



计算机科学

COMPUTER SCIENCE

基于动态图表示的设备故障推理预测方法

张慧, 张骁雄, 丁鲲, 刘姗姗

引用本文

张慧, 张骁雄, 丁鲲, 刘姗姗. 基于动态图表示的设备故障推理预测方法[J]. 计算机科学, 2024, 51(7): 310-318.

ZHANG Hui, ZHANG Xiaoxiong, DING Kun, LIU Shanshan. [Device Fault Inference and Prediction Method Based on Dynamic Graph Representation](#) [J]. Computer Science, 2024, 51(7): 310-318.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[融合多图卷积与层级池化的文本分类模型](#)

Text Classification Method Based on Multi Graph Convolution and Hierarchical Pooling
计算机科学, 2024, 51(7): 303-309. <https://doi.org/10.11896/jsjcx.230400164>

[融合多源图特征的Kcore-GCN反欺诈算法研究](#)

Study on Kcore-GCN Anti-fraud Algorithm Fusing Multi-source Graph Features
计算机科学, 2024, 51(6A): 230600040-7. <https://doi.org/10.11896/jsjcx.230600040>

[基于自适应上下文匹配网络的小样本知识图谱补全](#)

Adaptive Context Matching Network for Few-shot Knowledge Graph Completion
计算机科学, 2024, 51(5): 223-231. <https://doi.org/10.11896/jsjcx.230200012>

[基于多视图自编码器的多被试者脑影像功能校准](#)

Multi-view Autoencoder-based Functional Alignment of Multi-subject fMRI
计算机科学, 2024, 51(3): 141-146. <https://doi.org/10.11896/jsjcx.230600166>

[低资源场景事件抽取研究综述](#)

Survey of Event Extraction in Low-resource Scenarios
计算机科学, 2024, 51(2): 217-237. <https://doi.org/10.11896/jsjcx.221200142>

基于动态图表示的设备故障推理预测方法

张慧 张骁雄 丁鲲 刘姗姗

1 国防科技大学第六十三研究所 南京 210007

2 国防科技大学大数据与决策实验室 长沙 410073

(zhanghui82@nudt.edu.cn)

摘要 高效的设备运维可以保障设备的正常运行。然而,随着设备复杂程度越来越高,设备的维护和故障排查的复杂度和难度也不断增加。因此,人工方式越来越不能满足智能化设备的运维需要。智能运维将人工智能等新兴技术运用于运维过程,可以作为设备运维的有力支撑。但现有的很多方法依旧存在着未考虑动态性等不足。针对上述问题,提出了一种基于动态知识图谱表示学习的设备故障推理预测方法,用于预测目标设备是否与故障设备存在潜在的关联。该方法结合动态知识图谱表示学习和图表示推理模型,可以利用实时数据更新图网络,并运用图表示推理模型对新的故障数据进行推理。首先,使用动态知识图谱来表示设备运维数据,记录设备随时间的演化过程,从而有效地表达设备之间关系的动态变化性;然后,通过表示学习获得动态知识图谱中源故障设备和目标设备的时间感知表示;最后,将时间感知表示作为输入进行故障推理预测,判断设备之间是否存在潜在的关联。预测结果可以辅助运维人员解决相应的设备故障问题。在多个数据集上进行了实验,验证了所提方法的有效性。

关键词: 动态知识图谱;表示学习;链接推理预测;时间感知;设备运维

中图分类号 TP391

Device Fault Inference and Prediction Method Based on Dynamic Graph Representation

ZHANG Hui,ZHANG Xiaoxiong,DING Kun and LIU Shanshan

1 The Sixty-Third Research Institute,National University of Defense Technology,Nanjing 210007,China

2 Laboratory for Big Data and Decision,National University of Defense Technology,Changsha 410073,China

Abstract Effective equipment operation and maintenance is able to ensure the proper operation of equipment. Nevertheless,as the equipment becomes more and more sophisticated,the complexity and difficulty of maintaining and troubleshooting these devices are constantly increasing. As a result,equipment operation and maintenance mode that only rely on manual efforts is gradually unable to meet the requirements of intelligent equipment. Intelligent operation and maintenance that applies many new emerging technologies such as artificial intelligence to process of operation and maintenance can be used as a strong support for equipment operation and maintenance task. However,many existing methods still have deficiencies such as lack of considering dynamic characteristics. In order to solve these problems,a device fault inference and prediction method that is based on dynamic knowledge graph representation learning is proposed. The method can predict whether a target device is potentially associated with a faulty device time during the operation and maintenance process. The proposed method combines dynamic knowledge graph representation learning with graph representation inference models,updates the graph network based on real-time data,and employs graph representation inference models to infer new fault data. Firstly,it takes advantage of a dynamic knowledge graph to represent the equipment operation and maintenance data,so as to records the evolution of the equipment over time. The representation effectively denote dynamic changes of the relationship between the devices. Next,the time-aware representations of the source faulty equipment and the target equipment in the dynamic knowledge graph are obtained through representation learning. Finally,the time-aware representations are used as inputs for fault inference prediction,which predicts whether there exists any potential correlation between the equipment so as to assist the operation and maintenance engineer in solving the corresponding equipment fault problems. Experiments on multiple datasets verify the effect of the proposed method.

Keywords Dynamic knowledge graph,Representation learning,Link inference prediction,Time awareness,Device operations and maintenance

到稿日期:2023-10-31 返修日期:2024-03-02

基金项目:**** 能力 ** 服务提升(24220407)

This work was supported by the Project of **** Capacity ** Service Improvement(24220407).

通信作者:丁鲲(dingkun18@nudt.edu.cn)

1 引言

智能化设备在各领域的应用越来越广泛,如各类通信设备、运输设备等。这些设备的维护以及故障排查的复杂度和难度也不断增加^[1]。人工故障排查和维护需要耗费大量时间和人力^[2];而且随着设备复杂性越来越高,单纯的人工方式已经很难适应智能化设备维修的需要。近年来,随着人工智能方法和技术不断被运用于设备运维领域,智能设备故障推理预测技术开始出现^[3-7],其不仅可以对故障部位进行定位,对产生故障的原因进行分析,还可以评估设备出现问题的风险,预测潜在的故障。在实际的应用中,有效的故障推理预测技术可以帮助运维人员提前发现故障并采取相应的维修措施,减少停工时间和维修成本,提高运维效率和质量,降低发生更严重故障的可能性^[8]。

目前,许多设备故障推理预测技术基于图智能^[9]的方法。图方法以图的形式表示运维数据,能够有效地捕捉设备之间复杂的关系^[10],可以帮助运维团队更好地理解这些关联,以便采取适当的措施来应对潜在问题。但这类方法也存在着一一些问题,主要表现为未考虑时间因素^[11]。而事实上,由于设备的运行状态和性能通常会随着时间而发生变化,且设备故障通常与先前的事件或条件有关,这些关系很可能需要在时间轴上得到体现。例如,某些故障可能会在设备长时间运行后才逐渐显现。如果在模型设计时未能充分考虑时间因素,则很可能会导致无法准确地对故障进行推理预测。在推理预测过程中引入时间因素,则可以更好地理解设备故障的演变过程,包括可能的先兆和趋势。有些研究人员意识到运维数据的时序性是重要的特征,将时间作为属性纳入到建模过程^[12-13]。如 Zhou 等认为运维数据具有时间依赖的特点,将其作为重要的特征进行建模^[12];Qi 等设计了门控卷积网络,从设备使用情况的时序数据中提取时间维度特征来检测设备的异常^[13]。但是这些方法还普遍存在一些问题,主要表现为:首先,纳入时间属性后,存在着模型的训练和推理计算成本过大的问题;其次,针对设备使用过程中的状态变化对设备拓扑图带来影响的适应能力较差。例如,局域网中某个路由器出现了故障或新增了路由器,对应的设备图拓扑结构会发生变化。传统的方法难以处理这类问题。

综上所述,传统的静态方法很难适应设备的动态运行和表达设备之间随时间变化的关联性,因此需要设计一种新的方法。动态图是一种动态计算模型,其计算图的结构可以在运行时动态地改变。动态图模型所具备的可以随着时间的推移而演变的能力,使之能捕捉设备在运行时的状态变化。动态知识图谱(Dynamic Knowledge Graph)在知识图谱的基础上引入了时间维度^[14],它可以记录和表示知识的演化过程,为各种应用场景提供了更强大的知识表示和分析能力。但是在利用动态知识图谱进行推理预测时,如何有效地进行表示学习是亟需解决的关键问题^[15]。

为解决传统设备故障推理预测技术不能较好地处理动态信息的问题,本文提出了一种基于动态知识图谱表示学习的设备故障推理预测方法。该方法主要包括动态知识图谱表示学习和故障推理预测两个部分,其首先采用动态知识图谱来

表示设备运维数据,然后在动态知识图谱表示学习部分,通过提取源设备和目标设备的一阶节点来进行交互并对其进行编码和分片,再将其输入到 Transformer 中学习源设备和目标设备的时间感知表示,最后将时间感知表示和所需故障实体向量、关系向量和三元组向量输入到推理预测部分,不断地进行迭代学习,来预测源设备和目标设备是否存在潜在的关联,从而辅助运维人员解决相应的设备故障问题。

综上所述,本文的主要工作包括:

- 1)使用动态知识图谱来表示设备运维数据,有效考虑了设备之间关系的动态变化性;
- 2)提出了一种图表示学习方法,能够从动态知识图谱中提取知识实体的时间感知表示;
- 3)设计了一种故障推理预测方法,可以准确、及时地预测设备之间的潜在关联性;
- 4)实验验证了本文方法能较好地运用于设备运维任务。

本文第 2 章介绍了与本文相关的一些现有工作;第 3 章详细描述了所提方法,包括方法的总体框架、动态知识图谱表示学习和设备故障推理预测;第 4 章介绍了实验所用的运维数据集和模型评价指标;第 5 章给出了实验结果;最后对全文进行了总结。

2 相关工作

2.1 动态知识图谱表示学习

动态知识图谱由于其动态性,相比静态的知识图谱,在表示学习上更具挑战性。Seo 等^[16]提出的图卷积递归网络(Graph Convolutional Recurrent Networks,GCRN)可有效地进行动态知识图谱表示学习。GCRN 使用切比雪夫多项式图卷积^[17]处理图空间信息,使用长短期记忆网络(Long Short Term Memory network,LSTM)处理时间信息。Yu 等^[18]提出了时空图卷积网络(Spatio-Temporal Graph Convolutional Networks,STGCN),空间层卷积使用的是切比雪夫多项式近似的卷积,而时间维卷积使用的是门函数卷积^[19]操作,在交通预测中取得了很好的效果。但是,总体而言,上述方法没有考虑到隐藏的节点管理或者链接缺失的情况。针对此,Wu 等^[20]提出了一种用于深度时空建模的 Graph Wavenet,其通过学习一个自适应邻接矩阵来反映隐藏的空间依赖,同时使用图卷积神经网络(Graph Convolutional Network,GCN)和扩张因果卷积(Dilated Casual Convolution)^[21]来获取长时空相关性;Zheng 等^[22]提出了一种用于交通预测的图多注意力网络(Graph Multi-attention Network,GMAN),其空间维度和时间维度全部采用注意力模型。

2.2 链接推理预测

链接推理预测的一个研究难点是如何在变化的网络中进行有效的预测。Lei 等^[23]引入了一种新型非线性模型 GCN-GAN(Generating Adversarial Network),用于解决加权动态网络中具有挑战性的时态链路预测任务。所提模型充分利用了 GCN,LSTM 以及 GAN 的优势。De 等^[24]进行了有监督的时间链接预测的研究,使用一种新方法研究离散事件建模网络和持久关系建模网络之间的区别。该方法系统地利用了过去所有事件的信息,从而大大提高了链接预测性能。Lv

等^[25]提出了一种新颖的图正则化非负矩阵因式分解算法,利用 PageRank 中心性和非对称链接聚类系数来解决时空链接预测问题。

3 基于动态知识图谱表示学习的预测方法

本章将详细介绍所提的基于动态知识图谱表示学习的设备故障推理预测方法,包括所提方法的整体架构,以及动态图网络表示学习和故障推理预测的具体细节。

3.1 整体架构

本文面向设备故障预测,使用动态知识图谱表示设备运维数据,其中节点(实体)表示设备,它们直接的链接表示两设备之间随时间在机械和电子层面上构建的联系。基于此,本文提出了一种基于动态知识图谱表示学习的设备故障推理预测方法。

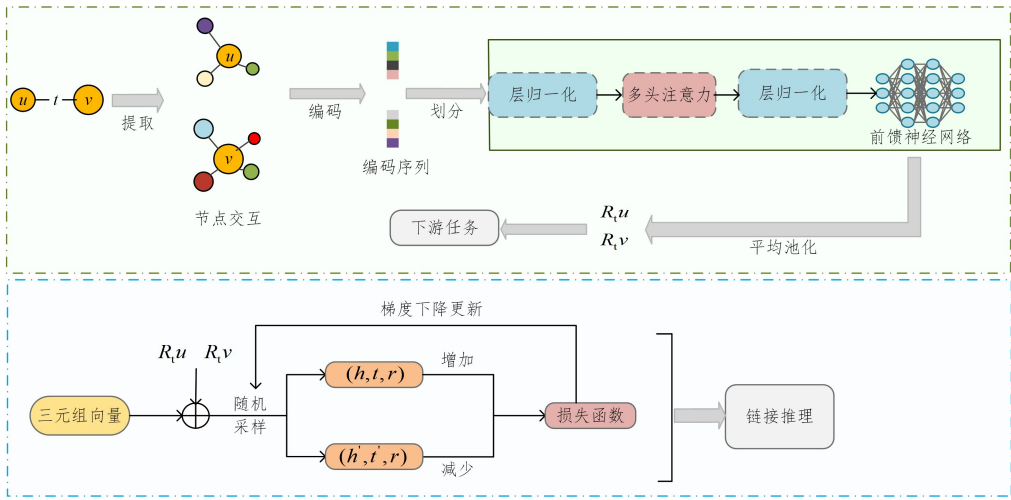


图1 模型架构图

Fig. 1 Architecture diagram of the proposed model

3.2 动态知识图谱表示学习

3.2.1 一阶交互序列提取

现在很多方法以多阶的方式提取源节点 u 和目标节点 v 的交互序列,这会极大地提高模型的计算复杂度。本文方法只从节点的一阶交互序列中学习,将复杂的动态图表示学习任务变成了一个简单的序列学习问题^[27]。在数学上,给定一个与时间相关的源节点 u 和目标节点 v 相互作用 (u, v, t) ,时间戳 t 之前的源节点 u 和目标节点 v 的交互序列定义如下:

$$S_u^t = \{(u, u', t' | t' < t)\} \cup \{(u', u, t' | t' < t)\} \quad (1)$$

$$S_v^t = \{(v, v', t' | t' < t)\} \cup \{(v', v, t' | t' < t)\} \quad (2)$$

3.2.2 交互序列编码

在得到两个交互序列 S_u^t 和 S_v^t 后,需要计算每个序列的节点编码、链接编码、时间间隔编码和邻居共现编码。动态图通常与节点和链接的特征相关联^[28]。因此,对于源节点 u 和目标节点 v 的节点编码和链接编码,只需根据给定的特征,检索序列 S_u^t 和 S_v^t 中涉及的邻居和链路的特征来表示它们的编码,分别表示为 $X_{\#,N}^t \in \mathbb{R}^{|S_u^t| \times d_N}$ 和 $X_{\#,E}^t \in \mathbb{R}^{|S_v^t| \times d_E}$ 。然后,通过下述函数对时间间隔 $\Delta t = t - t'$ 进行编码来学习周期时间模式^[29]。

本文方法的整体架构如图1所示,主要包括动态知识图谱表示学习和故障推理预测两个部分,目标是在源设备 u 发生故障时,推测目标设备 v 与源设备 u 是否有关联。动态知识图谱表示学习部分主要包含4个阶段,分别是提取阶段、编码阶段、划分阶段和表示学习阶段。在提取阶段,从历史相互作用序列 (u, v, t) 中提取源设备 u 和目标设备 v 的一阶交互序列 S_u^t 和 S_v^t ;在编码阶段,对一阶交互序列进行编码,得到编码序列;在划分阶段,将编码序列划分成多个小块输入到设计的 Transformer 模型^[26]中;在表示学习阶段,利用 Transformer 进行表示学习,并对其输出进行平均池化,获得时间感知表示 R_u^t 和 R_v^t 。故障推理预测部分通过输入3组向量和时间感知表示 R_u^t 和 R_v^t ,来不断进行迭代学习,以预测在源设备 u 发生故障时,目标设备 v 与源设备 u 是否有潜在联系,从而推理目标设备 v 是否可能发生故障。

$$\sqrt{\frac{1}{d_T}} [\cos(\omega_1 \Delta t'), \sin(\omega_1 \Delta t'), \dots, \sin(\omega_{d_T} \Delta t')] \quad (3)$$

其中, $\omega_1, \dots, \omega_{d_T}$ 是可训练的参数, d_T 是时间间隔编码的维度。因此,时间间隔编码定义为 $X_{\#,T}^t \in \mathbb{R}^{|S_{\#}^t| \times d_T}$ 。

但是,上述编码将节点、链接和时间间隔分别进行编码,没有考虑它们之间的相关性,因此本文使用了一个邻居共现编码方案来解决这个问题。该方法的主要假设是,一个邻居在序列中出现的频率表明了它的重要性,而一个邻居在 u 和 v 序列中的出现(即共现)可以反映 u 和 v 之间的相关性。即,如果 u 和 v 在其序列中具有更多共同的共现情况,则它们在未来更有可能发生相互作用。形式上,给定交互序列 S_u^t 和 S_v^t ,我们将计算其中邻居出现的次数,并导出一个二维向量;再将所有的邻居向量整合在一起,可以得到 u 和 v 的邻居共现特征,表示为 $C_u^t \in \mathbb{R}^{|S_u^t| \times 2}$ 和 $C_v^t \in \mathbb{R}^{|S_v^t| \times 2}$ 。于是,邻居共现编码定义如下:

$$\mathbf{X}_{\#,O}^t = f(\mathbf{C}_{\#}^t[:, 0]) + f(\mathbf{C}_{\#}^t[:, 1]) \in \mathbb{R}^{|S_{\#}^t| \times d_O} \quad (4)$$

其中,函数 f 的输入和输出维度分别为1和 d_O ,是通过带有 ReLU^[30] 的双层感知器来实现的; $\mathbf{C}_{\#}^t$,表示计算交互序列 S_u^t 和 S_v^t 每个邻居出现次数得到的二维向量; $\#$ 表示源节点 u 或目标节点 v 。

3.2.3 修补技术

在进行编码后,本文利用修补技术将编码序列划分为多个不重叠的小块,可有效地解决传统方法无法有效捕获节点之间的长期依赖性的问题。 P 定义为块的大小,因此每个块由 P 个时间上相邻的相互作用组成,具有平坦的编码,并且可以保持局部时间上的邻近性。例如,节点编码 $X_{u,N}^t \in \mathbb{R}^{|S_u^t| \times d_N}$ 总共可被划分成 $l_u^t = |S_u^t|/P$ 个块,定义为 $M_{u,N}^t \in \mathbb{R}^{l_u^t \times d_N \cdot P}$ 。其他编码序列处理方式相同。

3.2.4 表示学习

在进行节点的表示学习之前,需要使用可训练的权值 $W_* \in \mathbb{R}^{d_* \cdot P \times d}$ 和 $b_* \in \mathbb{R}^d$ 将修补的编码对齐到相同的维度 d 上,得到 $Z_{u,*}^t \in \mathbb{R}^{l_u^t \times d}$ 和 $Z_{v,*}^t \in \mathbb{R}^{l_v^t \times d}$ 。详细定义如下:

$$Z_{u,*}^t = M_{u,*}^t W_* + b_* \in \mathbb{R}^{l_u^t \times d} \quad (5)$$

$$Z_{v,*}^t = M_{v,*}^t W_* + b_* \in \mathbb{R}^{l_v^t \times d} \quad (6)$$

其中, $*$ 可代表节点 N 、链接 E 、时间间隔 T 或邻居共现 O 。

然后,将有关 u 和 v 的 4 种编码的对齐编码连接起来,得到 Z_u^t 和 Z_v^t ,其具体定义如下:

$$Z_u^t = [Z_{u,N}^t; Z_{u,E}^t; Z_{u,T}^t; Z_{u,O}^t] \in \mathbb{R}^{l_u^t \times 4d} \quad (7)$$

$$Z_v^t = [Z_{v,N}^t; Z_{v,E}^t; Z_{v,T}^t; Z_{v,O}^t] \in \mathbb{R}^{l_v^t \times 4d} \quad (8)$$

接着,使用 $Z' = [Z_u^t; Z_v^t]$ 作为 Transformer 的输入来捕获节点之间的时间依赖性。Transformer 编码器由多头自注意力 (Multi-head Self-Attention, MSA)^[31] 和前馈网络 (Feed-Forward Network, FFN) 块^[32] 堆叠而成,每块之前使用层归一化 (Layer Normalization, LN)^[33],每块之后使用残差连接。经 Transformer 处理后,时间感知表示 R_u^t 和 R_v^t 定义为:

$$R_u^t = \text{MEAN}(H^t[; l_u^t; :])W_{\text{out}} + b_{\text{out}} \in \mathbb{R}^{d_{\text{out}}} \quad (9)$$

$$R_v^t = \text{MEAN}(H^t[l_v^t; l_v^t; :])W_{\text{out}} + b_{\text{out}} \in \mathbb{R}^{d_{\text{out}}} \quad (10)$$

其中, H^t 为 Transformer 的输出, $W_{\text{out}} \in \mathbb{R}^{4d \times d_{\text{out}}}$ 和 $b_{\text{out}} \in \mathbb{R}^{d_{\text{out}}}$ 为可训练的权值, d_{out} 为输出的维度, MEAN 为求平均数操作。

3.3 故障推理预测

故障推理预测的主要任务是将学习到的时间感知表示 R_u^t 和 R_v^t 作为输入来推理预测目标实体 u 与源故障实体 v 是否存在关联。此外,故障推理预测还需要用到 3 个向量,包括故障实体向量(实体名, id)、故障关系向量(关系名, id)和故障三元组向量(h, t, r)。训练前需要对实体和关系各列表进行初始化,得到两个向量字典。初始化过程中使用了 L2 范数归一化,每训练一个 epoch 后,都对实体和关系进行一次 L1 范数归一化。 $h+r$ 与 t 之间的距离是 $d(h+r, t) = L(h+r-t)$ 。

模型训练时,在实体字典中随机不重复抽样得到 batch-size 个正确三元组 $S(h, t, r)$ 和错误三元组 $S'(h', t', r)$, S 和 S' 合并得到元组 $T_{\text{batch}}((h, t, r)(h', t', r))$, 模型训练的损失函数为:

$$\text{Loss} = \sum_{((h, r, t), (h', r', t')) \in T_{\text{batch}}} \nabla[\gamma + d(h+r, t) - d(h'+r, t')]_+ \quad (11)$$

如果正例 d 特别小, 负例 d 特别大, 损失几乎为 0 或小于 0, 则取 0 不叠加; 否则, 计算每个 batch 的损失并做叠加。其中, γ 代表正负三元组最大的距离。在损失大于 0 时, 采用

随机梯度下降算法 (Stochastic Gradient Descent, SGD) 更新参数, 需要对 Loss 采用 L1 范数求导, 先对 L2 范数求导, 逐元素判断其值为正或为负, 如果值为正, 则赋值为 1, 如果值为负, 则赋值为 -1。每一次 SGD 更新的参数就是一个 batch 中所有实体和关系的 Embedding 值。更新参数 h' 或 t' , 重叠的实体或关系需要更新两次, 针对正确的三元组更新一次, 针对错误的三元组更新一次。训练完成后, 即可进行设备故障的推理预测。本文方法的描述如算法 1 所示。

算法 1 基于动态图表示的设备故障推理预测

输入: 数据集中时间交互三元组序列 $S\{(u, v, t)\}$, 关系三元组 $G\{(h, t, r)\}$

输出: 训练好的设备故障推理预测模型

1. 初始化: 训练轮次 epoch、批次大小 batchsize、块大小 P、关系总数 $|R|$ 、关系集 R 、链接总数 $|E|$ 、链接集合 E 、序列 $S_{\#}^t$ 的邻居出现次数的二维向量 $C_{\#}^t$ ($\#$ 代表 u 或 v);
2. for $ep \leftarrow 1$ to epoch do
3. for $b \leftarrow 1$ to batchsize do
4. $l^t \leftarrow |S^t|/P$
5. $M_{u,*}^t, M_{v,*}^t \leftarrow \text{patching}(X_{u,*}^t, X_{v,*}^t), l^t$
6. $h_*^t \leftarrow \text{Transformer}(M_{u,*}^t, M_{v,*}^t)$
7. end for
8. for each $r \in R$
9. $r \leftarrow \text{uniform}\left(-\frac{6}{\sqrt{k}}, \frac{6}{\sqrt{k}}\right)$
10. end for
11. for each $r \in R$
12. $r \leftarrow \frac{r}{\|r\|}$
13. end for
14. for each $e \in E$
15. $e \leftarrow \text{uniform}\left(-\frac{6}{\sqrt{k}}, \frac{6}{\sqrt{k}}\right)$
16. end for
17. for each $e \in E$
18. $e \leftarrow \frac{e}{\|e\|}$
19. end for
20. $e \leftarrow e \cup h_u^t \cup h_v^t$
21. for $b \leftarrow 1$ to batchsize do
22. $S_{\text{batch}} \leftarrow \text{sample}(G, b)$
23. $T_{\text{batch}} \leftarrow \emptyset$
24. for $(h, t, r) \in S_{\text{batch}}$ do
25. $(h', t', r) \leftarrow \text{sample } S'_{(h,t,r)}$
26. $T_{\text{batch}} \leftarrow T_{\text{batch}} \cup \{(h, t, r), (h', t', r)\}$
27. end for
28. 更新模型
29. end for
30. end for
31. return 训练好的设备故障推理预测模型

3.4 复杂度分析

在经典的基于图神经网络的表示学习方法中, 图神经网络的时间复杂度以节点数 ($|N|$) 和链接数 ($|E|$) 为基础, 时间复杂度表示为 $O(f(|N|, |E|, |L|))$, 其中 $|L|$ 是层数, f 是与消息传递和更新相关的函数; 而空间复杂度则以节点数和

每个节点表示的维度为基础,表示为 $O(|N| \times |D| \times |L|)$,其中 $|N|$ 是节点数, $|D|$ 是每个节点表示的维度, $|L|$ 是层数。

本文方法只需从节点的一阶交互序列中学习,使用基于 Transformer 的表示学习在处理图数据时具有更高的效率。其时间复杂度以序列长度和层数为基础,先对时间和邻居进行编码,时间复杂度分别为 $O(|S| \times |T|)$ 和 $O(|S| \log |S|)$;然后通过 Transformer 编码器学习得到节点的特征表示,时间复杂度为 $O(g(P, |S|) \times |L|)$,其中 $|S|$ 是序列长度, $|T|$ 是时间维度, P 为块大小, $|L|$ 是层数, g 是与自注意力机制相关的函数,对多头注意力的计算复杂度为 $O(P \times |S|)$,因此通过 Transformer 编码器学习得到节点特征表示的时间复杂度为 $O(P \times |S| \times |L|)$ 。综上可得,本文方法的时间复杂度为三者之和: $O(|S| \times |T|) + O(|S| \log |S|) + O(P \times |S| \times |L|)$ 。本文所提方法的空间复杂度以序列长度和每个位置的表示维度为基础,表示为 $O(|S| \times |D| \times |L|)$,其中 $|S|$ 是序列长度, $|D|$ 是每个节点表示的维度, $|L|$ 是层数。

4 实验设置

本章介绍了实验所使用的运维数据集,然后给出了模型的参数设置,最后介绍了模型度量标准。

4.1 数据集介绍

4.1.1 VtE-I 数据集

VtE-I 数据集是一个现实生活中的车辆设备间交互的动态连接网络。在这个网络中,节点表示组成车辆的设备,它们直接的链接表示两设备之间随时间以日志和运行逻辑的形式在机械和电子层面上构建的联系;同时,该数据集中还含有动态的故障标签,用于提醒用户设备可能面临的故障。

4.1.2 CDNA 数据集

CDNA 数据集是一个通信设备网络流量数据集。该通信图网络中的实体包含多个通信设备节点,每个节点代表一个不同类型的通信设备,如路由器、交换机、服务器等。边记录了设备之间的网络通信流量信息,包括数据包传输、带宽利用率和延迟等性能指标。该数据集是动态图结构数据集,允许分析通信流量和操作的时间动态性。

4.1.3 维基百科数据集

维基百科(Wikipedia)数据集包括一个双方交互图,记录了一个月内对维基百科页面所做的编辑。其中,节点代表用户和页面,它们之间的链接表示编辑操作,并附有时间戳,每个链接都与 172 维的语言调查和字数统计(LIWC)特征相关联^[34]。此外,该数据集还包含动态标签,用于指示用户面临临时编辑禁令的情况。

4.1.4 Can. Parl. 数据集

Can. Parl. 数据集是一个动态政治网络,记录了 2006—2019 年加拿大国会议员之间的互动。该网络中,节点代表不同选区的国会议员,当议员都对一项法案投“赞成”票时,议员之间就建立了联系。每个连接的强度由一名议员在特定年份内对另一名议员投“赞成”票的频率决定。

4.1.5 数据集统计分析

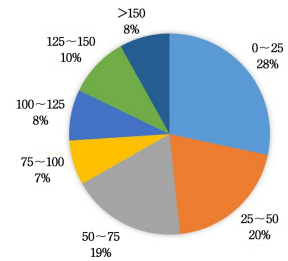
表 1 列出了在实验中所使用的 4 个数据集的统计分析信息。

表 1 数据集的相关统计分析信息

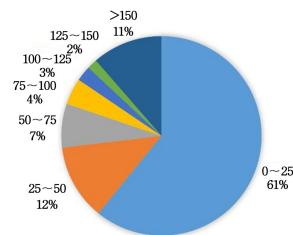
Table 1 Relevant statistics of datasets

数据集	节点	链接	时长
VtE-I	945	1 783	6 hours
CDNA	1 899	9 835	7 days
Wikipedia	9 227	157 474	1 month
Can. Parl.	734	74 478	14 years

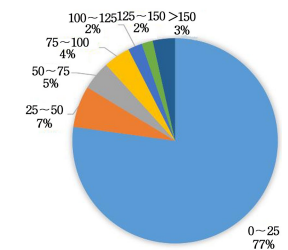
通过统计分析可以得出 4 个数据集中节点的度的分布情况,如图 2 所示。可以看出,节点的度分布主要包括 0 到 25 之间、25 到 50 之间、50 到 75 之间、75 到 100 之间、100 到 125 之间、125 到 150 之间以及大于 150 这 7 种情况。度数在 25 以上的节点数量在各个数据集中都占有一定比重,表明了实验所用数据集中链接的复杂性。



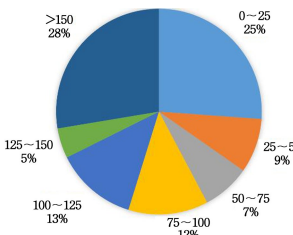
(a) VtE-I 数据集



(b) CDNA 数据集



(c) Wikipedia 数据集



(d) Can. Parl. 数据集

图 2 4 个数据集中节点的度的分布

Fig. 2 Distribution of degrees of nodes in four datasets

4.2 模型参数设置

本文对模型进行了 200 轮次的训练,采用 Adam 优化器进行优化,学习率为 0.0001。为避免模型在到达最佳性能之前出现过拟合,使用早期停止策略,设置耐心值为 20。当模型在训练过程中的性能在耐心值轮次(20 轮次)内没有显著

改善时,提前停止训练以避免过度拟合,从而提高模型的泛化能力。不同的数据集在训练时,输入的交互序列长度和块大小不一,在 VtE-I,CDNA 和 Wikipedia 数据集上分别为 32 和 1,在 Can. Parl. 数据集上分别为 2048 和 64。同时,在评估模型性能时,本文使用不同的随机种子对所提模型运行 10 次,计算模型的平均结果以消除偏差。具体的超参数设置如表 2 所列。

表 2 模型的超参数设置

Table 2 Model hyper-parameter settings

超参数	值
训练轮次 epoch	100
批次大小 batchsize	200
学习率 α	0.0001
块大小 P	1 或 64
交互序列长度	32 或 2048
耐心值	20
时间间隔编码维度 d_T	100
邻居共现编码维度 d_O	50
对齐编码维度 d	50
时间感知表示维度 d_{out}	172
多头自注意力数量	2
Transformer 层数量	2

4.3 模型度量标准

本文基于文献[35]中的方法,对模型在动态链接预测方面的性能进行了评估,使用多层感知器,将两个节点的串联表示作为输入,并输出源实体和目标实体之间存在链接的概率。同时,使用平均精度(Average Precision, AP)和接收者工作特性曲线下面积(Area Under the Receiver Operating Characteristic Curve, AUC-ROC)作为评估性能的指标。

在车辆设备图网络中,实体间的推理测试模块使用平均排名 MR(Mean Rank)和 HITS@10 两个指标来评价车辆故障的预测准确度。

MR 指预测排名中的正确预测所在的排名的平均值,计算思想是正确结果排名之和与总查询次数的比值,排名越靠前,rank 越小,求和也就越小,指标越小越好。MR 的计算式如下:

$$MR = \frac{1}{|S|} \sum_{i=1}^{|S|} rank_i \quad (12)$$

其中, $rank_i$ 是第 i 个三元组的链接预测排名; S 是三元组集合, $|S|$ 是三元组集合中元素的数量。

HITS@10^[36] 指预测排名中正确预测排在前 10 位占总查询次数的比例,计算式如下:

$$HITS@10 = \frac{1}{|S|} \sum_{i=1}^{|S|} I(rank_i \leq 10) \quad (13)$$

其中, I 是 indicator 函数,若条件真则函数值为 1,否则为 0。

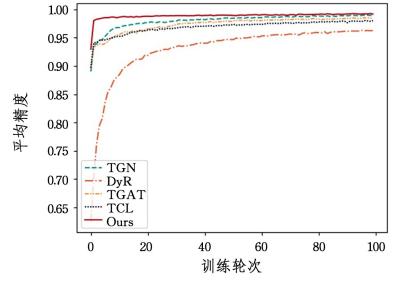
5 实验

首先,在 Wikipedia 和 Can. Parl. 数据集上进行了实验,测试链接预测任务下的模型性能;然后,以车辆设备运维数据集 VtE-I 为例,验证了所提方法在设备故障推理预测上的有效性;最后,在 CNDA 数据集上,通过预测通信设备之间是否存在潜在的关联来进一步验证模型的有效性。

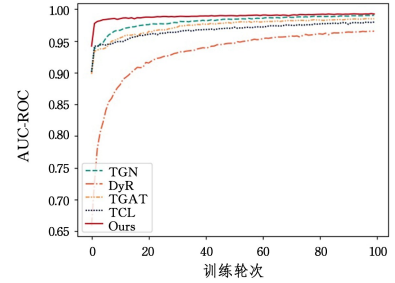
5.1 链接预测任务下的模型性能比较

本文共采用了 DyRep^[37], TGAT^[29], TGN^[38] 和 TCL^[39] 4 个模型与所提方法进行对比训练,链接预测结果的评价指标

采用 AUC 和 AP。每个模型运行 10 次,取其平均值作为最终绘制曲线的依据,如图 3 所示。



(a) 平均精度变化



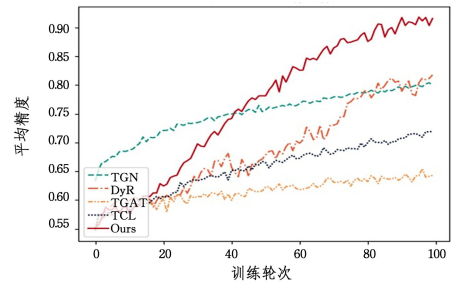
(b) AUC-ROC 变化

图 3 Wikipedia 数据集上不同模型的训练过程

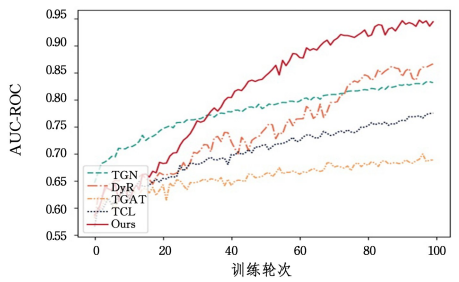
Fig. 3 Training process of different models on Wikipedia dataset

显而易见的是,相比另外 4 种模型,本文方法在 Wikipedia 数据集上训练的初始几轮,AP 和 AUC-ROC 值就达到了一个很高的水平。此外,值得注意的是,由于 DyRep 使用了一种双时间尺度的深度时间点过程模型(Two-time Scale Deep Temporal Point Process Model),该模型由一个时间注意力表征网络进一步参数化,该网络将时间上不断演变的结构信息编码到节点表征中,而节点表征反过来又驱动着动态图结构的非线性演化,因此模型需要更多的训练来适应图结构的动态变化,从而达到更好的表现。

在 Can. Parl. 数据集上的训练过程如图 4 所示。



(a) 平均精度变化



(b) AUC-ROC 曲线变化

图 4 Can. Parl. 数据集上不同模型的训练过程

Fig. 4 Training process of different models on Can. Parl. dataset

可以观察到 TGN 模型在训练轮次小于 35 时,模型的 AP 和 AUC-ROC 值在 5 个模型中表现得最好。但随着训练轮次的增加,本文模型和 DyRep 模型的性能逐渐超越 TGN 模型。同时,相比 DyRep 模型,所提方法实现了更为出色的性能。这是因为 Can. Parl. 数据集中记录的链接时长达到了 3 年,而本文方法使用修补技术,能够有效地利用更长的历史记录并捕获更加长期的时间依赖性。

综上所述,相比其他 4 种模型,所提模型在 Wikipedia 和 Can. Parl. 数据集上训练时都取得了最好的效果。但在实际应用的下游任务中,还必须认真考虑应用数据集的质量和规模,选择合适的参数配置,以确保模型的鲁棒性和泛化能力。

表 3 列出了不同方法最终在测试集上的 AP 度量。从最终的测试结果可以看出,本文方法优于其他方法。这是因为所提方法使用了一阶关联节点编码方案,通过分析源节点和目标节点在同一空间的共同关联节点,根据节点间的相关性来预测它们未来的关联概率;其次,所提方法还采用了修补技术,这使得它能够有效地利用更长的历史记录并捕获长期的时间依赖性。而尽管基于记忆网络的方法(即 DyRep 和 TGN)能够在可承受的计算成本范围内处理较长的历史信息,但由于梯度消失和爆炸的问题,它们依然无法从更长的历史中获益。

表 3 测试集上链路预测的 AP

Table 3 AP of link prediction on test sets

方法	Wikipedia	Can. Parl.
本文方法(Ours)	0.9804	0.9623
DyRep	0.8969	0.5689
TGAT	0.9569	0.6148
TGN	0.9562	0.5983
TCL	0.9615	0.6106

5.2 车辆设备运维数据的图网络表示学习

本节以真实收集的车辆设备运维数据集 VtE-I 为例,展现了所提模型在 VtE-I 数据集上的训练、验证以及测试过程,如图 5 所示。显而易见的是,验证集和训练集的 AUC-ROC 之间存在约 0.3 的差距。这是因为本文所使用的车辆设备网络运维数据集 VtE-I 规模相对较小,容易出现过拟合现象。但在最后对测试结果进行计算后,得到的 HITS@10 指标可以达到 0.8087,MR 指标达到了 2.1119,这表明所提方法在真实的设备故障推理中是有效的。

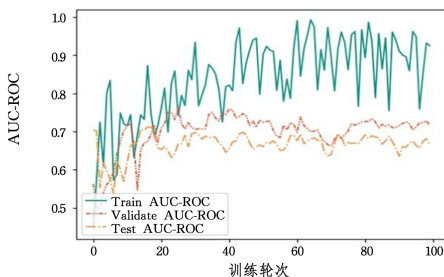


图 5 VtE-I 数据集上的训练、验证和测试过程

Fig. 5 Training, validation and testing process on VtE-I dataset

本文模型可以将车辆设备运维网络中的节点编码为有效的潜在表示,其中不同的节点在实际场景中代表着不同的设备。在动态运维过程中,通过模型学习可以得到源故障设备

和目标设备的时间感知表示。故障的推理预测在图结构层面对应链接预测,如果预测得到源故障设备与目标设备之间存在链接关系,则说明目标设备当前的亚健康状态或故障情况可能是由于源设备节点的故障所致,可考虑安排运维人员对源设备的状态进行检查。

图 6 给了一个案例分析。如图 6 所示,本文所提模型将设备节点编码为时间感知表示,当发现燃油泵出现故障时,基于学习到的潜在表示对燃油泵这一设备节点进行链接预测(即故障推理预测),发现喷油器设备节点与燃油泵设备节点之间可能存在故障链接关系,说明目标设备(燃油泵)当前的亚健康状态或故障情况可能是由于源设备节点(喷油器)的故障所致。

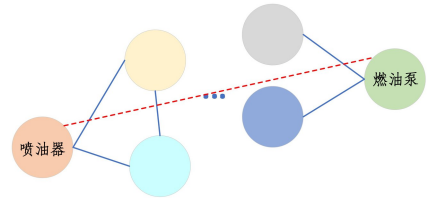
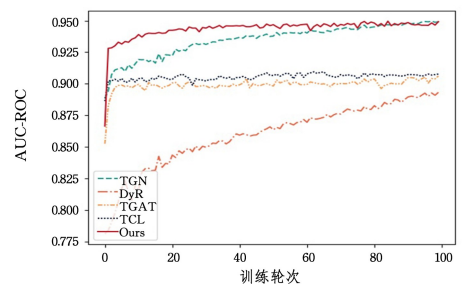


图 6 案例分析

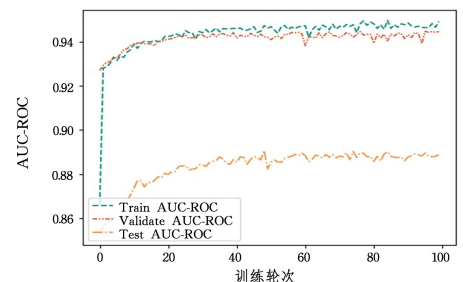
Fig. 6 Case study

5.3 基于通信流量网的设备间关系预测

本文在通信流量网数据集 CDNA 上进行了多个模型的对比实验,并对所提模型在训练、评估、测试过程的变化也进行了可视化,如图 7 所示。从图 7(a)可以看出,所提方法可以更有效地对动态知识图谱进行表示学习,这有助于应用于更多的通信设备和网络流量分析方面的下游任务中。从图 7(b)中可以发现,训练和验证阶段取得了较高的 ACU-ROC 值,但在测试阶段性能会出现较为明显的下降(但是 AUC-ROC 仍能达到 0.88 以上的较高水平),这说明模型可以较好地结合流量流动的细节,预测通信设备间是否可能存在相关的联系。



(a) AUC-ROC 变化



(b) 所提模型在 CDNA 数据集上的训练、验证和测试过程

图 7 CDNA 数据集上的实验结果

Fig. 7 Experimental results on CDNA dataset

结束语 本文设计了一个基于动态知识图谱表示学习的推理预测模型,其可用于设备间关联的推理预测。动态知识图谱表示学习和图表示推理模型的组合可以实现实时的风险检测,可以根据实时的数据更新图网络,并通过运用图表示推理模型对新的故障数据进行推理。如果系统发现潜在的运维风险,它可以及时进行警报并进行检查和维修。通过基于动态知识图谱的推理预测模型,对大量车辆设备故障数据的分析,发现设备间的链接模式和规律,从而可以协助工作人员定位故障部位和原因,可以提高运维效率,降低故障发生的风险。

所提方法虽然在一定程度上可以有效地对源设备和目标设备间的关联进行预测,但仅仅依赖从知识图谱中提取时间感知表示来进行预测,会导致预测的准确性较低。因此,未来工作的重点在于综合设备多种模式^[40]下的数据,如视频数据、图像数据以及传感器数据和维修记录等,再进行特征提取后与时间感知表示一起输入到模型中,从而提高推理预测的准确性。

参 考 文 献

- [1] WANG S, LU H, WANG S, et al. Research progress of KPI anomaly detection in intelligent operation and maintenance[J]. *Telecommunications Science*, 2021, 37(5): 42-51.
- [2] XIA L, LIANG Y, LENG J, et al. Maintenance planning recommendation of complex industrial equipment based on knowledge graph and graph neural network[J]. *Reliability Engineering and System Safety*, 2023, 232: 109068.
- [3] SI X S, HU C H, ZHOU Z J. Fault prediction model based on evidential reasoning approach [J]. *Science China Information Sciences*, 2010, 53(10): 2032-2046.
- [4] WANG R G, WU J, LIU C, et al. Aircraft equipment failure cause discrimination method based on maintenance log[J]. *Journal of Software*, 2019, 30(5): 1375-1385.
- [5] FAN Y, NOWACZYK S, RÖGVALDSSON T. Transfer learning for remaining useful life prediction based on consensus self-organizing models[J]. *Reliability Engineering and System Safety*, 2020, 203: 107098.
- [6] CHUI K T, GUPTA B B, VASANT P. A genetic algorithm optimized RNN-LSTM model for remaining useful life prediction of turbofan engine[J]. *Electronics*, 2021, 10(3): 285.
- [7] GUO R, WANG Y, ZHANG H, et al. Remaining useful life prediction for rolling bearings using EMD-RISI-LSTM[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-12.
- [8] REN S, SHI L, LIU Y, et al. A personalized operation and maintenance approach for complex products based on equipment portrait of product-service system[J]. *Robotics and Computer-Integrated Manufacturing*, 2023, 80: 102485.
- [9] CHOI E, XU Z, LI Y, et al. Learning the graphical structure of electronic health records with graph convolutional transformer [C]// *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI, 2020: 606-613.
- [10] WU L, CUI P, PEI J, et al. Graph neural networks: foundation, frontiers and applications[C]// *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM, 2022: 4840-4841.
- [11] DORNAIKA F, DAHBI R, BOSAGHZADEH A, et al. Efficient dynamic graph construction for inductive semi-supervised learning[J]. *Neural Networks*, 2017, 94: 192-203.
- [12] ZHOU X H, WANG Y J, XU H Z, et al. Fusion learning based unsupervised anomaly detection for multi-dimensional time series[J]. *Journal of Computer Research and Development*, 2023, 60(3): 496-508.
- [13] QI Q, SHEN R Y, WANG J Y. GAD: topology-aware time series anomaly detection[J]. *Journal on Communications*, 2020, 41(6): 152-160.
- [14] SHEN Y H, JIANG X H, WANG Y Z, et al. A review of inference research on temporal knowledge graph [J]. *Journal of Computing*, 2023, 46(6): 1272-1301.
- [15] QUAN T, ZHU F, LING X, et al. Learning fair representations by separating the relevance of potential information[J]. *Information Processing & Management*, 2022, 59(6): 103103.
- [16] SEO Y, DEFFERRARD M, VANDERGHEYNST P, et al. Structured sequence modeling with graph convolutional recurrent networks[C]// *The 25th International Conference on Neural Information Processing (ICONIP 2018)*. ICONIP, 2018: 362-373.
- [17] HE M, WEI Z, WEN J R. Convolutional neural networks on graphs with chebyshev approximation, revisited[C]// *Advances in Neural Information Processing Systems*. NIPS, 2022, 35: 7264-7276.
- [18] YU B, YIN H, ZHU Z. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting[C]// *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI 2018)*. IJCAI, 2018: 3634-3640.
- [19] CHEN D, HUANG T, SONG Z, et al. AGG-Net: Attention Guided Gated-convolutional Network for Depth Image Completion[C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*. IEEE, 2023: 8853-8862.
- [20] WU Z, PAN S, LONG G, et al. Graphwavenet for deep spatio-temporal graph modeling[C]// *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. 2019: 1907-1913.
- [21] BI J, YUAN H, XU K, et al. Large-scale Network Traffic Prediction With LSTM and Temporal Convolutional Networks [C]// *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022: 3865-3870.
- [22] ZHENG C, FAN X, WANG C, et al. Gman: A graph multi-attention network for traffic prediction[C]// *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI, 2020: 1234-1241.
- [23] LEIK, QIN M, BAI B, et al. GCN-GAN: A non-linear temporal link prediction model for weighted dynamic networks [C] // *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019: 388-396.
- [24] DE BRUIN G J, VEENMAN C J, VAN DEN HERIK H J, et al. Supervised temporal link prediction in large-scale real-

- world networks[J]. *Social Network Analysis and Mining*, 2021, 11(1):80.
- [25] LV L, BARDOU D, HU P, et al. Graph regularized nonnegative matrix factorization for link prediction in directed temporal networks using PageRank centrality[J]. *Chaos, Solitons & Fractals*, 2022, 159:112107.
- [26] CASOLA S, LAURIOLA I, LAVELLI A. Pre-trained transformers: an empirical comparison[J]. *Machine Learning with Applications*, 2022, 9:100334.
- [27] JI S, PAN S, CAMBRIA E, et al. A survey on knowledge graphs: Representation, acquisition, and applications[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 33(2):494-514.
- [28] BOSAGHZADEH A, DORNAIKA F. Incremental and dynamic graph construction with application to image classification[J]. *Expert Systems with Applications*, 2020, 144:113117.
- [29] XU D, RUAN C, KORPEGLU E, et al. Inductive representation learning on temporal graphs[C]// 8th International Conference on Learning Representations (ICLR 2020). ICLR, 2020: 1-19.
- [30] YAROTSKY D. Error bounds for approximations with deepReLU networks[J]. *Neural Networks*, 2017, 94:103-114.
- [31] XU Y, HUANG H, FENG C, et al. A supervised multi-head self-attention network for nested named entity recognition[C]// Proceedings of the AAAI Conference on Artificial Intelligence. AAAI, 2021:14185-14193.
- [32] LIU S, LIN T, HE D, et al. Paint transformer: Feed forward neural painting with stroke prediction[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE, 2021:6598-6607.
- [33] XU J, SUN X, ZHANG Z, et al. Understanding and improving layer normalization[C]// Proceedings of the 33rd International Conference on Neural Information Processing Systems(NIPS 2019). NIPS, 2019:4381-4391.
- [34] TA H T, RAHMAN A B, MAJUMDER N, et al. WikiDes: A Wikipedia-based dataset for generating short descriptions from paragraphs[J]. *Information Fusion*, 2023, 90:265-282.
- [35] LIU M, LIU Y. Inductive representation learning in temporal networks via mining neighborhood and community influences [C]// Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2021:2202-2206.
- [36] JIANG X, WANG Q, WANG B. Adaptive convolution for multi-relational learning[C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics; Human Language Technologies. ACL, 2019: 978-987.
- [37] TRIVEDI R, FARAJTABAR M, BISWAL P, et al. Dyrep: Learning representations over dynamic graphs[C]// International Conference on Learning Representations (ICLR 2019). ICLR, 2019:1-25.
- [38] SONG C, SHU K, WU B. Temporally evolving graph neural network for fake news detection[J]. *Information Processing & Management*, 2021, 58(6):102712.
- [39] LIU J, LI D, GU H, et al. Parameter-free Dynamic Graph Embedding for Link Prediction[J]. *Advances in Neural Information Processing Systems(NIPS 2022)*. NIPS, 2022:27623-27635.
- [40] HAN Z, YANG F, HUANG J, et al. Multimodal dynamics: Dynamical fusion for trustworthy multimodal classification[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2022:20707-20717.



ZHANG Hui, born in 1982, postgraduate, assistant researcher. His main research interests include intelligent data processing and data engineering.



DING Kun, born in 1978, postgraduate, researcher. His main research interests include data engineering and network management.

(责任编辑:柯颖)