

文章编号: 1000-8055(2023)10-2516-11

doi: 10.13224/j.cnki.jasp.20220007

基于 KG-KGCN 的某型动力装置风险分析方法

陈国兵, 曾国庆, 王 越, 王学峰, 谢旭阳, 杨自春
(海军工程大学 动力工程学院, 武汉 430033)

摘 要: 针对动力装置危险因素众多、风险数据繁杂且利用率不高等问题, 以故障模式、影响及危害性分析等数据为基础, 提出基于知识图谱和图神经网络的系统风险分析方法及算法, 利用事件的知识完成链接预测, 推理出引发事件的原因及影响等, 以实现其精准的统计分析。本文以某型发动机为对象, 对其进行风险分析和推理。结果表明: 该方法有效解决了传统方法存在的表单繁琐、分析单一等问题, 结果更加准确和全面, 可使分析效率提升 60% 以上; 并可及时开展风险推理和预测, 实现从事后分析向主动预防的转变, 为动力装置的智能运维提供了有力支持。

关 键 词: 风险分析; 知识图谱; 知识图卷积网络; 动力装置; 故障模式、影响及危害性分析 (FMECA)
中图分类号: V240.2; TP391 **文献标志码:** A

Risk analysis method of a certain type of power plant based on KG-KGCN

CHEN Guobing, ZENG Guoqing, WANG Yue, WANG Xuefeng,
XIE Xuyang, YANG Zichun

(College of Power Engineering, Naval University of Engineering, Wuhan 430033, China)

Abstract: Targeting the problems of numerous risk factors, complex risk data and low utilization rate of power plant, and based on the data of failure mode, impact and criticality analysis, a method and algorithm of system risk analysis was proposed using knowledge graph and graph neural network, and the knowledge of events were utilized to complete link prediction and deduce the cause and impact of events, so as to achieve accurate statistical analysis. A certain type of engine was taken as the object to carry out risk analysis and reasoning. Results showed that this method effectively solved the problems of the traditional method, such as cumbersome forms and single analysis, the results were more accurate and comprehensive, and the analysis efficiency can be improved by more than 60%; and it can carry out risk reasoning and prediction in time, and realize the transformation from post analysis to active prevention, providing a strong support for the intelligent operation and maintenance of power plants.

Keywords: risk analysis; knowledge graph; knowledge graph convolutional network;
power plant; failure mode, effects and criticality analysis (FMECA)

随着现代工业系统智能化、集成化的不断发展, 系统越来越复杂, 受复杂性、随机性、动态性等多方面因素的影响, 事故发生的结果往往难以让人接受, 付出的代价往往很大。例如, 1985 年

8 月 12 日, 日本航空 123 号班机由于液压系统故障失效发生坠毁, 520 人遇难; 1986 年 1 月 28 日, 美国“挑战者”号 (HMS Challenger) 航天飞机升空时发生爆炸, 究其原因是右侧助推器的密封垫圈

收稿日期: 2022-01-05

基金项目: 国家自然科学基金 (51609251)

作者简介: 陈国兵 (1982—), 男, 教授、博士生导师, 博士, 主要从事舰船动力及热力设备的科学管理研究。E-mail: chenguob@163.com

通信作者: 杨自春 (1967—), 男, 教授、博士生导师, 博士, 主要从事舰船动力及热力设备的科学管理研究。E-mail: yangzichun11@sina.com

引用格式: 陈国兵, 曾国庆, 王越, 等. 基于 KG-KGCN 的某型动力装置风险分析方法 [J]. 航空动力学报, 2023, 38(10): 2516-2526. CHEN Guobing, ZENG Guoqing, WANG Yue, et al. Risk analysis method of a certain type of power plant based on KG-KGCN [J]. Journal of Aerospace Power, 2023, 38(10): 2516-2526.

老化。在美国空军飞机的使用经历中, 战斗机与攻击机有 20% 的灾难性事故与发动机有关, 如近年来美军 B-1B 轰炸机发动机故障、B-52H 轰炸机发动机叶片问题等。对于船舶而言, 由于海洋环境的特殊性, 随着海洋活动的日趋频繁, 动力瘫痪等事故发生的可能性也随之升高。根据国际海事组织(International Marine Organization, IMO)统计, 1991~2017 年的 20 多年内共计发生各类船舶事故 8 781 起, 由动力系统引发的事故造成了较大的经济损失。2012 年, 俄罗斯为印度改建的航母试航过程中, 动力系统 8 台锅炉 7 台坍塌, 损失数亿美元并延期一年交船, 引发全世界关注。2016 年以来, 英国造价高达 10 亿英镑的 45 型驱逐舰“无畏号”, 因动力问题一直处于停泊状态。2020 年以来, 美军造价 5 亿美元的“自由”级濒海战斗舰因动力问题等饱受诟病, 在海上发生一系列动力故障。由此可见, 系统及设备的风险分析和精准管控, 对提高其整体安全性水平至关重要。

目前, 在动力装置风险分析中, Bahrebar 等^[1]指出船舶动力系统风险分析中危险源识别分析(HAZID)^[2]、危险与可操作性分析(HAZOP)^[3]、结构化假设分析(SWIFT)以及失效模式与影响分析(FMEA)方法高度依赖于初始设计阶段和制造阶段的定性评估; 并在考虑一般的 2 型模糊逻辑集下开发了一种鲁棒 FMEA 来识别船舶动力系统的潜在危险条件。徐立等^[4]针对大型民用核动力船舶的安全性问题, 应用层次分析方法(AHP)对民用船舶核动力装置的二回路系统进行故障分析。总体来说, 这些传统的动力装置风险分析方法和思路, 主要侧重于对已发生事故的原因分析和规律的研究, 对风险分析、评估^[5]与控制等方面更多依靠经验积累等, 缺乏更加科学合理的风险分析及管控的技术手段。如最常用的 FMEA 及其改进的 FMECA 等, 一方面复杂系统 FMECA 表单往往过于复杂, 难以提取有用信息, 对其数据处理和分析缺乏有效手段; 另一方面其分析结果主要为关键风险清单和风险矩阵图等, 大量数据和信息难以得到有效利用; 同时 FMECA 方法难以解决系统存在的多状态、相关性和及时更新等问题, 难以满足当前复杂系统风险分析及管控的需求。因此, 对最为基础的 FMECA 等数据进行挖掘, 采用基于数据和知识的双驱动智能混合模型来应对复杂系统的风险分析具有重要的研究

意义。

近年来, 随着人工智能技术的快速发展, 知识图谱(knowledge graph, KG)技术已经应用到各个研究领域。知识图谱作为“图”这种直观的数据结构来描述知识, 构建事物之间的概念、实体、事件和相互关系等, 能够表达关系复杂的知识, 很容易挖掘数据之间的潜在联系^[6]。一个全面的知识图谱具有以下特点: ①系统化、规范化和智能化地将现有的数据和知识进行互连, 形成一个丰富的知识关系网, 弱化甚至消除“知识孤岛”现象^[7]; ②图数据的更改和增删更加便捷, 易于丰富现有的知识图谱^[8], 更加易于模型的更新; ③具有较强的可推理性, 可解释性及可预测性^[9]。在航空领域知识图谱中, 邱凌等^[10]对比航空制造知识图谱与通用知识图谱构建技术上的异同, 并结合综述情况提出了知识表示与抽取、知识融合和知识推理等切合实际的航空制造知识图谱应用方向及其解决方案。郭静^[11]提出一种改进的 FP(frequent pattern) growth 算法, 通过引入跨层次频繁项和修补项, 每层设置不同的支持度自底向上对编码后的多层次属性进行挖掘, 最后得到事件原因、事件结果、运行阶段、事件类型等多种事件属性之间的多层关联规则。杜党党等^[12]针对航空发动机状态参数和故障机理的特点, 提出一种基于故障彩色图谱的发动机故障模式识别方法, 通过计算待检测故障彩色图谱与标准故障彩色图谱库中各元素之间的相似度, 输出故障类型。

图神经网络(graph neural networks, GNN)模型兼顾了大规模数据的输入、异质信息网络和更进一步的拓扑等问题, 在深层挖掘、捕获数据特征和处理海量数据等方面具有独特能力。但是, 初期的研究模型以循环和迭代的计算方式消耗非常大。2013 年, Bruna 等^[13]将卷积操作与图神经网络结合, 提出图卷积网络概念(graph convolution network, GCN)与模型。自此之后, 很多研究学者提出改进、拓展这种基于频域图卷积的神经网络模型, Wang 等^[14]融合知识图谱的特点以及图神经网络(graph convolutional networks)方法得到了知识图卷积网络(knowledge graph convolutional networks, KGCN)。例如: Zhang 等^[15]通过建立 Fast GCRNN(fast graph convolution recurrent neural network)模型应用于道路网络交通图, 该模型能够提取大规模路网交通流的时间和空间属性来预测交通状况。Kapoor^[16]研究了一种使用图

时空网络和移动数据预测 COVID-19(corona virus disease 2019)病例的预测方法。该方法从单一的大尺度时空图中学习,其中节点代表区域层次的人的流动性,空间边缘代表基于区域间连接性的人的流动性,时间边缘代表节点的时间特征。刘欢等^[17]提出一种基于知识图谱驱动的端到图神经网络推荐模型 KGLN(knowledge graph driven learning network, KGLN)。使用单层神经网络框架对图中单个节点进行特征融合,并加入影响因素来改变不同邻居实体的聚合权重。史运涛等^[18]提出了一种基于图神经网络的社区户内燃气系统动态风险评估方法,将知识图谱用于社区户内燃气系统的场景构建,实现了社区户内燃气系统的安全动态风险评价。因此将知识图谱与神经网络结合能够很好的解决复杂系统风险分析的短板和难题。

虽然结合知识图谱与神经网络在解决复杂系统风险分析上有着明显的优势,但在动力装置方面的工程实际应用并不多,主要原因在于动力装置属于复杂机械系统,寿命周期的故障数据往往较少。FMECA(failure mode, effects and criticality analysis)作为动力装置设计、制造、使用维护中最为基础的风险数据,目前缺乏在其基础上结合知识图谱与神经网络开展相关研究,而针对 FMECA 这种带有因果关系的结构化数据,可以很大程度上增强结合知识图谱与图神经网络方法的鲁棒性和可解释性。因此,本文以 FMECA 等风险数据为基础,提出基于知识图谱和图神经网络的系统风险分析方法及算法,建立符合智能化、信息化、数字化的风险评估模型,并以某型动力装置发动机为例进行分析。一方面充分利用了现有风险分析数据,使得风险评估更加科学有效;另一方面开展了系统风险的智能分析及评估,可进一步强化系统的风险管理,对实现系统风险的智能预防预报、控制及应急处置等具有重要意义。

1 基于 KG-KGCN 的动力装置风险分析方法

1.1 知识图谱构建方法

1) 构建流程

知识图谱构建方法分为两种方式,其中自上而下(top-down)是指先设置好语义信息和逻辑规则,然后将获取的实体元素以及之间的关系存入知识图谱中^[19]。自下而上(bottom-up)刚好相反,

是先提取的实体和关系,再归纳本体和语义。专业性较强的领域知识图谱的构建一般选择自上而下效果更佳。所以,本文采用自上而下的方式进行构建。

构建系统风险知识图谱整体方法如图 1 所示。具体流程如下:

(1) 根据 FMECA 风险清单的基础数据和风险分析的专业知识进行知识建模,确定知识图谱模式层的结构。本文选取 Neo4j 图数据库进行知识存储,不需要显示本体,因此不需要像采用 RDF(resource description framework)存储知识图谱时进行本体建模。

(2) 由于本文涉及到的是 FMECA 风险清单是结构化数据,采用手动处理的方式更加简单方便且严谨。因此通过处理 FMECA 数据得到相关的实体和关系,并将实体和对应的关系采用三元组(h, r, t)的形式存储,其中 head(头实体), relation(实体的关系), tail(尾实体)。这些三元组就是知识图谱的一种表现形式。为了更好地可视化表现和反映实体之间的关系,将得到的实体以图的形式展现,也就是采用 Neo4j 来对三元组进行存储和可视化。

(3) 在完成知识图谱可视化后,根据实体之间的关系可以进行数据之间的查询和统计,也可以根据已有的知识图谱扩展更多数据之间的关系,来进一步丰富和完善知识图谱的结构。

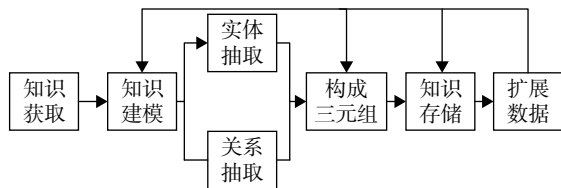


图 1 系统风险知识图谱构建

Fig. 1 Process of building knowledge graph of system risk

2) 建立实体和关系

由于 FMECA 风险清单是结构化数据,因此在进行实体关系抽取时通过人工操作即可获取关联性强的实体关系,且准确度更高。本文基于 FMECA 风险清单的知识图谱中包含了 6 类实体、5 类关系以及 6 类属性。

数据存储在 Neo4j 图数据库中,其属性不会主动显示,可以通过查询查看。通过抽取和整理之后,实体及其之间的关系如表 1 所示。根据上述操作完成 FMECA 风险清单的实体和关系抽取工作,其结构图如图 2 所示。

表 1 不同类型实体及其关系

Table 1 Different types of entities and their relationships

模型数据	数据结构	作用	对象范围
实体	标签	可被标记多个类别标签	系统、故障模式、故障原因、故障影响
	属性	可存储多个相关属性	编号、工况、严酷度等
关系	类型	所属关系类型	导致关系、包含关系
	属性	存储关系属性	关系名称、关系 ID

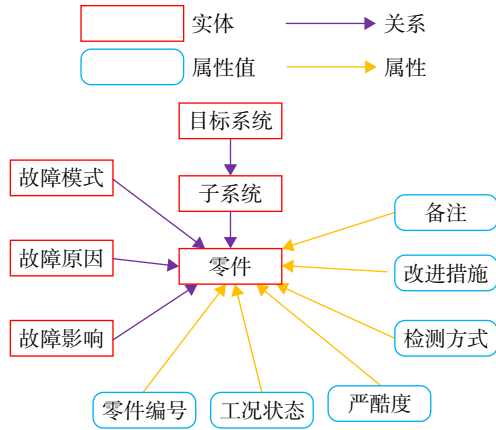


图 2 实体及其关系结构图

Fig. 2 Entities and relationships structure diagram

1.2 知识图神经网络算法

1) 知识图卷积网络

知识图卷积网络的推理是一种端到端的框架,用于捕捉知识图谱实体之间的高阶结构相似性,并提供推理合理的解释性。模型框架图如图 3 所示。

为了能够捕获目标的高阶结构信息和语义信息,KGCCN 受到 GCN 的启发,将目标中的实体邻域中抽取样本作为它们的感受野(receptive field)。当计算其中一个实体表示的时候,就可以综合邻域和偏置信息。当然,感受野可以自己设定大小,用来扩展到多层捕获高阶信息,使得目标实体获得更丰富的信息,并完成潜在的推理预测。

2) 模型分析

数学定义如下:

(1) 实体 u 的集合 $U = \{u_1, u_2, \dots, u_M\}$, M 表示 u 的个数。

(2) 实体 v 的集合 $V = \{v_1, v_2, \dots, v_N\}$, N 表示 v 的个数。

(3) 交互矩阵 $Y \in \mathbf{R}^{M \times N}$, 其中 $y_{uv} = 1$ 表示 u 中的 u 对 v 中的 v 产生了交互关联, 否则 $y_{uv} = 0$ 。

(4) 知识图谱 G , 该知识图谱由三元组 (h, r, t) 组成, 其中 $h \in \mathcal{E}$ 表示三元组的头部, $t \in \mathcal{E}$ 表示三元

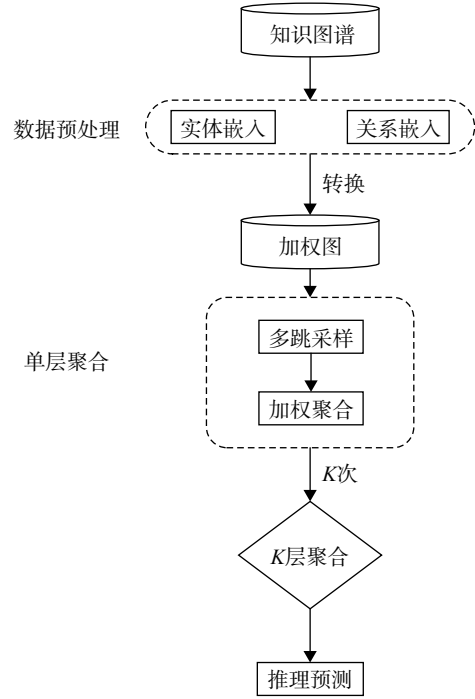


图 3 KGCCN 的知识图谱推理过程

Fig. 3 Reasoning process of knowledge graph of KGCCN

组的尾部, $r \in \mathbf{R}$ 表示三元组中头部实体 h 和尾部事件 t 的关系, $\mathcal{E}$ 和 $\mathbf{R}$ 分别表示实体集合和关系集合。

对于实体 $u-v$ 的 Y 和 G , 是为了推理 u 中的 u 的是否对 v 中的 v 产生关联。即, 模型的目标是学习预测函数^[14]

$$\hat{Y}_{uv} = F(u, v | \Theta, Y, G) \quad (1)$$

其中 $\hat{Y}_{uv}$ 表示 u 中的 u 对 v 中的 v 的产生关联的可能性, 而 Θ 表示函数 F 的模型参数。

3) 单层 KGCCN

对于单个 KGCCN 层, 考虑实体 u 和实体 v 的一个交互对 $(u-v)$, 令直接与 v 相连的实体集合为 $N(v)$, 用 $r_{ei, ej}$ 表示实体 e_i 和 e_j 之间的关系。在知识图谱中的边是没有显示权重的, 只代表了一个关系类型。采用函数 $g: \mathbf{R}^d \times \mathbf{R}^d \rightarrow \mathbf{R}$ (例如, 内积操作) 的方式, 用来计算实体 u 与关系 $r_{ei, ej}$ 的分数, 对每个关系进行打分, 从而把知识图谱转换成权重图。该函数的结果表示该关系对实体 u 的重要程度^[14]

$$\pi_r^u = g(u, r) \quad (2)$$

式中 u 和 r 是由嵌入向量表示, $u \in \mathbf{R}^d$ 和 $r \in \mathbf{R}^d$ 。 π_r^u 作为结果量化了关系 r 对实体 u 的重要程度。

由于实体 v 的邻接节点和其本身有若干关系, 为了拓扑该邻近结构, 将上述得到的分数 π_r^u 与实体的嵌入向量相乘, 完成了实体 v 的邻域的线性

聚合^[14]

$$\mathbf{v}_{N(v)}^u = \sum_{e \in N(v)} \tilde{\pi}_{rv,e}^u \mathbf{e} \quad (3)$$

式中 $\tilde{\pi}_{rv,e}^u$ 是对实体 u 对关系 r 的分数的归一化,可以将其作为“Attention”也就是注意力机制,用来区分邻居的重要性, $\tilde{\pi}_{rv,e}^u$ 的计算公式^[14]为

$$\tilde{\pi}_{rv,e}^u = \frac{\exp(\pi_{rv,e}^u)}{\sum_{e \in N(v)} \exp(\pi_{rv,e}^u)} \quad (4)$$

式中 $\mathbf{e}$ 是初始实体的表示,即当 $K=0$ 时中心实体的向量表示。

将实体 v 的邻居表示计算为 $\mathbf{V}_{S(v)}^u$, $S(v) \triangleq \{e|e \sim N(e)\}$ 并且 $|S(v)| = K$ 是一个可配置常量。在 KGCN 中, $S(v)$ 也被称为实体 v 的感受野,因为实体 v 的最终表示非常依赖这些区域。如图 4 所示,是给定一个实体的两层感受野^[14],其中 $K=2$,实体为蓝色,感受野为绿色。

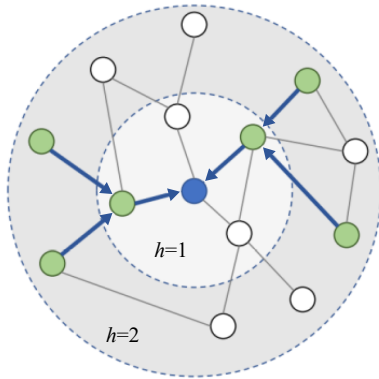


图 4 知识图谱实体两层感受野^[14]

Fig. 4 Two-layer of receptive fields of the entity in the knowledge graph^[14]

经过上述步骤对每个与实体 u 相连的实体 v 进行聚合操作,这样就完成了单层 KGCN。最后一步是通过一个单层感知机,即将实体 v 及其邻域表示 $\mathbf{v}_{S(v)}^u$ 进行聚合,得到聚合后的嵌入向量。聚合是 KGCN 中的关键一步,通过聚合将实体与其邻居联系在一起。聚合方式为加法聚合:将两个表示相加,进行非线性变换

$$\mathbf{Z}_{\text{sum}} = \sigma[W \cdot (\mathbf{v} + \mathbf{v}_{S(v)}^u) + b] \quad (5)$$

式中 $W \in \mathbf{R}^{d \times d}$ 是权重, $b \in \mathbf{R}^d$ 是偏差项, σ 是非线性激活函数,例如 ReLU(rectified linear units)激活函数。

4) KGCN 算法描述

通过单个 KGCN 层的操作,一个实体的最终表示由它自己和它的直接邻居决定,称之为 1-or-

der 实体表示。接着将 KGCN 从一层延伸到多层,那么将以更广和更深的方式进行合理和有效的推理预测。将每个实体的初始表示(记为 0-order 表示)传播到它的邻居,可以得到 1-order 实体表示;重复上述操作,进行进一步传播和聚合 1-order 表示,即可得到 2-order 表示。因此,一个实体的 h -order 表示是由它自身的初始表示(0-order 表示)和 h 跳邻居的混合表示。

1.3 基于 KG-KGCN 的风险分析方法

1) 基于知识图谱可视化的风险统计与查询

在知识图谱上进行系统风险的统计与查询,可以通过输入 Cypher 语句来完成。Cypher 语言是 Neo4j 主要支持的查询语言,操作简单,不仅有丰富的表现力还具有较高的准确率^[20]。基于 FMECA 风险清单构建的系统风险知识图谱不仅储存了完整的故障数据,借助 Cypher 还可以进行系统整体情况,子系统之间,零件之间和具体故障情况的查询,具体流程如图 5 所示。

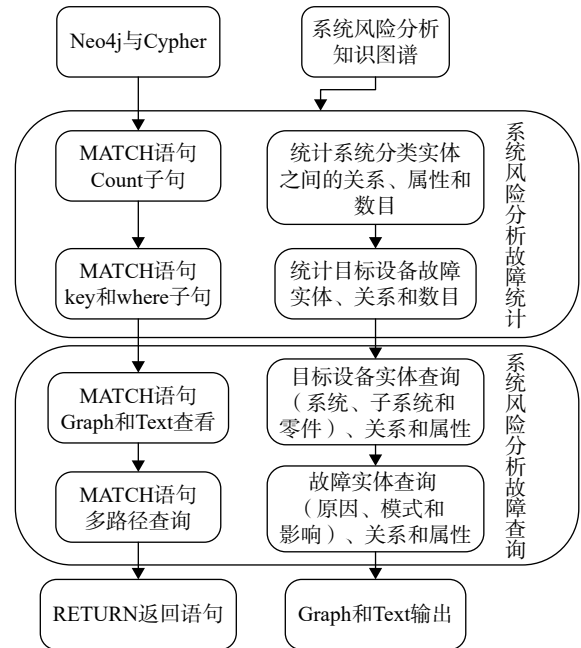


图 5 系统风险分析知识图谱统计与查询流程

Fig. 5 Process of system risk analysis knowledge graph statistical and query

当系统复杂,子系统以及零件繁多且随着数据的增多,FMECA 的风险清单数据将是巨量的,此时对系统风险知识图谱进行查询和统计是非常有必要的。对知识图谱可视化后进行统计与查询分析,能够快速看到相关信息的直接展示。可视化的查询与统计分析通过将各类信息转化为视觉图像的方式呈现,不仅提高了相关工作人员对抽

象信息的理解和判断能力, 而且使得在数据采集和挖掘的评估、改进等方面更加方便^[21]。

2) 基于 KGCN 的系统风险的推理

系统风险事件推理的详细过程如图 6 所示, 主要包括数据预处理、模型的建立和学习、模型的测评三大模块。

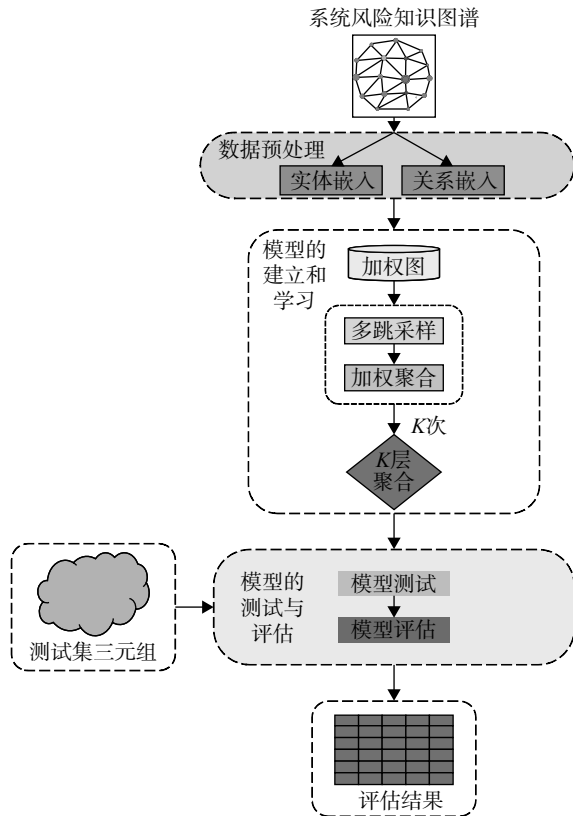


图 6 基于 KGCN 的推理过程

Fig. 6 Reasoning process based on KGCN

(1) 数据预处理: 考虑到 FMECA 风险清单的特点, 故障现象在发生时是直接可观的, 故障原因是引起风险发生的根本, 选取故障模式作为故障现象和故障原因的关系。因此本文选取知识图谱三元组(故障现象, 故障模式, 故障原因), 即故障现象作为头部实体 h , 故障模式作为实体关系 r , 故障原因作为尾部实体 t 。

首先将实体和实体之间的关系嵌入到低维向量中。独热编码(one-hot encoding)具有将离散特征向量映射到欧式空间的特点, 使得特征之间的距离计算更加合理, 因此本文选择独热编码的编码形式处理实体和实体之间关系向量化。

(2) 模型的建立和学习: 风险事件中故障的发生往往是由相关原因引起的, 及时发现故障原因并采取干预措施是防止风险发生的根本方法, 因此对故障发生的原因必须重视。通过使用

KGCN 模型对风险事件的学习, 发现相关事件的风险因素。由于故障原因涉及到各方面, 因此对与事件有关系的所有故障原因进行统一建模分析。

(3) 模型的评测: 使用测试集中的数据进行故障事件的原因及结果影响进行推理, 然后对推理的结果进行分析。

基于系统风险知识图谱进行事件原因及影响的推理过程如下:

(1) 首先使用命名实体识别、实体融合和关系提取等相关技术, 从系统的 FMECA 风险清单中提取风险事件的实体和关系, 并生成相应的三元组信息。

(2) 将三元组构建的知识图谱中的实体及其关系按照独热编码进行向量初始化表示, 并通过模型训练得到实体向量表示, 然后以函数内积的形式将知识图谱转化成加权图, 得到节点的邻接列表和关系的邻接列表向量表示。

(3) 将实体和实体关系向量表示作为模型训练的初始嵌入向量, 以一个实体为起点传播两次。每一个实体的特征矢量为与该实体直接相连的外层实体特征矢量的和, 并引入注意力机制。通过得分函数与实体的嵌入向量相乘得到实体邻域的线性聚合, 其中得分函数的作用类似于注意力机制, 以区别每个邻居的重要性。通过上述操作, 来更新事件的特征向量, 从而融合了更丰富的语义信息。

(4) 聚合时, 为了保证计算模式固定以及计算效率更高, 对知识图谱中每个风险事件实体的邻域三元组, 以固定大小的感受野随机均匀地采样和聚合。

(5) 经过上述过程的多次迭代, 最终和候选实体的特征向量进行内积, 获得在相应实体和关系下的真实得分情况。

(6) 将获得的得分情况进行softmax函数操作, 最终获得每个三元组为真的概率数值。

(7) 将获得的推理概率最大的 t , 作为该事件最有可能发生的故障原因。

3) 基于 KG-KGCN 的风险分析流程

本文提出的 KG-KGCN 系统风险分析方法流程如图 7 所示。首先, 进行数据准备工作, 即对系统进行结构解析、FMECA 分析和数据预处理; 随后, 基于系统 FMECA 数据完成系统知识图谱的构建和风险查询分析; 然后, 基于 KGCN 方法完成系统风险推理和分析。

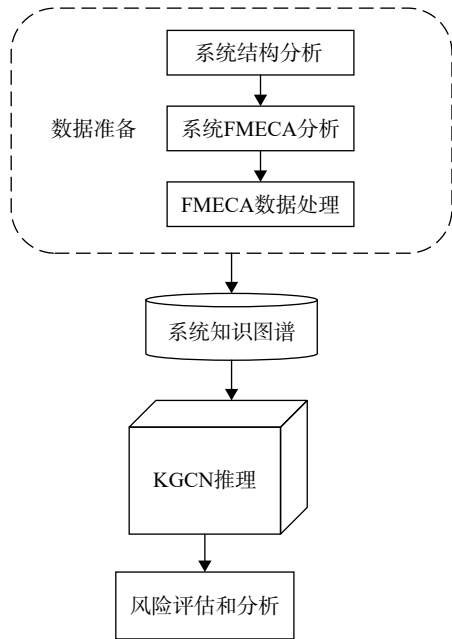


图 7 KG-KGCN 系统风险分析

Fig. 7 Model of KG-KGCN system risk analysis

2 基于 KG-KGCN 的某型动力装置风险分析

本文以某型动力装置发动机作为研究对象, 基于本文的 KG-KGCN 方法, 对其进行风险分析

和推理。

2.1 某型发动机 FMECA 分析结果

船用某型发动机 FMECA 分析结果, 包括压气机、燃烧室、涡轮和附件传动装置, 具体层次约定如图 8 所示。

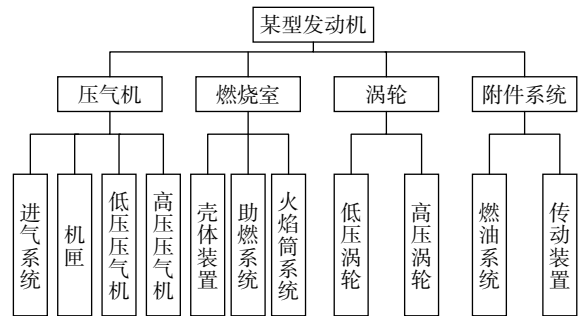


图 8 发动机 FMECA 层次约定

Fig. 8 Conventional level of engine FMECA

2.2 发动机风险知识图谱的构建及分析

1) 发动机风险知识图谱的构建

通过 Python 编程读取从数据库中处理过后的发动机 FMECA 风险清单, 然后提取及整理相关的数据, 添加系统结构顺序读取, 最终完成发动机风险知识图谱的搭建, 其知识图谱如图 9 所示。

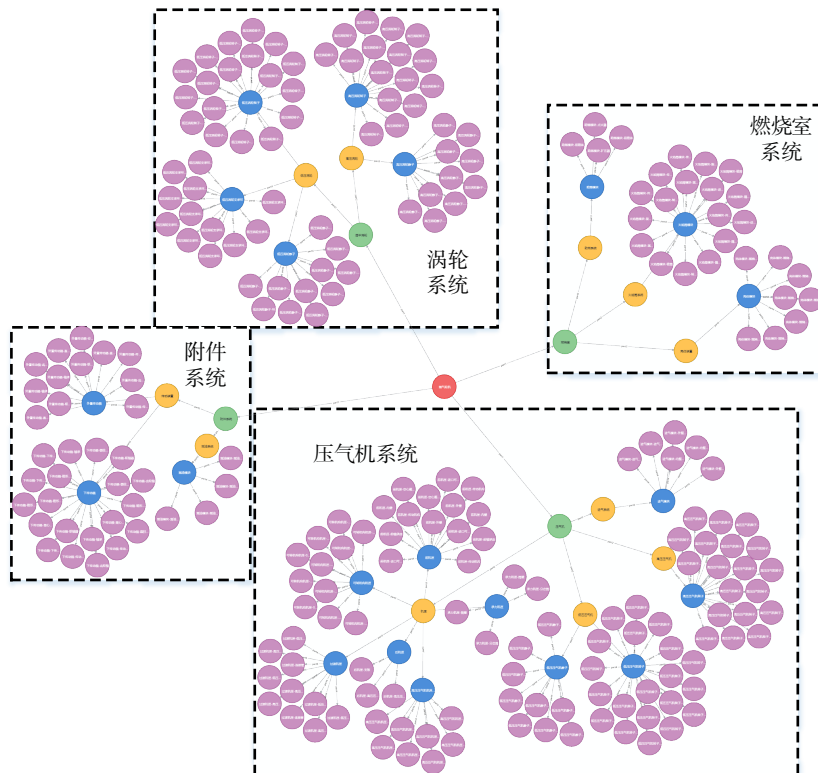


图 9 发动机 FMECA 整体知识图谱

Fig. 9 Overall knowledge graph of engine FMECA

2) 发动机风险知识图谱的分析

发动机知识图谱按照系统结构发散模式进行构建,中间节点为发动机主系统,通过包含关系与压气机、燃烧室、涡轮以及附件系统 4 个分系统建立关系;每个分系统通过包含关系与不同分系统下的子系统建立关系;每个子系统通过包含关系与相应模块系统进行关系的建立;每个模块系统通过包含关系与相应零件进行关系的建立。零件是终端包含了每个零件发生的故障详细信息即相对应的 FMECA 数据,包含了故障模式、故障现象、故障原因、工况状态、故障影响、严酷度、检测方式、改进措施和关键清单等内容,模型如图 10 所示。

2.3 发动机风险知识图谱统计与查询

发动机系统风险知识图谱统计,可以采用 count 等命令语句对相关数据进行统计分析。

对于相关设备故障数目的统计,例如:统计下传动箱共有多少零件发生故障,输入命令“match p=(sub_system {name:'下传动箱'})-[r]->() return count(p)”,可以看出低压压气机转子故障零

件有 20 个;输入相似命令“match p=(sub_system {name:'下传动箱'})-[r]->() return p”,可以得到下传动箱的故障零件的知识图谱,并且从图谱中可以看到每一个故障零件位置的详细信息,如图 11 所示。

对发动机相关设备进行分类统计,例如查询分系统的情况,可以输入“match (m:part_system) <-[r]->() return m”,可以得到分系统的图谱展示,如图 12 所示。

发动机系统风险分析知识图谱查询主要是对相关设备之间的查询和对具体故障的查询。其中相关设备之间的查询是对整体设备和设备之间包含关系进行查询。

例如,发动机设备查询时,输入命令“match p=()-[r:part3]->() return p”,会得到子系统与模块之间的知识图谱,得到包含的 11 个子系统和 21 个模块之间的关系图,如图 13 所示。

发动机具体故障的查询是对某模块具体设备的查询。在查询时,本文的故障信息是按图 10 所示的结构进行的信息存储,查询结果也遵循查询

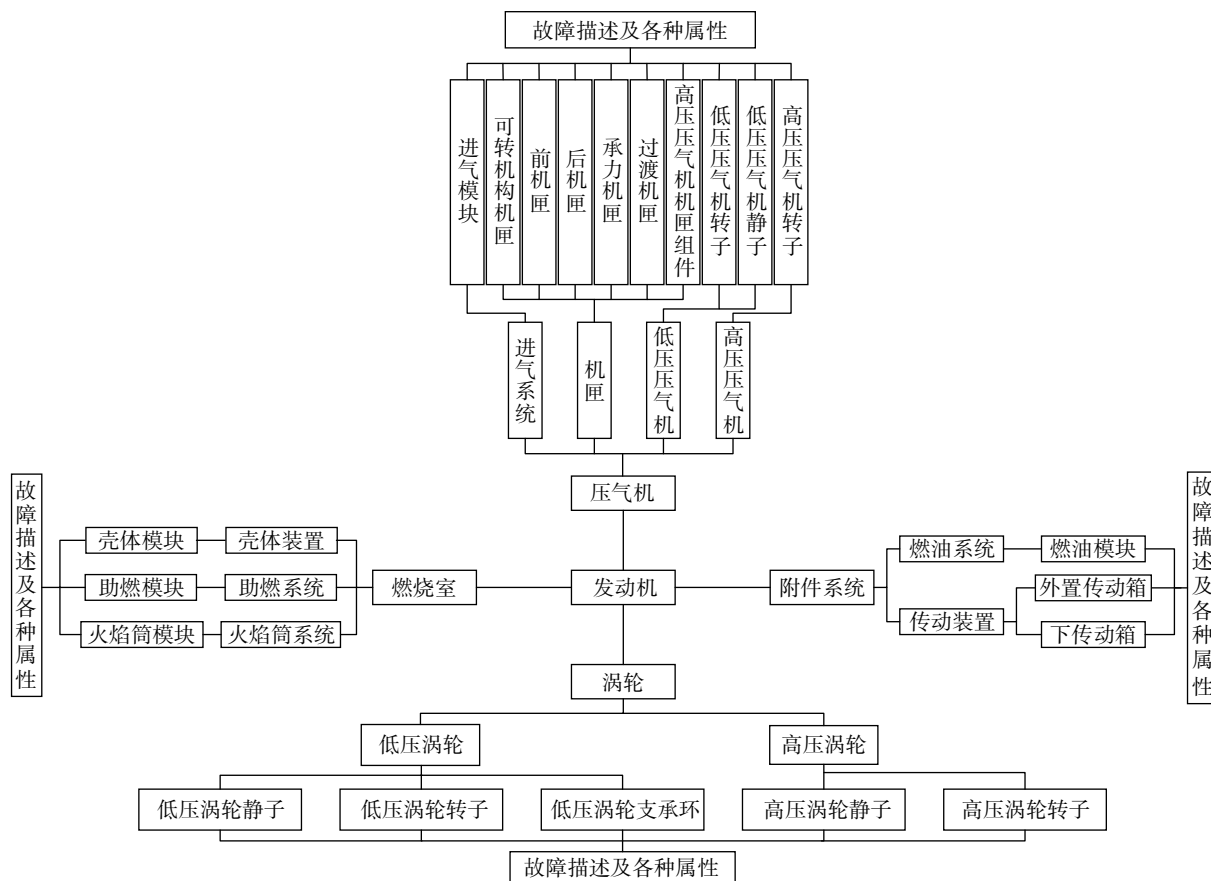


图 10 知识图谱结构模型

Fig. 10 Structure model of knowledge graph

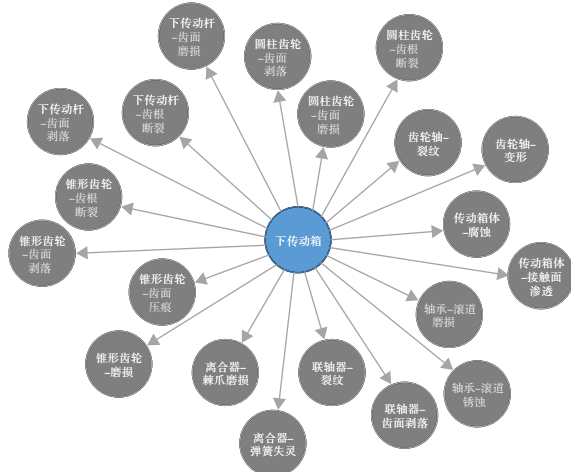


图 11 下传动箱故障零件统计图谱

Fig. 11 Faulty parts statistics graph in the lower gearbox



图 12 分系统设备统计图谱

Fig. 12 Equipment statistics graph of the part system

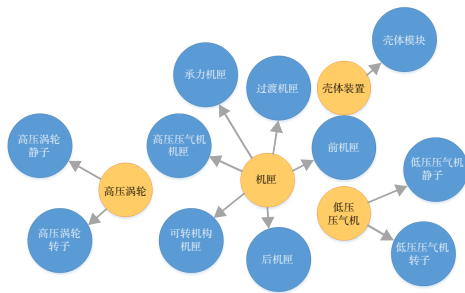


图 13 子系统与模块之间查询界面

Fig. 13 Interface of knowledge graph query between subsystems and modules

的结构。例如, 低压压气机转子出现故障, 输入命令“`match p=(sub_system {name:'低压压气机转子'})-[r]->( ) return p`”进行多路径查询, 就会看到关于低压压气机转子模块的零件的各种故障信息。其中低压压气机转子的第 0~2 级动叶出现裂纹、断裂的故障现象, 其故障模式是疲劳, 故障原因是低周疲劳, 最终影响会导致系统停车, 后

果严重, 需要请专业工程师对相关零件进行更换, 并提高零件的质量。如图 14 所示。

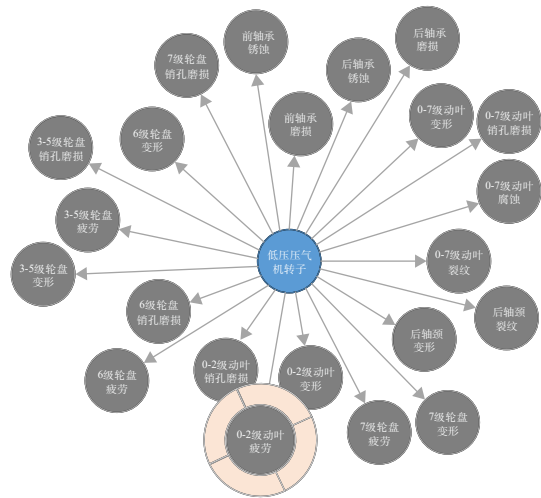


图 14 低压压气机转子具体故障查询结果界面

Fig. 14 Interface of query results of specific faults of the low-pressure compressor rotor

发动机风险知识图谱描述了故障事件的基本信息、故障现象、故障原因、故障模式和故障影响等风险因素。学习和训练系统风险事件, 即使使用图卷积神经网络训练知识图谱包含的所有三元组信息, 通过捕获事件中邻域的特征向量, 使得事件实体特征获得了更加全面的语义信息, 可以使得系统风险分析更加全面和准确。因此避免出现“知识孤岛”的情况, 解决了分析片面故障原因和故障影响, 个例事件存在的信息不对称和不完整等问题。可以发现系统的风险事件的基本信息和各种风险因素所反映的真实情况和规律。

相比于传统表单式风险分析方法, 基于知识图谱的统计分析更加自动和便捷, 可使分析效率提升 60% 以上, 其可视化分析结果, 便于工作人员进一步加深理解。建立起发动机系统风险分析数据知识库, 可为相关工作人员提供可靠的风险分析和管控措施, 也为真实工作中的预防发动机故障风险事件的发生提供帮助。

2.4 发动机系统风险推理结果与分析

1) 模型学习

通过调用知识图卷积网络对发动机系统风险知识图谱包含的基本三元组信息的学习, 全方位考虑故障信息与发生原因之间的潜在关系, 从而根据故障事件的风险因素推理出引发事故发生的根本原因以及引发的事故影响, 提前采取相关手段和措施, 减少故障发生带来的损失。

对于超参数, 采取网格搜索的方法。训练

批次从 [10,20,30,40,50] 中选取, 采样大小 K 从 [1,2,3] 中选取, 学习率从 [0.1,0.01,0.001,0.000 1] 中选取。通过训练和选取, batch_size 为 10, 采样大小 K 为 2, 学习率为 0.01。另外, 激活函数是 ReLU, 聚合函数采用加法聚合

$$Z_{\text{sum}} = \sigma [W \cdot (v + v_{S(v)}^u) + b] \quad (6)$$

2) 推理结果与分析

图 15 是本文通过知识图卷积网络对高压压气机静子放气阀发生漏气事件与类似故障事件一起学习, 得到的发生故障的原因以及产生的影响。

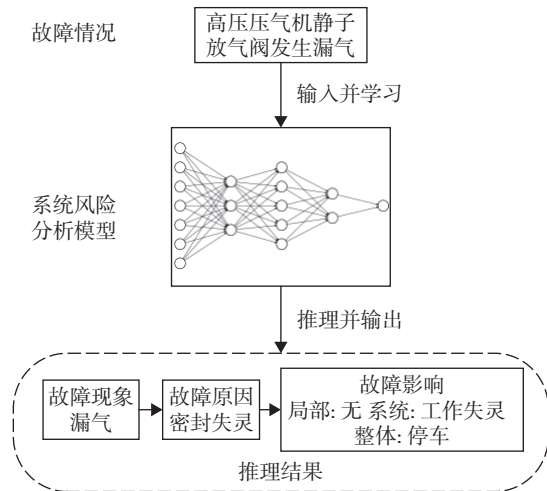


图 15 放气阀漏气事件推理结果

Fig. 15 Inference result of the air release valve leakage event

起初获取高压压气机静子放气阀漏气事件的特征向量, 将该实体信息的三元组与发动机系统风险知识图谱中的所有事件共同学习和训练。当该实体与邻域三元组的聚合学习过程中, 其邻域中的实体与更外层的事件相关联, 从而使得故障事件高压压气机实体静子放气获得更多的语义联系。然后对事件实体的三元组进行卷积操作, 高压压气机实体静子放气阀漏气的更新特征融合了已有事件的特征和规律。最后根据已知头部实体、故障现象和关系、故障模式, 推理出尾部实体, 也就是高压压气机实体静子放气阀漏气可能的故障原因和故障影响。

为了验证推理的准确性和合理性, 采用故障树 (fault tree analysis, FTA) 方法由上往下对高压压气机静子放气阀漏气事件进行分析, 可以发现漏气原因之一就是密封失灵, 其分析结果如图 16 所示。

高压压气机实体静子放气阀漏气产生的故障原因是密封失灵; 故障的局部影响是没有影响,

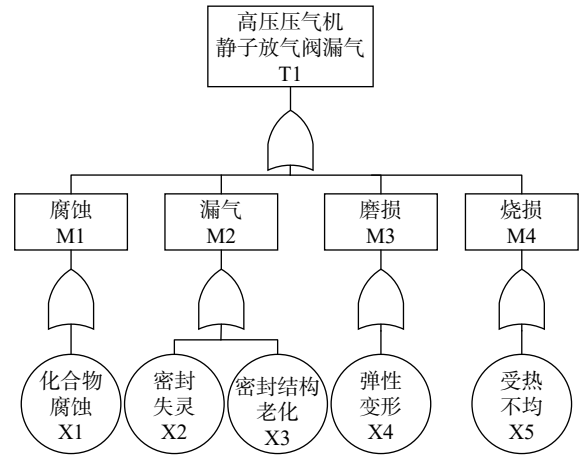


图 16 高压压气机静子放气阀漏气故障树

Fig. 16 FTA of the air release valve leakage event

系统影响是工作失灵以及整机影响是系统停车。根据对事件高压压气机实体静子放气阀漏气的推理, 相关维修人员需要对放气阀采取修复或者更换等措施, 防止故障的发生。

可以看出, 基于 KG-KGCN 的风险推理, 有效解决了传统风险分析方法侧重于事后原因分析, 难以对系统风险开展主动预防和及时更新的难题, 可及时对风险原因及后果进行推理和预测, 实现从事后分析向主动预防的转变, 有效提升了动力装置的运行管理水平。

3 结 论

1) 以系统 FMECA 数据为基础, 构建系统风险知识图谱。首先对系统进行结构分析, 约定系统层次进行 FMECA 分析; 接着对系统确立实体、属性和关系类型, 从 FMECA 中提取事件的相关信息以及风险因素构成知识图谱三元组; 然后使用 Neo4j 对系统风险知识图谱完成存储、可视化和查询等, 易于掌握和理解, 有效解决了传统 FMECA 分析方法存在的表单繁琐、容易出错、难以分析和理解等问题, 结果更加准确和全面, 可使分析效率提升 60% 以上。

2) 提出 KG-KGCN 的系统风险分析方法。将知识图卷积网络应用于系统风险分析的推理和预测, 在分析系统风险知识图谱特点的基础上, 将知识图谱中的实体和实体之间的关系向量化, 通过学习实体包含的风险因素的三元组信息, 发现事件之间存在的潜在关系和演变特征, 从而推理出可能引起风险事件发生的原因以及产生的影响, 以提前采取相关措施避免风险事件的发生, 克服了 FMECA 等传统风险分析无法充分数据、

分析结果单一、难以考虑系统复杂性相关性等问题。本文提出的方法可有效实现复杂装置系统风险分析及精准管控,为其智能运维奠定了坚实的基础。

参考文献:

- [1] BAHREBAR S, BLAABJERG F, WANG Huai, et al. A novel type-2 fuzzy logic for improved risk analysis of proton exchange membrane fuel cells in marine power systems application[J]. *Energies*, 2018, 11(4): 721.
- [2] MORENO V C, COZZANI V. Integrated hazard identification within the risk management of industrial biological processes[J]. *Safety Science*, 2018, 103: 340-351.
- [3] DUNJÓ J, FTHENAKIS V, VÍLCHEZ J A, et al. Hazard and operability (HAZOP) analysis. A literature review[J]. *Journal of Hazardous Materials*, 2010, 173(1/2/3): 19-32.
- [4] 徐立, 邓儒超, 徐楚, 等. 船舶核动力装置二回路故障分析及安全性研究[J]. 哈尔滨工程大学学报, 2015, 36(1): 130-133.
XU Li, DENG Ruchao, XU Chu, et al. Failure analysis and safety studies of the second loop system of civil marine nuclear power plant[J]. *Journal of Harbin Engineering University*, 2015, 36(1): 130-133. (in Chinese)
- [5] 屠秋野, 张伟, 蔡元虎, 等. 民用航空动力技术风险评估的量化方法[J]. 航空动力学报, 2010, 25(11): 2403-2408.
TU Qiuye, ZHANG Wei, CAI Yuanhu, et al. Quantity method of technical risk assessment for civil aero-engine[J]. *Journal of Aerospace Power*, 2010, 25(11): 2403-2408. (in Chinese)
- [6] AMIT S. Introducing the knowledge graph: things, not strings[EB/OL]. [2022-02-15]. <https://www.blog.google/products/se-arch/introducing-knowledge-graph-things-not>.
- [7] KRIZHANOVSKY A A, SMIRNOV A V. An approach to automated construction of a general-purpose lexical ontology based on Wiktionary[J]. *Journal of Computer and Systems Sciences International*, 2013, 52(2): 215-225.
- [8] 魏卉子. 煤矿安全融合知识图谱构建研究[D]. 江苏 徐州: 中国矿业大学, 2020.
WEI Huizi. Study on the construction of coal mine safety integration knowledge map[D]. Xuzhou Jiangsu: China University of Mining and Technology, 2020. (in Chinese)
- [9] IYYER M, BOYD-GRABER J, CLAUDINO L, et al. A neural network for factoid question answering over paragraphs[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Stroudsburg, US: Association for Computational Linguistics, 2014: 633-644.
- [10] 邱凌, 张安思, 李少波, 等. 航空制造知识图谱构建研究综述[J]. 计算机应用研究, 2022, 39(4): 968-977.
QIU Ling, ZHANG Ansi, LI Shaobo, et al. Survey on building knowledge graphs for aerospace manufacturing[J]. *Application Research of Computers*, 2022, 39(4): 968-977. (in Chinese)
- [11] 郭静. 基于知识图谱的航空安全事件关联分析方法研究[D]. 天津: 中国民航大学, 2020.
GUO Jing. Research on the correlation analysis of aviation safety incidents based on knowledge graph[D]. Tianjin: Civil Aviation University of China, 2020. (in Chinese)
- [12] 杜党党, 贾晓亮, 王卓. 基于故障彩色图谱的航空发动机故障模式识别[J]. 计算机集成制造系统, 2018, 24(9): 2297-2305.
DU Dangdang, JIA Xiaoliang, WANG Zhuo. Fault model identification method for aero engines based on fault color picture[J]. *Computer Integrated Manufacturing Systems*, 2018, 24(9): 2297-2305. (in Chinese)
- [13] BRUNA J, ZAREMBA W, SZLAM A, et al. Spectral networks and locally connected networks on graphs[EB/OL]. [2021-12-08]. <https://arxiv.org/abs/1312.6203>.
- [14] WANG Hongwei, ZHAO Miao, XIE Xing, et al. Knowledge graph convolutional networks for recommender systems[C]//Proceedings of the 19th World Wide Web Conference. New York: ACM, 2019: 3307-3313.
- [15] ZHANG Ya, LU Mingming, LI Haifeng. Urban traffic flow forecast based on FastGCRNN[J]. *Journal of Advanced Transportation*, 2020, 2020: 1-9.
- [16] KAPOOR A, BEN Xue, LIU Luyang, et al. Examining COVID-19 forecasting using spatio-temporal graph neural networks[EB/OL]. [2021-12-15]. <https://arxiv.org/abs/2007.03113>.
- [17] 刘欢, 李晓戈, 胡立坤, 等. 基于知识图谱驱动的图神经网络推荐模型[J]. 计算机应用, 2021, 41(7): 1865-1870.
LIU Huan, LI Xiaoge, HU Likun, et al. Knowledge graph driven recommendation model of graph neural network[J]. *Journal of Computer Applications*, 2021, 41(7): 1865-1870. (in Chinese)
- [18] 史运涛, 党亚光, 雷振伍, 等. 基于图神经网络的社区户内燃气系统动态风险评估[J]. 安全与环境工程, 2021, 28(5): 1-9.
SHI Yuntao, DANG Yaguang, LEI Zhenwu, et al. Dynamic risk assessment of community indoor gas system based on graph neural network[J]. *Safety and Environmental Engineering*, 2021, 28(5): 1-9. (in Chinese)
- [19] 孟寒姣. 基于知识图谱的铁路电务事故智能问答系统的设计与实现[D]. 石家庄: 河北科技大学, 2020.
MENG Hanjiao. Design and implementation of intelligent question answering system for railway electrical accidents based on knowledge graph[D]. Shijiazhuang: Hebei University of Science and Technology, 2020. (in Chinese)
- [20] 何巍. 社交媒体数据可视化分析综述[J]. 科学技术与工程, 2020, 20(32): 13085-13090.
HE Wei. An overview on social media visual analytics[J]. *Science Technology and Engineering*, 2020, 20(32): 13085-13090. (in Chinese)
- [21] 王红, 林海舟, 卢林燕. 基于 Att_GCN模型的知识图谱推理算法[J]. 计算机工程与应用, 2020, 56(9): 183-189.
WANG Hong, LIN Haizhou, LU Linyan. Knowledge graph inference algorithm based on Att_GCN model[J]. *Computer Engineering and Applications*, 2020, 56(9): 183-189. (in Chinese)

(编辑: 王碧珺)