

# 专用网络异常诊断机制的设计与实施

王中仁、张玉阳、董平、张晓雅

北京交通大学电子与信息工程学院 21120143@bjtu.edu.cn, zhyy@bjtu.edu.cn, pdong@bjtu.edu.cn,  
18111040@bjtu.edu.cn

**摘要**—随着网络技术的不断发展, 人们正在尝试构建各种专用网络, 用于加密通信和应急通信场景。路由设备是专用网络使用中必不可少的核心设备, 监控路由设备的工作状态, 诊断并修复异常情况, 对于确保通信的快速恢复, 提高通信系统的稳定性至关重要。然而, 随着人们对通信需求的不断提高, 网络设备逐渐向复杂化、定制化和专业化方向发展。在不同场景下, 网络设备的结构和功能可能会有很大差异, 这给用户和维护人员诊断网络故障带来了很大困难。因此, 为了辅助决策, 实现对网络设备及时、准确的故障诊断和维护, 提出了一种基于知识图谱的故障诊断机制。基于网络设备的各类故障知识, 构建知识图谱, 根据异常现象从知识图谱中定位可能的原因和可能性, 并针对每个原因提供相应的解决方案。

**关键词**—网络设备; 故障诊断; 知识图谱

## I. 导言

专用网络技术在日常工作和生活中的应用越来越广泛[1-2]。然而, 作为网络连接媒介的路由设备, 其结构和功能的复杂程度随着人们的需求不断提高。一旦出现异常, 故障诊断和修复所需的时间和成本将逐渐增加, 导致网络通信系统受到影响。因此, 有必要设计一种可靠的异常诊断机制, 以提高网络通信系统的可靠性。网络系统的维修效率。

在网络协议中, SNMP 协议 [3] 可用于故障诊断。不过, SNMP 更适用于传统的通用网络和多网络节点的异常诊断和管理。在专用网络的路由设备中, 网络通信程序、软硬件和整个系统结构都是定制的, 使用 SNMP 协议来排除可能存在的不兼容问题。

随着人工智能技术的发展, 知识表示和分析

技术已经有了很大改进。比较合适的技术是知识图谱[4]。它可以通过语料库构建图结构关系, 从而以通用的方式融合现有的故障知识, 并根据图结构进行故障诊断。知识图谱在故障诊断中的应用非常广泛。在工业物联网的电机故障诊断场景中, 利用知识图谱和卷积神经网络设计了一种故障诊断机制[5]。在云平台的运维场景中[6], 知识图谱的使用改进了原有的故障跟踪系统, 提高了溯源结果的准确性。在工程应用中的旋转机械设备领域, 提出了基于本体的故障诊断模型, 并基于耶拿工具库实现了知识推理[7]。

然而, 在网络领域利用知识图谱进行故障诊断推理的研究相对较少。由于知识图谱的关系性, 其在故障知识存储和表达推理方面具有优势, 可以深入挖掘已有的故障知识, 建立通用的故障表示方法和诊断模型。本文针对专用网络中的定制化网络设备, 提出了一种基于知识图谱的问答诊断机制, 并介绍了其在专用网络诊断场景中的应用。

## II. 基于知识图谱的故障诊断机制介绍

### A. 机制结构

专网路由设备的故障信息文本类型多样, 包含了大量计算机、通信等领域的专业术语。有必要利用知识图谱的智能方法, 充分挖掘和管理以往技术人员对该设备的调试经验信息, 有利于技术人员和用户今后对网络设备的维护和管理。

本文通过源数据构建了特种网络设备的故障诊断机制。

如图 1 所示，知识图谱包括数据层、数据处理层、知识图谱构建层和业务层。

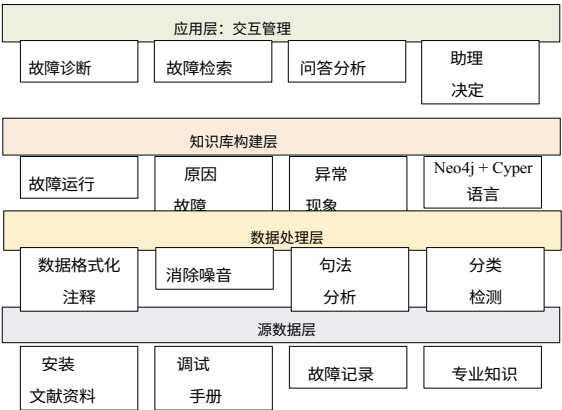


图 1. 知识图谱构建框架

数据层是整个知识地图构建的基础，主要包括故障分析记录、设计文件、设备安装手册以及开发和维护人员的专家经验等，并将这些内容作为构建层的基础数据。

数据处理层是构建知识地图的重要组成部分。通过使用标签分类和构建分类模型、数据文件格式化、噪声去除、语法分析等操作，将元数据转化为便于计算机程序处理的数据。

构建层包括以下内容：首先，根据原始数据和业务表单设计合理的知识图谱结构；其次，将知识图谱原始数据中的故障现象条目分为设备类型、异常运行、异常现象和异常报警等四个维度，使异常描述更适合知识图谱结构；最后，根据数据层的基础数据进行拆分，得到知识图谱中各节点的数据集和节点间三重关系的数据集。并使用赛弗尔语言完成知识图谱的构建。

应用层是指利用构建的知识图谱解决相应的异常诊断任务，从而达到辅助人工决策的目的。本文设计了异常原因概率判断的优化机制，与基本的 Cyper 查询相比，可以提高诊断的准确性和结果的完整性，同时利用图结构展示诊断过程中的推理过程。此外，应用层还支持基本的知识图谱推理问答和专网故障搜索与管理功能。

## B. 设计知识图谱结构

通过对专用网络路由设备的诊断知识本体信息和重要概念进行分类整理，可以剔除与故障诊断关联性差、冗余的知识。将故障诊断的相关知识实体归纳为设备信息和异常信息。其中，设备信息类分为网络设备类型和异常设备位置两个子类；异常信息分为异常现象、异常告警、异常操作、异常分析、异常原因五个类别。

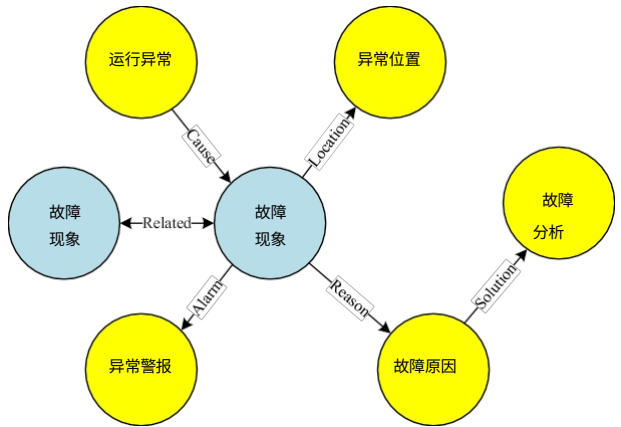


图 2. 知识地图结构

在划分和构建故障知识实体类别后，根据应用场景构建了实体之间的关系模板。通过匹配模板中的关系关键字，将不同类型的实体联系起来。构建的数据通常以节点、关系和结点的形式存储在图数据库中。根据对故障机制的分析和诊断知识的描述，可以总结出以下几种关系类型。构建的结构如图 2 所示。

原因：表示异常现象与异常运行之间的从属关系。表示为：（异常运行）-[原因]->（故障现象）；

报警信息：表示异常现象与报警信息之间的从属关系。表示为(故障现象)-[警报]->（异常报警）；

异常部分：表示异常现象与异常部分之间的从属关系。表示为(异常现象)-[位置]->（异常部位）；

直接/间接原因：表示异常现象与异常原因之间的从属关系。表述为(异常现象)-[原因]->（异常原因）；

相关性：表示一种异常现象与另一种异常现象之间的从属关系。表示为(异常现象) - [相关] -> (异常现象)；

### C. 异常诊断与异常问题解答机制设计

#### 1) 基于相似性的实体对齐

在根据输入描述搜索知识图谱的过程中，由于用户的异常描述与知识图谱中的数据存在偏差，很难直接将描述与知识图谱中的节点内容进行匹配。本文采用基于属性相似性的实体配准方法，有效提高了实体配准的准确性，提高了专网知识地图的界面处理效率。

首先，计算词频向量。利用自然语言处理技术将输入的两个句子分割成单词，然后对联合集进行去重复处理。在统计了要比较的两个句子的词频后，根据联合集计算出词频向量 A 和 B，如式 (1) 所示。

$$\text{节点}_A = [v_1^A, v_2^A, \dots, v_n^A]$$

A 和 B 结合后的维数， $v_i^A$  表示

A 中出现的第 i 个词在两个词条结合处的频率。

将实体名称 A 和 B 转化为词向量 A 和 B 后，根据余弦相似度计算两个名称属性之间的相似度。余弦值越接近 1，表示包含的角度越接近 0 度，即两个名称属性越相似。

载体。其中有  $\|\text{节点}_A\|$  和  $\|\text{节点}_B\|$  代表分别表示两个句子向量的模块长度。 $\text{Sim}(A,B)$  表示句子 A 和句子 B 之间的相似度。

$$\text{Sim}(A,B) = \frac{\|\text{节点}_A\| \cdot \|\text{节点}_B\|}{\sqrt{\sum_{i=1}^n (v_i^A)^2} \times \sqrt{\sum_{i=1}^n (v_i^B)^2}} = \frac{\sum_{i=1}^n v_i^A \times v_i^B}{\sqrt{\sum_{i=1}^n (v_i^A)^2} \times \sqrt{\sum_{i=1}^n (v_i^B)^2}} \quad (1)$$

#### 2) 类别分割

根据知识图谱的结构，异常现象与异常运行、异常报警和异常部件有关，而用户输入的描述可能比较复杂并包含一些无关信息。为了提高异常诊断的准确性，需要对复杂的异常描述输入进行一定的处理。

为了对输入的故障描述进行更准确的分类，需要对其进行预处理，包括单词分割和停顿词过滤，然后由 CNN 对预处理后的输入异常描述进行分类，即

分为三类：设备类型、执行操作和故障现象。

对于分离执行的操作：使用第 III 节中的相似性比较方法，与异常操作中的相似性进行比较。

然后，根据集合 X 中的元素，按照 (异常操作) - [原因] -> (故障现象) 的关系，从知识图谱中查询异常现象集合 A。

对于分离出来的故障现象，通过分词技术从异常现象中获取故障位置和故障代码。然后利用图数据库，根据 (故障现象) - [位置] -> (异常位置) 和 (故障现象) - [警报] -> (异常警报) 之间的关系，从知识图谱中查询异常现象集 B 和异常现象集 C。

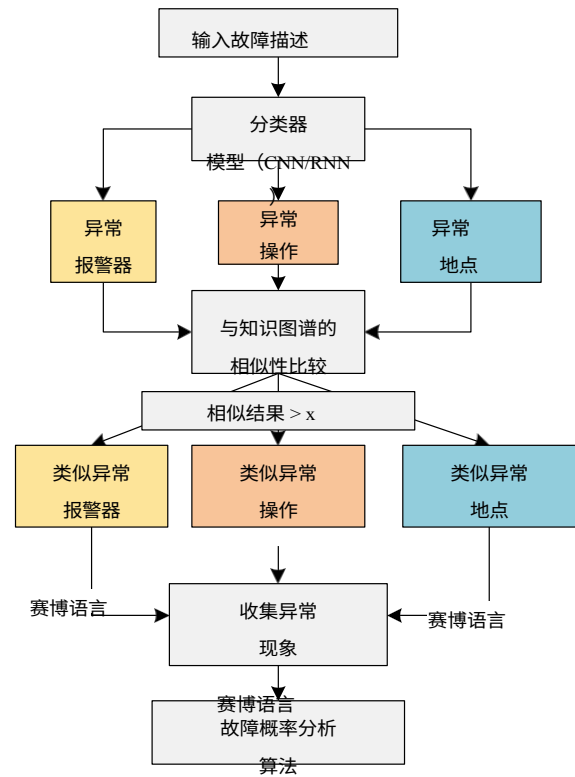


图 3. 分类匹配流程图

然后合并查询到的异常现象集。由于前面提到的图中涉及到三种关系，因此有三种类型的集 (A,B,C) 需要合并，如式 (2) 所示：其中 x 为阈值， $A_i$  为描述集， $\text{get\_type\_items}()$  是根据节点类型从知识图谱中查询该类型所有元素的操作， $\text{get\_items}()$  是根据关系使用 Cyper 语言查询关系另一端元素集的操作， $\text{get\_similar\_items}()$  是计算

第一参数和第二参数之间的句子相似度，并保存大于  $x$  的阈值。

$$a = \bigcup_{i=1}^3 \text{获取项目} \quad (2)$$

$$(A_i, \text{get\_type\_items}(A_i), x)$$

最后，对于用户输入分类的异常点，通过第 1) 节中的方法比较上一步合并的异常集  $a$  的相似度，以集合  $A$  中元素阈值大于  $x$  作为最终异常集  $c$ ，如式 (3) 所示：

$$c = \text{获取\_类似\_项目}(a, b, x) \quad (3)$$

通过以上步骤，得到处理后的异常现象集。该集合经过自然语言处理，同时更适合知识图谱，输入的现象也得到了合理的扩展，增强了系统的智能性。整体流程如图 3 所示。

### 3) 最佳拟合概率判断异常的诊断算法

对于处理后的输入异常现象集，我们可以从知识图谱结构中得知，异常现象和异常原因是多对多的关系，每个现象可能对应多个原因。同时，知识图谱中的一个异常原因  $C_x$  也可能对应多个输入的异常现象  $A_i$ ，如图 4 所示。这个数字是通过基于 Cyper 语言的查询得到的，计为  $num$ 。对于这些输入异常的集合  $A$ ：[ $A_1, A_2, \dots, A_n$ ] ( $1 < i < n$ )，存在以下三元组关系： $A_i - [\text{reason}] \rightarrow C_i$  ( $C_i \in C$ )，适用于任意  $A_i$ 。  $C$  是知识图谱中与输入现象集相对应的异常原因集。同样，根据 Cyper  $n$  与  $y$  的商为用户输入的异常现象的各异常原因的重合概率。该值越大，代表异常原因的概率越高，从而可以根据图谱关系，根据异常原因找到相应的解决方案。

从上述机制可以知道，最佳判断概率可以根据输入现象的重合数和知识图谱中对应现象的数量，最后将超过阈值  $x$  的若干  $p_i$  按逆序排列，然后每个实体与这些  $p_i$  对应的是一组可能的异常原因，如式 (4) 所示， $\text{get\_counts}()$  是具有关系  $rel$  的查询实体  $a_i$  的数量。

$$P = [p_1, \dots, p_n], (p_i = \frac{n}{\text{获取\_计数}(a_i, rel)}, p \geq x, a_i \in A) \quad (4)$$

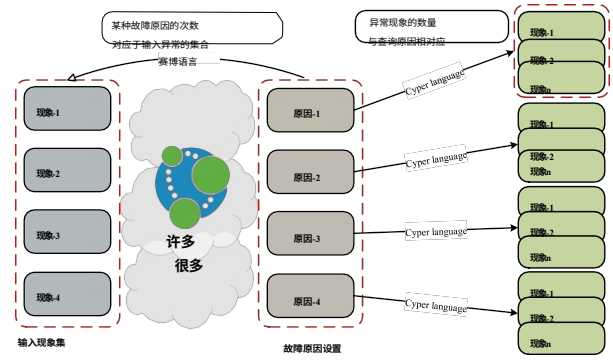


图 4. 基于拟合概率的异常诊断方法

## III. 专用网络特殊路由器场景中的构造示例和应用

### A. 专用网络路由设备故障数据介绍

以智能汇聚网络路由设备典型异常案例和近十年调试记录数据为原始语料，整理出800余条数据，利用第II部分提出的知识图谱构建方法，构建了专网异常诊断知识图谱。下表以非结构化语料为例，通过异常描述定义了执行操作、异常告警、异常位置和异常现象四种实体类型。其次，定义了异常原因和解决方案两种实体类型。如表I所示。

表 I. 语料库数据示例

类型	数据示例
异常原因	设备没有拨号，也没有连接互联网。获取时间的命令是启动项目没有时间。
异常描述	盒式设备在断电和重新启动后，时间将恢复为 0。
解决方案	重复获取时间命令。

表 II. 结构化数据实例

类型	数据示例
异常 [原因]-> 异常现象	如果没有通话和服务，所有网卡信号都将消失。
异常现象-[位置]-> 异常反应	拨通号码后，界面屏幕也会显示灰色、位置、屏幕。
异常现象-[相关]-> 异常现象	启动时数据库连接失败，显示 1044 错误码，拒绝访问，相关、远程机器没有访问权限。
异常现象-[原因]-> 异常原因	界面显示卡槽未插入，原因是 ttyUSB 未在卡槽中驱动，无法拨号上网（ttyUSB 未完全插入）。



基于语料库数据，"章节"中的实体分类方法可以获得知识图谱节点数据，并在此基础上获得节点间关系的结构化数据，如表"所示，包括五种类型的节点和五种三元组的推理规则。

B. 实验环境和描述 分类效果分析

实验使用的操作系统是 Linux ubuntu4.4.0，处理器是 Intel (R) Xeon (R) CPU E5- 26xv4，基于 python3.6.12 的深度学习框架是 tensorflow 1.14.0。数据经过预处理后分为训练集、测试集和验证集，三者的比例分别为 80%、10% 和 10%。

在""部分，使用了深度学习分类技术。因此，需要提取非结构化数据中的异常描述，进行语义拆分，再进行人工标注。主要可以将设备类型、异常现象和异常操作进行分离，异常报警和异常部件的分离比较简单，可以通过正则化和字典来完成。因此，本节主要讨论前三类要素的分类效果。

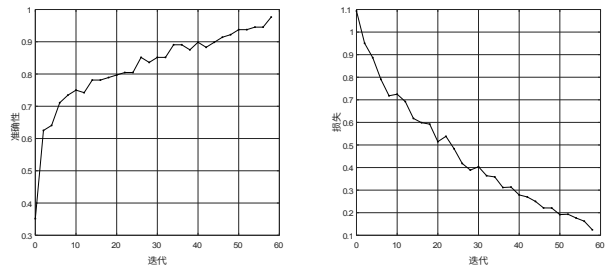


图 5.不同迭代下精度和损失函数值的变化。

经过多次试验，最终确定的模型参数如表III所示。该参数下损失函数的迭代效果和精度如图 5 所示。

表 III. CNN 模型的分类参数

参数	设定值	解释
embedding_dim	64	词向量维度
序列长度	600	序列长度
过滤器数量	256	卷积核数量
内核大小	五个	卷积核大小
hidden_dim	128	全连接层神经元
辍学率	0.5	辍学率
学习率	0.001	学习率
批量大小	64	每批训练的规模
数次	20	迭代轮总数

C. 知识地图的构建效果和业务应用

根据智能融合路由设备的调试记录，提取路由设备异常实体 224 个，其中异常运行实体 53 个、异常部件 24 个、异常原因 60 个、异常现象实体 67 个、异常告警 20 个。然后，根据实体之间的关系，按照五条三重推理规则建立了如图 6 所示的完整知识图谱，节点之间的关系共计 549 条，为后续应用提供了数据基础。

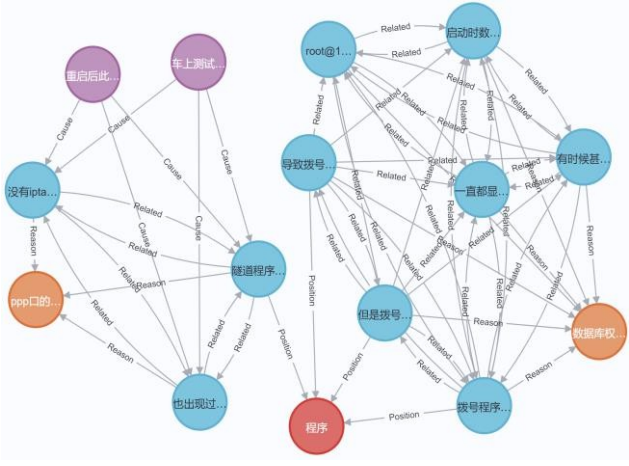


图 6.知识图谱构建示例

基于专网路由设备故障知识图谱的内涵及其对异常诊断场景的适应性，本研究认为专网异常诊断知识图谱将有以下应用：异常原因诊断、在线问答、维护措施查询。

在线问答：在专用网络知识图谱的基础上，通过在线用户将一些简单的单跳关系呈现给用户和运维人员，帮助他们了解知识图谱对应设备的故障知识。根据知识图谱的结构，单跳知识图谱的简单问答可以分为以下四类：发生故障会造成哪些现象？执行操作时会遇到哪些错误？某部件常见的故障有哪些？故障报警的含义是什么？例如，当用户想知道路由设备管理界面有哪些常见故障时，可以直接输入到知识地图后台，根据输入的问答方式匹配不同的Cyper语句进行查询，便于操作和判断。如表所示

IV.

表 IV. 在线问答流程示例

答案集	查询集
某一区域经常出现哪些故障	MATCH (n1)-[rel {type:" location"}]->(n2 {title: "area"})返回 n1,rel,n2
在某个 某些操作	匹配 (m:" 非正常操作") 返回 m 输入 m -> 实体对齐 -> 输出 n MATCH (n1 {title:m})-[rel {type: "Cause"}]-> (n2) 返回 n1,rel,n2
什么原因会导致什么现象? 故障	匹配 (m: "原因") 返回 m 输入 m -> 实体对齐 -> 输出 n MATCH (n1)-[rel {type: "Cause"}]-> (n2 {title:n}) 返回 n1,rel,n2

异常原因诊断和异常解决方案查询：基于专网路由设备异常诊断部分的知识图谱，用户和运维人员可根据设备的异常现象描述异常情况并输入知识图谱，根据历史故障经验知识查询可能的异常原因，并结合设备的实际情况推断可能的异常原因。

例如，异常描述为 "界面显示异常，所有卡均无法拨号，无信号？"针对该问题，先用上述方法得到异常现象集，然后对异常现象集进行分类，再用 " "部分的方法得到最终结果，然后根据异常原因和解决方案之间的关系，对每种可能的原因给出相应的解决方案。整体流程如下图所示。故障原因查询和异常解决方案查询可以为今后提高特种网络设备的维护效率和运行可靠性提供参考信息。如表 V 所示。

表 V. 诊断过程示例

计划流程	示例集
分类	界面显示异常；所有卡均无法拨号；无信号
Match1：查询所有故障与现象 x 相对应的原因	匹配 (n1 {无信号}) - [rel {类型: "原因"}]-> (n2)-> 返回 n1,rel,n2
Match2：与故障原因 x 与故障原因 x	匹配 (n)-[r {类型: " 网络访问模块 硬件故障"}]->(m {title: "Cause"}) WITH count(n) as relationCount 返回关系数
诊断结果包括故障原因、概率失败的原因和解决办法	原因1:HTTP 服务器程序异常退出；可能性: 0.67；解决方案: 数据库处理错误人脉

IV. 结论

本文提出的基于知识图谱的异常诊断机制有效克服了传统故障诊断方法存在的问题，充分利用了以往的故障诊断经验，提高了专网中定制化网络设备故障诊断过程的智能性和效率，提高了专网运行的可维护性。

鸣谢

本研究部分得到中国博士后科学基金（2021M690343）和中央高校基本科研业务费（2023JBGP005）的资助。

参考资料

[1] J.Nemčik, P. Kánuch and I. Kotuliak, "Content Distribution in Private Networks," 2022 International Symposium ELMAR, Zadar, Croatia, 2022, pp、

[2] S.Budiyanto, C. S. Aprihansah, L. M. Silalahi, I. U. Vistalina Simanjuntak, F. A. Silaban and A. D. Rochendi, "Auto Discover Virtual Private Network Using Border Gateway Protocol Route Reflector," 2022 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), Solo, Indonesia, 2022, pp、

[3] E.P. Duarte 和 T. Nanya, "基于 SNMP 的局域网故障管理自适应分布式系统级诊断算法实施", NOMS '96 - IEEE 网络运行与管理研讨会论文集, 日本京都, 1996 年, 第 530-539 页, 第 2 卷、

[4] K.Bollacker, R. Cook, and P. Tufts, "Freebase : A shared database of structured general human knowledge," in AAAI-07/IAAI-07 Proceedings: 22nd AAAI Conference on Artificial Intelligence and the 19th Innovative Applications of Artificial Intelligence Conference, July 22, 2007 - July 26, 2007, Vancouver, BC, Canada, 2007, vol. 2: American Association for Artificial Intelligence, in Proceedings of the National Conference on Artificial Intelligence, pp.

[5] D.Geng, L. Zhou, Y. Zhao and H. Li, "Fault Diagnosis of Motor based on Knowledge Graph and CNN," 2022 第三届国际大数据、人工智能和物联网工程大会 (ICBAIE), 中国西安, 2022 年, 第 742-745 页、

[6] Y.Yang and L. Yang, "Cloud Platform Component Fault Tracing System Based on Operation and Maintenance Knowledge Map," 2022 2nd International Conference on Algorithms, High Performance Computing and Artificial Intelligence (AHPCAI), Guangzhou, China, 2022, pp、

[7] R.Chen 等人, "基于本体的旋转机械故障诊断知识建模", 2015 IEEE 第 13 届国际工业信息学会议 (INDIN), 英国剑桥, 2015 年, 第 1050-1055 页。