

翁少轩,王奋,刘志钢. 地铁运营调度应急处置知识图谱构建[J]. 智能计算机与应用, 2025, 15(4): 198–203. DOI: 10.20169/j. issn. 2095–2163. 250428

地铁运营调度应急处置知识图谱构建

翁少轩, 王 奋, 刘志钢

(上海工程技术大学 城市轨道交通学院, 上海 201620)

摘 要:城市轨道交通调度员需依据实际运营状况填写调度日志。日志中通过故障现象、处置方法、事件对运营的影响情况等信息描述故障处置事件。针对当前调度细则中构建的故障场景与实际情况存在偏差不够全面,应急处置方法匹配效率低等问题,调度日志中的填记的内容更贴近实际故障场景,相关处置操作具有更高的参考价值。本文以某地铁公司某条线路调度日志中的应急处置文本构建应急处置知识图谱,提供信息检索功能,利用知识图谱丰富的语义表达能力展示不同故障事件的演变情况,为调度员提供相关处置信息。

关键词: 地铁调度; 应急处置; 信息检索; 知识图谱

中图分类号: U298

文献标志码: A

文章编号: 2095–2163(2025)04–0198–06

Construction of knowledge graph for subway operation dispatch and emergency response

WENG Shaoxuan, WANG Fen, LIU Zhigang

(School of Urban Railway Transportation, Shanghai University of Engineering Science, Shanghai 201620, China)

Abstract: The dispatcher of urban rail transit needs to fill in the dispatching log according to the actual operation status. The log describes the fault handling events through information such as fault phenomena, disposal methods, and the impact of events on operations. In view of the problems that the deviation between the fault scenario constructed in the current dispatching rules and the actual situation is not comprehensive enough, and the matching efficiency of emergency disposal methods is low, the content filled in the dispatching log is closer to the actual fault scenario, and the relevant disposal operation has higher reference value. In this paper, an emergency disposal knowledge graph is constructed based on the emergency disposal text in the dispatching log of a certain line of a subway company, which provides information retrieval function. The rich semantic expression ability of the knowledge graph is used to show the evolution of different fault events and provide relevant disposal information for dispatchers.

Key words: subway dispatch; emergency disposal; information retrieval; knowledge graph

0 引 言

城市轨道交通运营调度主要负责行车运营监护、行车组织调整和处理突发事件^[1–2]。随着城市轨道交通突发事件的处置效率和安全要求的不断提高,如何快速调用应急预案和以往的处置经验就显得尤为重要。

为有效降低突发事件带来的影响,实现运营快速恢复,地铁运营机构依据不同类型事件制定了调度工作细则和应急预案。调度员需熟练背记不同突发事件的处置措施,针对常见故障往往凭借经验及背记的内容做出决策,遇到偶发故障时需要检索细

则中的处置措施。然而,单纯的背记不仅工作量巨大且存在遗忘的风险,在检索时以非结构化形式存储的调度细则和预案文本也很难直接为调度员提供有效处置信息。

随着运营场景的不断细分,调度细则及预案中设定的情形与突发事件场景的偏差不断扩大。同时,调度规程中的要求也在逐渐增加,仅凭人工经验处置、人工经验检索信息的方式已无法满足高质量的运营需求,人工智能技术在调度应急处置领域的应用成为了发展趋势^[3]。

作者简介: 翁少轩(1998—),男,硕士研究生,主要研究方向:自然语言处理,知识图谱。

通信作者: 王 奋(1986—),女,博士,副教授,主要研究方向:城市轨道交通运营安全。Email: wangf_2016@163.com。

收稿日期: 2023–09–08

1 调度应急处置现状分析

生产管理中应急处置可归纳为事件触发、方案拟定、方案执行、运营恢复四个阶段。当接收到告警信息时调度员需结合不同故障场景对事件进行预判并选择合适的应急处置预案;预案启动后调度员依据实际运营状况拟定行车调整方案、现场抢修方案及客运组织方案并按照上级指示调整方案,确定后立即实行;最终调度员依据事件影响、应急处置及运营恢复情况向相关单位申请预警解除。

调度日志记录运营中发生的各类运营事件,调度员需如实填写运营事件发生的时间、地点,涉及故障处置事件需填记故障场景、处置措施、事件影响等信息。相比于调度细则及预案中对于故障现象的简单描述,调度日志文本中记录的内容更为详细贴近实际生产场景。因此本文利用日志文本构建地铁运营调度处置知识图谱,实现了调度应急处置信息的结构化存储,利用 Neo4j 图数据库实现了处置案例的可视化展示,同时提供了信息检索的功能。

2 知识图谱及知识抽取

2.1 知识图谱构建流程

2012 年,知识图谱的概念被谷歌公司首次提出^[4],旨在有效提升数据的管理效率^[5],使用户能够快速准确地找到所需信息。

通用领域知识图谱构建过程主要包括本体构建、实体关系定义、命名实体识别、关系抽取等步骤。首先利用自然语言处理技术对文本进行预处理去除噪音数据;其次依据文本特点构建实体及不同实体间的关系并结合数据特征选择合适的方法完成知识提取,然后利用 Neo4j 图数据库存储信息。最终通过不断更新领域知识完善知识图谱。常见知识图谱的构建过程如图 1 所示。

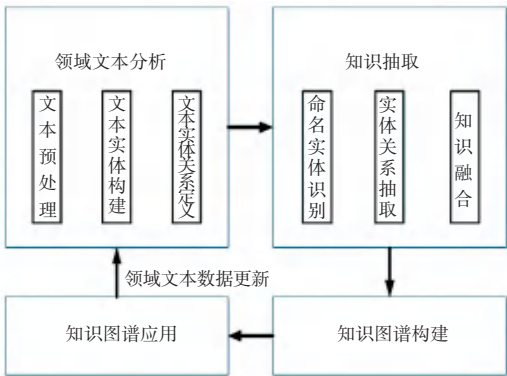


图 1 知识图谱构建流程

Fig. 1 Construction process of knowledge graph

2.2 知识图谱构建方法

知识图谱技术在轨道交通生产管理中应用广泛。郭恒等学者^[6]研究了多域数据融合的列车维修性知识图谱构建,用于辅助设计人员进行可维修性设计。姚剑等学者^[7]构建了面向铁路客站设备健康管理的知识图谱,用于保障设备运行的稳定性。许慧等学者^[8]构建了城市轨道交通运营风险事件知识图谱,优化了风险事件的统计路径及可视化效果。

构建知识图谱的关键在于实现关键信息的抽取,早期的研究中主要使用基于规则和词典的方法来实现。丁君军等学者^[9]通过构建情报学领域中学术概念属性的描述规则对相关领域文献进行属性抽取,最终实验准确率达到 60% 以上。基于规则和词典的抽取方法需耗费大量时间设计规则且泛化程度低。相比之下,基于机器学习的方法有效降低了人工工作量并提高了模型的通用性^[10]。曾军等学者^[11]利用最大熵及支持向量机两种机器学习算法提出了一种支持多语言信息提取的框架。近年来,深度学习算法的不断发展为知识抽取任务带来了全新的机遇。深度学习能够自主学习数据特征,模型具有更高的泛化能力。刘东等学者^[12]利用 BiLSTM-CRF 算法提取信息构建了电网调度知识图谱供调度员信息检索。基于 BiLSTM-CRF 的文本提取方法首先需要将文本内容转变为向量形式,但该方法转变的文字向量并未考虑上下文之间的关系对识别效果具有一定的影响。2018 年,Devlin 等学者^[13]提出了 BERT 预训练模型,有效提升了词向量模型的泛化能力,被广泛应用于自然语言处理技术中。赵鹏飞等学者^[14]分别利用 BiLSTM-CRF 和 BERT-BiLSTM-CRF 模型对农学文本进行信息抽取,实验结果表明采用预训练模型的方法在相同评价指标情况下具有更高的识别准确率。

综上所述,为了实现调度处置信息的可视化展示,提供信息检索功能,本文研究利用基于 BERT-BiLSTM-CRF 的方法实现调度日志中关键信息的识别。

3 数据处理与应急处置知识图谱构建

3.1 日志文本数据分析

调度日志是对当日运营工作的总结,依据记录文本的特征记录的内容分为一般记事、提醒关注及闭环跟踪三类。详细记录内容示例见表 1。

根据表 1 可知,不同类别的记录数据包含了多元化的信息。为了准确提取与应急处置相关的记录内容,需对调度日志中的提醒关注类和一般叙事类

进行筛选,删除与应急处置无关的记录内容。与此同时,利用自然语言处理技术统一文本数据格式。

表 1 调度日志内容
Table 1 Dispatch log content

信息类别	详细信息
闭环跟踪	记录突发事件中设施设备的详细故障情况(列车关门灯不亮)及相应的应急处置措施(退出运行、载客跳停等)
提醒关注	设施设备检查或事件处置已完成
一般记事	记录本班次车调联控的情况 运营线路设施设备测试状况 记录运营规定要求的执行情况 列车设施设备故障或其他运营事件调整措施 因其他客观原因对列车运行进行调整 ...

表 2 处置图谱实体类别及示例
Table 2 Disposal of graph entity categories and examples

实体类别	示例	英文缩写
设备所属单位	车辆	Equipment Unit
设备类型	车载 ATC	Equipment Type
时间	13:43	Time
列车车次号	21801 次车 0266#	Train Number
列车运行方向	xxx 站下行	Direction
列车故障现象	面板显示该扇车门故障	Symptom
处置	发布口头命令 111#准 xx 站上行 21801 次 0266#号车退行至 xx 站上下客	Disposal
事件影响	列车晚点 2 min、造成换表 1 次	Event Influence
事件闭环	驻勤回复经更换电路板后重启设备 正常上试车线路测试正常	Closed Loop

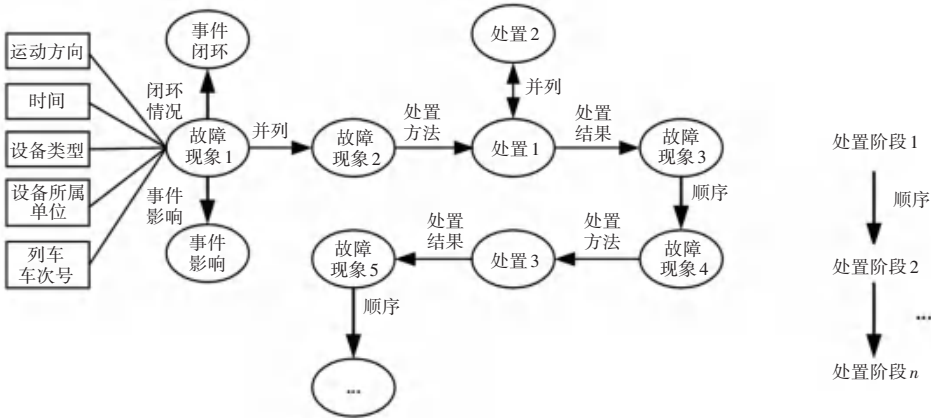


图 2 实体关系

Fig. 2 Entity relationship

- 同一处置事件中调度员处置流程遵循以下步骤：
- (1) 现场人员向调度员报告故障现象。

(2) 调度员依据实际情况采取措施。

(3) 现场人员反馈处置情况。

(4) 调度员与现场人员重复上述 3 个步骤,直到故障解决为止。

3.3 调度日志文本实体识别

将预处理后的调度日志文本按照 BIO 的格式进行标注。B(Begin)代表实体的首个字,I(Inside)代表实体的中间字,O(Other)则代表非实体的文字部分。具体标注方法见表 3。

表 3 调度日志文本 BIO 标注

Table 3 Dispatch log text BIO annotation

文本	标注
5	B-Time
:	I-Time
3	I-Time
2	I-Time
x	B-Direction
x	I-Direction
站	I-Direction
上	I-Direction
行	I-Direction
X	B-TN
X	I-TN
#	I-TN
列	O
车	O
司	B-Symptom
机	I-Symptom
报	I-Symptom
列	I-Symptom
车	I-Symptom
上	I-Symptom
E	I-Symptom
B	I-Symptom

完成文本标注后,按 8:1:1 的比例划分为测试集、验证集和测试集。本文采用基于预训练模型的命名实体识别方法,由 BERT 模型、BiLSTM 模型和 CRF 模型三个部分组成。模型的结构如图 3 所示。

Bert 模型能够捕捉到丰富的语义和上下文信息,从字、词、位置等多个维度综合考虑文本特征,将文本转换为向量表示。随后以生成的向量为输入利用 BiLSTM 模型双向学习文本完成特征提取。

虽然 BiLSTM 模型能够直接进行标签预测,但预测时忽略了相邻标签之间的约束关系,会产生不符合标注规则的标签序列。为了解决这个问题,引入 CRF 模型在训练过程中增加标签的约束条件,通过联合概率计算出所有可能的标签序列,并选择具有最大联合概率的标签序列作为最终的预测结果。CRF 模型中文本序列标签得分计算公式如下所示:

$$Score(X,Y)=\sum_{i=1}^n P_{i,y_i}+\sum_{i=1}^n A_{y_i,y_{i+1}} \tag{1}$$

其中, P_{i,y_i} 表示第 i 个字符的标签是 y_i 的概率, $A_{y_i,y_{i+1}}$ 表示第 y_i 个标签转移到第 y_{i+1} 个标签的

得分。

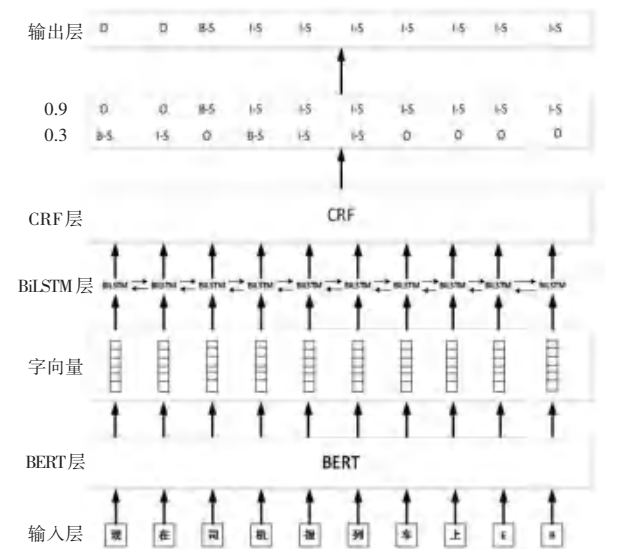


图 3 BERT-BiLSTM-CRF 模型

Fig. 3 BERT-BiLSTM-CRF model

本实验在 Windows10 系统上进行,使用 Python3.9,GPU 版本为 NVIDIA GeForce RTX 1050,在 Pytorch 框架下搭建命名实体识别模型,模型中最大长度序列为 120, $batch_size = 50$, $Epochs = 30$, $lr = 0.0002$ 。

使用调和平均数 ($F1$)、召回率($Recall$)及精确率($Precision$)为命名实体识别任务的评价指标。指标计算公式具体如下:

$$Recall = T_p / (T_p + F_n) \tag{2}$$

$$Precision = T_p / (T_p + F_p) \tag{3}$$

$$F1 = 2(P \times R) / (P + R) \tag{4}$$

其中, T_p 表示模型正确识别的实体个数; F_p 表示模型错误识别的实体个数; F_n 表示模型未能正确识别的实体个数。

验证集中不同指标随着训练轮次增加的变化情况如图 4 所示。

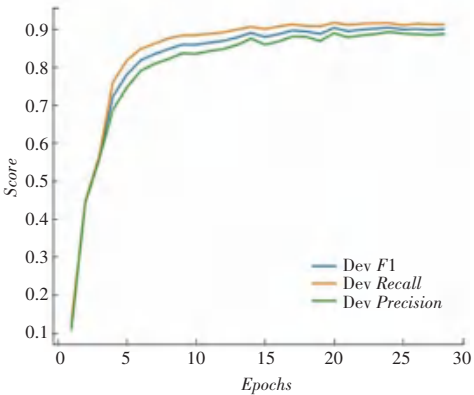


图 4 训练指标变化趋势

Fig. 4 Trend of training indicators

由图 4 可知,前期模型指标得分上升趋势较为明显,随着训练轮次的增加,后期模型指标趋于稳定。

为了更好地验证模型效果,本文选择 BERT-LSTM-CRF 及 BERT-BiLSTM 为参照模型,对比试验的结果见表 4。

表 4 模型对比试验结果
Table 4 Model comparison test results

模型	F1	召回率	精确率
BERT-BiLSTM	0.88	0.91	0.86
BERT-LSTM-CRF	0.89	0.91	0.87
BERT-BiLSTM-CRF	0.91	0.92	0.91

根据结果可知基于 BERT-BiLSTM-CRF 的方法具有一定的优势,结合模型测试集中评价指标结果分别为: $F1_score = 0.88$, $Recall_score = 0.90$,

$Precision_score = 0.86$, 本文训练出的实体识别模型基本满足调度应急处置文本实体提取的需求。

3.4 调度日志文本关系抽取

实体关系抽取是构建知识图谱的另一重要环节。常见的关系抽取技术主要包括基于规则、基于机器学习及基于神经网络的方法。

调度日志中的记录文本是对整个故障事件的描述,文本的组成具有一定的结构性,实体类别以及不同实体间的关系都是较为固定的。因此本文选择基于规则的方法实现文本中实体关系的提取。

以不同故障处置案例中提取出的实体为输入进行关系匹配。首先,在输入的案例实体集中确定第一个故障现象实体的位置,并以其为中心检索其余实体。本文定义的关系抽取模板见表 5。

表 5 调度应急处置案例关系抽取模板
Table 5 Dispatch emergency disposal case relationship extraction template

步骤(Step)	输出
Step1:检索处置案例中第一个故障现象实体出现的位置	故障现象实体
Step2:检索实体集中的时间、运行方向、车次号、设备类型、设备所属单位实体	作为第一个故障现象实体的实体属性进行存储
Step3:检索处置实体	建立关系 {故障现象 处置方法 处置}
Step4:检索处置实体后的故障现象实体	建立关系 {处置 处置结果 故障现象}
Step5:检索事件影响实体	建立关系 {故障现象 事件影响 事件影响}
Step6:检索事件闭环实体	建立关系 {故障现象 闭环情况 事件闭环}

注:连续相同类别实体之间的关系为并列,事件影响及事件闭环实体仅与第一个故障现象建立关系。

4 运营调度应急处置知识图谱应用分析

本文共提取出 8 752 个实体,建立了 7 882 个实体关系。所构建的知识图谱借助 Neo4j 图数据库以事件链的形式展示故障事件的演变规律,同时提供信息检索功能。

4.1 基于知识图谱的故障处置事件分析

借助 Neo4j 中的 Cypher 语句能够实现故障处置案例相关信息的检索。从设备类型、故障设备所属类别、故障时间、运行方向四个维度对故障处置事件的分布情况进行分析。以故障所属类别为例,查询结果如图 5 所示。

常见故障设备所属类别包括通告、车辆、工务、供电、客运五类,查询结果以节点的形式展示,每个节点代表一件故障处置实例,节点越多说明该类别下的故障事件越多,就应受到更多重视。通过该种形式能够使调度员直观地感受到故障事件的分布情况,加强对故障事件分布规律的认识。



图 5 故障处置事件分布情况

Fig. 5 Distribution of fault handling events

4.2 故障处置事件可视化展示

当发生故障事件时,调度员可依据故障现象查找相似案例。知识图谱中通过构建事件闭环实体来展示故障处置事件中最终的处置方法,包括故障原因分析、故障设备维修方法、故障场景中需要检修的设备等信息。调度员可通过查找不同故障事件的原因分析,并以此为依据有效提升处置效率。不同故

障现象的事件闭环实体查询结果如图 6 所示。



图 6 事件闭环实体查询

Fig. 6 Event closed-loop entity query

针对完整的故障处置事件知识图谱通过有向箭头建立不同实体间的关系,以链的形式展示事件的演变情况。

以图 7 中的列车区间运行 ATO 迫停处置事件为例,图数据库中以不同颜色的节点代表不同类别的实体。通过事件链的可视化展示能够加强调度员对于事件接报-处置决策-反馈-处置决策的故障处置流程的理解。同一事件中也更利于调度员查找不同故障现象所对应的处置方法,达到提升处置效率的目的。



图 7 故障处置实例

Fig. 7 Fault disposal example

5 结束语

本文研究了城市轨道交通调度员应急处置数据知识图谱的构建技术,描述了知识图谱中本体构建、本体关系构建及知识抽取的方法。构建的应急处置知识图谱提供了信息检索功能,通过完整的事件链展示了故障案例的演变规律。为提升适用性将进一步扩充处置数据完成知识图谱补全,并依据生产场景需求开发基于知识图谱的调度员辅助决策系统。

参考文献

[1] 刘开国. 重庆城市轨道交通调度指挥系统的网络化模式[J]. 城市轨道交通研究,2011,14(2):109-112.

[2] 刘志钢,胡华,黄远春,等. 城市轨道交通应急调度指挥系统的现状及发展趋势[J]. 城市轨道交通研究,2012,15(12):5-9.

[3] 张志博. 高速铁路应急调度处置流程重构与协同决策方法研究[D]. 成都:西南交通大学,2021.

[4] FENSEL D, ŞİMŞEK U, ANGELE K, et al. Introduction: what is a knowledge graph? [M]//Knowledge Graphs. Cham : Springer, 2020: 1-10.

[5] 李飞. 基于知识图谱的问答系统研究与实现[D]. 南京:南京邮电大学,2022.

[6] 郭恒,黎荣,张海柱,等. 多域融合的高速列车维修性设计知识图谱构建[J]. 中国机械工程,2022,33(24):3015-3023.

[7] 姚剑,白伟,杨国元,等. 铁路客站设备健康管理知识图谱构建与应用研究[J]. 铁道运输与经济,2023,45(3):94-102.

[8] 许慧,李树秀,邢镔. 基于知识图谱的轨道交通运营风险事件智能分析研究[J]. 铁道标准设计,2024,68(8):34-42.

[9] 丁君军,郑彦宁,化柏林. 基于规则的学术概念属性抽取[J]. 情报理论与实践,2011,34(12):10-14.

[10] 黄贺瑄,王晓燕,顾正位,等. 医学知识图谱构建技术及发展现状研究[J]. 计算机工程与应用,2023,59(13):33-48.

[11] 曾军,周国富. 基于机器学习的多语言文本抽取系统实现[J]. 计算机应用与软件,2017,34(4):87-92.

[12] 刘东,张越,皮俊波,等. 面向电网调控信息智能检索的知识图谱构建及应用[J]. 中国电力,2023,56(7):78-84.

[13] DEVLIN J, CHANG Mingwei, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv, 1810.04805, 2018.

[14] 赵鹏飞,赵春江,吴华瑞,等. 基于 BERT 的多特征融合农业命名实体识别[J]. 农业工程学报,2022,38(3):112-118.