库中提取的 JSON 格式数据，在 Neo4j 离线状态下完成实体查询，最后将结果用 jsonify 返回给前端展示。

通过油田集输与注水系统能源管控可视化平台，能够便捷地实现智能信息检索、异常原因推理与决策推荐功能，平台展示包括概念图谱和实体图谱 (图 8)。

(1) 智能信息检索

知识图谱通过整合不同来源的数据资料，以图的形式对能耗异常相关知识之间的语义关系进行表示。相较于没有明确关系概念的传统信息存储与检索方式，

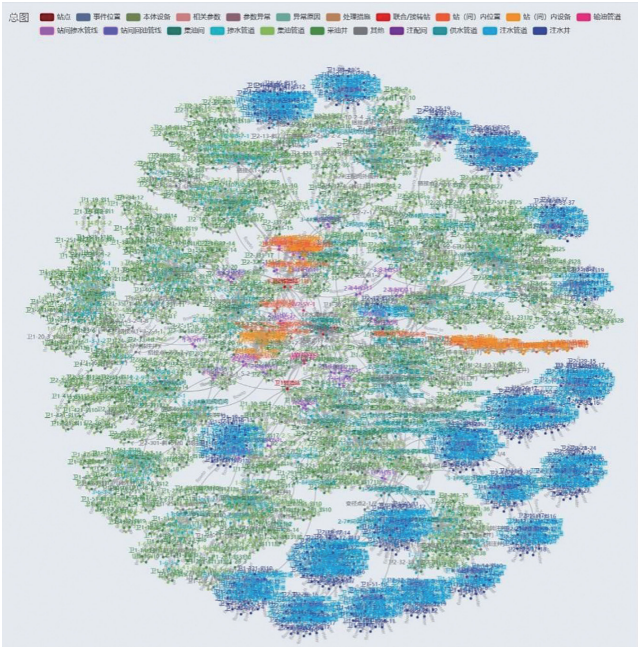


图 8 概念图谱和实体图谱

Fig. 8 Graph for concepts and pipe network in platform

知识图谱能够直接存储两个实体之间的关系，更直观灵活，有助于对知识的查询。当查询实体的某个特定关系对象时，只用查询该实体就能沿着“边”直接获取相关对象，避免传统关系检索过程中的各种表关联操作。

平台以知识图谱为基础，对使用者在搜索框内输入的关键字 (或点击页面展示标签) 进行解析，将其映射到具体的概念或实体上，返回所有与其相关的信息，实现智能信息检索。智能信息检索包括能耗概念关系

和管网实体关系检索，分别基于概念图谱和实体图谱搭建而成，能够实现对集输与注水系统能耗概念和管网实体的查询。

图 9-10 分别展示“注配间”这一概念以及“1-2#注配间”这一具体节点在概念图谱和实体图谱中的相关查询结果。根据查询结果显示，注配间内可能发生异常事件的位置有来水汇管、注水阀组、注水泵进口汇管和注水泵出口汇管，1-2#注配间的上游为1-2#间进间供水管道，下游为卫1-27-3等14条注水管道。

(2)异常原因推理与决策推荐

根据所构建的油田集输与注水系统能源管控知识图谱，能够建立能耗异常的关系链条。当监测到参数

波动异常时，通过关系链条进行路径分析，定位可能因素，从而针对性地采取安全措施控制能耗因素，避免参数持续波动造成大量能源浪费，或导致生产事故。平台提供实体图谱与概念图谱间的关系查询，输入管网实体节点可以查询该节点相关的监测参数，点击页面查询结果即跳转至“检索能耗概念关系”或“检索管网实体关系”页面进一步检索。

现场监测发现 1-2#注配间内注水泵出口汇管压力持续升高后，根据异常告警信息，通过平台检索确定其原因推理流程(图 11)。节点位置参数既可能因其自身情况变化导致，也可能受其上、下游节点的波动影响。图中的上、下游原因是以注水泵出口汇管作为本体并结合实际工艺流程确定的，注水泵出口汇管的上

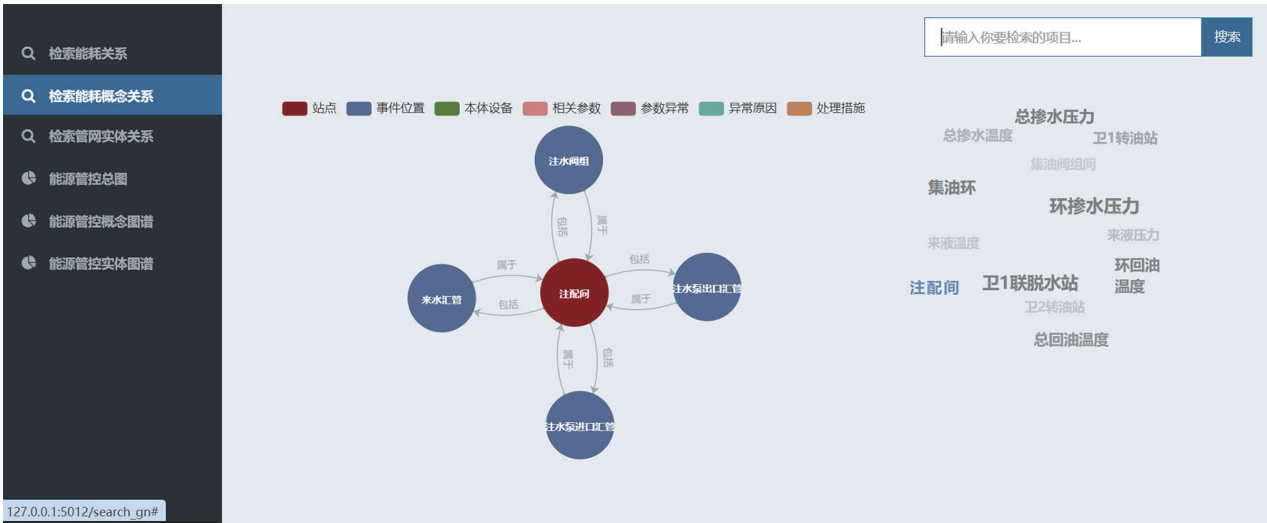


图 9 “注配间”概念关系查询结果

Fig. 9 Query results of the conceptual relationship between ‘injection room’

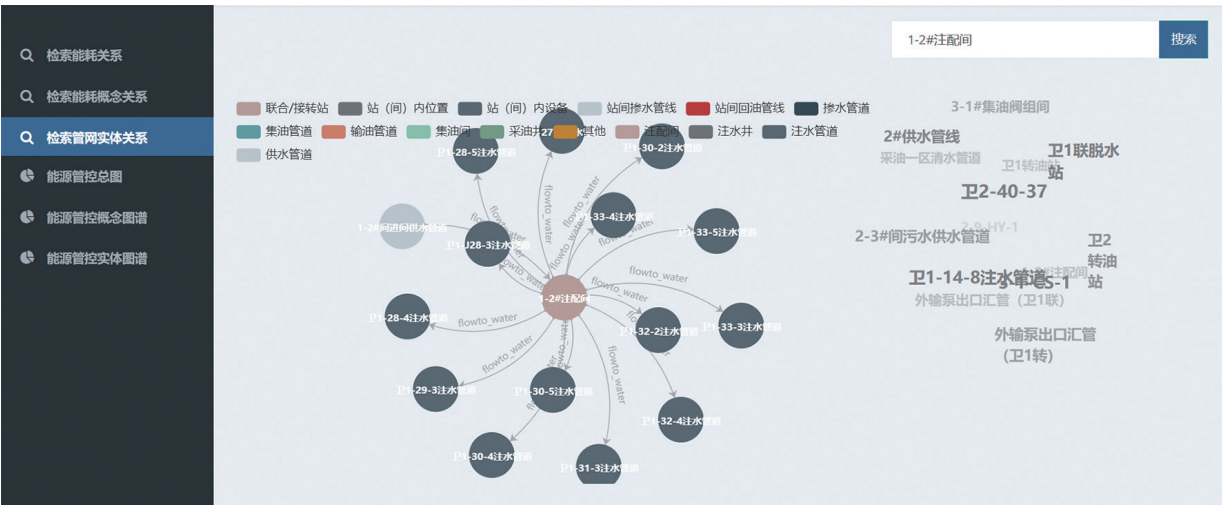


图 10 “1-2#注配间”实体关系查询结果

Fig. 10 Query results of the entity relationship between ‘1-2 # injection room’

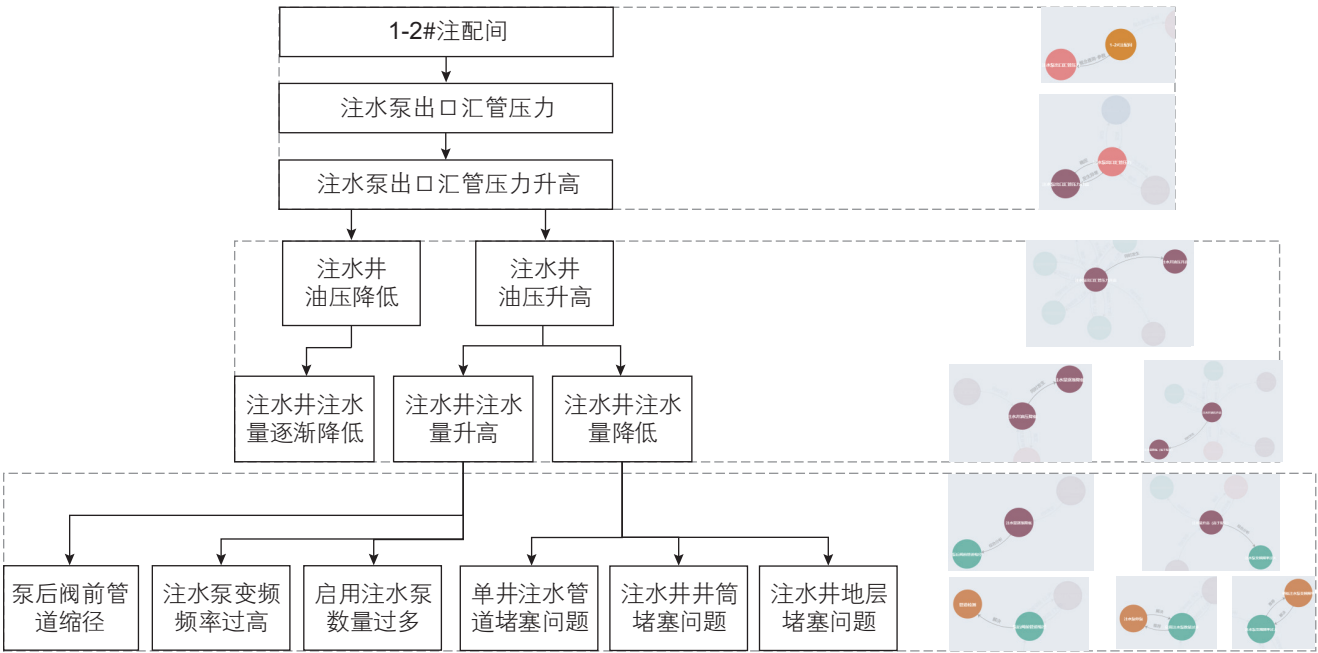


图 11 异常原因推理流程

Fig. 11 Process of abnormal cause reasoning

游为注水泵，下游为注水井。实际导致异常发生的可能是其中一个因素，也可能是多因素共同作用。当注水泵出口汇管压力升高时，根据关系链条确定的可能因素以及其它参数如注水井油压以及注水量的变化情况进行判断。如果注水井油压较高，并且流量计监测发现注水量较低不满足配注量，说明单井注水管道、井筒或地层存在堵塞问题，导致泵后管道憋压、能耗增加。在保证正常生产的情况下，推荐采用酸化开裂以及管线冲洗缓解堵塞。如果调整效果不明显，可以采取部分注水井停注、降注，或更换注水管线等进一步措施。如果注水井油压较高，但注水量大于正常配注量，则说明注水泵变频过大或启用注水泵数量过多，需要适当停泵、调低变频。如果注水井油压较低，并且注水量逐渐降低，说明可能是注水泵后汇管与该注水井调节阀之间的管道发生堵塞，需要对该段管道进行针对性的检测。

4 结论

针对油田集输与注水系统能源管控智能决策支持

面临的海量多源异构数据难题，本研究提出一种基于知识图谱的油田集输与注水系统能耗异常智能辅助决策方法，从数据和知识层面整体性地辅助指导集输与注水系统的能耗控制。

(1)采用BiGRU-CRF命名实体识别模型和BiGRU-Attention关系模型，对油田集输与注水系统能源管控领域文本信息进行知识抽取；根据知识抽取结果，利用Neo4j图数据库建立了概念图谱和实体图谱。

(2)基于所构建的知识图谱，采用Web技术研制以业务需求为导向的能源管控知识图谱可视化平台，实现能耗异常问题归因分析和决策推荐等功能。经验证，所搭建的平台在任务场景中能够为操作人员进行快速、全面、准确的能耗控制指导，对能耗异常事件进行有效处理。

(3)采用所提出的基于知识图谱的油田集输与注水系统能耗异常智能辅助决策方法，能够进行及时性、整体性、智能性的辅助决策推荐，提升油田集输与注水系统能源管控水平，为集输与注水系统能源管控决策支撑提供一定参考。

参考文献

[1] WU T X, QI G L, LI C, et al. A survey of techniques for constructing chinese knowledge graphs and their applications[J]. Sustainability, 2018, 10(9): 3245.

- [2] 刘俊楠, 刘海砚, 陈晓慧, 等. 面向多源地理空间数据的知识图谱构建[J]. 地球信息科学学报, 2020, 22(7): 476–1486 [LIU J N, LIU H Y, CHEN X H, et al. The construction of knowledge graph towards multi-source geospatial data[J]. Journal of Geo-Information Science, 2020, 22(7): 1476–1486.]
- [3] MARTINS V W B, RAMPASSO I S, ANHOLON R, et al. Knowledge management in the context of sustainability: literature review and opportunities for future research[J]. Journal of Cleaner Production, 2019, 229: 489–500.
- [4] 田玲, 张谨川, 张晋豪, 等. 知识图谱综述——表示、构建、推理与知识超图理论[J]. 计算机应用, 2021, 41(08): 2161–2186. [TIAN L, ZHANG J C, ZHANG J H, et al. Knowledge graph survey: Representation, construction, reasoning and knowledge hyper-graph theory[J]. Journal of Computer Applications, 2021, 41(8): 2161–2186.]
- [5] LI L F, WANG P, YAN J, et al. Real-world data medical knowledge graph: Construction and applications. Artificial Intelligence in Medicine, 2020, 103: 101817.
- [6] MACLEAN F. Knowledge graphs-and their applications in drug discovery[J]. Expert Opinion on Drug Discovery, 2021, 16 (9): –1057–1069.
- [7] 朱广宇, 张萌, 裔扬. 基于知识图谱的城市轨道交通突发事件演化结果预测[J]. 电子与信息学报, 2022: 1–9. [ZHU G Y, ZHANG M, YI Y. Prediction of Evolution Results of Urban Rail Transit Emergencies Based on Knowledge Graph[J]. Journal of Electronics & Information Technology, 2022: 1–9.]
- [8] 邹易呈. 基于多源数据融合的交通知识图谱构建及应用[D]. 大连: 大连理工大学, 2020. [ZOU Y C. Construction and application of traffic knowledge graph based on multi-source data Fusion[D]. Dalian: Dalian University of Technology, 2020.]
- [9] 张虹, 景欣, 阮梦宇, 等. 基于知识图谱的交直流大电网断面越限处置策略快速生成方法[J]. 现代电力, 2021, 38(4): 455–464. [ZHANG H, JING X, RUAN M Y, et al. A Method of Fast Eliminating Transmission Section Overload Based on Knowledge Graph in Large-scale AC-DC Power Grids[J]. Modern Electric Power, 2021, 38(4): 455–464.]
- [10] 郭榕, 杨群, 刘绍翰, 等. 电网故障处置知识图谱构建研究与应用[J]. 电网技术, 2021, 45(6): 2092–2100. [GUO R, YANG Q, LIU S H, et al. Construction and application of power grid fault handling knowledge graph[J]. Power System Technology, 2021, 45(6): 2092–2100.]
- [11] 张富利, 张恩莉, 向永慧, 等. 知识图谱技术在石油天然气勘探开发知识管理中的应用探讨[J]. 2020, 18(01): 128–131. [ZHANG F L, ZHANG E L, XIANG Y H, et al. Application of knowledge atlas technology in knowledge management of oil and gas exploration and development[J]. China CIO News, 2020(1): 128–131.]
- [12] 王智广, 文红英, 鲁强, 等. 地质领域开放式实体关系联合抽取[J]. 计算机工程与设计, 2021, 42(04): 996–1005. [WANG Z G, WEN H Y, LU Q, et al. Joint extraction of open entity relation in geological field[J]. Computer Engineering and Design, 2021, 42(04): 996–1005.]
- [13] 金泉. 基于深度学习的测井领域知识图谱构建与应用[D]. 大庆: 东北石油大学, 2021. [JIN Q. Construction and application of knowledge graph in logging domain based on deep learning[D]. Daqing: Northeast Petroleum University, 2021.]
- [14] 陈传刚, 胡瑾秋, 韩子从, 等. 恶劣环境条件下海外天然气管道站场事故演化知识图谱建模及预警方法[J]. 清华大学学报(自然科学版), 2022, 62(06): 1081–1087. [CHEN C G, HU J Q, HAN Z C, et al. Knowledge graph based early warning method for accident evolution in natural gas pipeline station abroad for harsh environmental conditions[J]. Journal of Tsinghua University(Science and Technology), 2022, 62(06): 1081–1087.]
- [15] ZHOU X G, GONG R B, SHI F G, et al. PetroKG: Construction and application of knowledge graph in upstream area of PetroChina[J]. Journal of Computer Science and Technology, 2020, 35(2): 368–378.
- [16] 张吉祥, 张祥森, 武长旭, 等. 知识图谱构建技术综述[J]. 计算机工程, 2022, 48(03): 23–37. [ZHANG J X, ZHANG X S, WU C X, et al. Survey of knowledge graph construction techniques[J]. Computer Engineering, 2022, 48(03): 23–37.]
- [17] 鄂海红, 张文静, 肖思琪, 等. 深度学习实体关系抽取研究综述[J]. 软件学报, 2019, 30(06): 1793–1818. [E H H, ZHANG W J, XIAO S Q, et al. Survey of entity relationship extraction based on deep learning[J]. Journal of Software, 2019, 30(06): 1793–1818.]
- [18] LI J, SHANG S, CHEN L S. Domain generalization for named entity boundary detection via metalearning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, PP(99): 1–12.
- [19] LAFFERTY J D, MCCALLUM A, PEREIRA F C N. Conditional random fields: Probabilistic models for segmenting and labeling sequence data [J]. Proceedings of ICML, 2002.
- [20] 赵鹏武, 李志义, 林小琦. 基于注意力机制和卷积神经网络的中文人物关系抽取与识别[J]. 数据分析与知识发现, 2022, 6(8): 41–51. [ZHAO P W, LI Z Y, LIN X Q. Identifying relationship of Chinese characters with attention mechanism and convolutional neural network[J]. Data Analysis and Knowledge Discovery, 2022, 6(8): 41–51.]
- [21] MNIH V, HEES N, GRAVES A, et al. Recurrent models of visual attention[J]. Advances in Neural Information Processing Systems, 2014, 3.
- [22] 王鑫, 邹磊, 王朝坤, 等. 知识图谱数据管理研究综述[J]. 软件学报, 2019, 30(7): 2139–2174. [WANG X, ZOU L, WANG C K, et al. Research on knowledge graph data management: A survey[J]. Journal of Software. 2019, 30(7): 2139–2174]

- [23] 蒋玮, 王铭华, 陈锦铭, 等. 基于Neo4j图数据库的配电网供电可靠性计算[J]. 电力系统自动化, 2022, 46(15): 104–111. [JIANG W, WANG M H, CHEN J M, et al. Calculation of power supply reliability for distribution network based on Neo4j graph database[J]. Automation of Electric Power Systems, 2022, 46(15): 104–111.]
- [24] 王昊奋, 丁军, 胡芳槐, 等. 大规模企业级知识图谱实践综述[J]. 计算机工程, 2020, 46(07): 1–13. [WANG H F, DING J, HU F H, et al. Survey on large scale enterprise-level knowledge graph practices[J]. Computer Engineering, 2020, 46(07): 1–13.]

(编辑 马桂霞)

第一作者: 王文君(1998 年—), 在读博士研究生, 主要从事油气储运系统节能与安全研究, 470519480@qq.com。

通信作者: 侯磊(1966 年—), 博士, 教授, 博士生导师, 主要从事复杂流体相态与流动特性, 油气集输与管道输送工艺, 油气储运系统节能与安全研究, houleicup@126.com。