

基于 Transformer 和强化学习的时序知识图谱补全方法

张亚睿¹, 谢璐², 吕佳琪², 雒雄艳², 陈桂军²

¹(太原理工大学 集成电路学院, 山西 晋中 030600)

²(太原理工大学 电子信息工程学院, 山西 晋中 030600)

E-mail: xiejun@tyut.edu.cn

摘要: 针对现有时序知识图谱推理存在捕获长距离依赖关系能力不足以及可解释性缺乏的问题, 提出了一种结合 Transformer 和强化学习的时序知识图谱补全模型. 该模型使用强化学习设计了一种具有高度解释性的新型策略网络, 由时序感知编码器、路径上下文编码器和动作评分器 3 个核心组件构成. 首先, 时序感知编码器利用自注意力机制将时间信息嵌入到关系表示中, 增强了对时间动态性的处理能力; 其次, 路径上下文编码器利用 Transformer 高效编码历史事件序列, 捕捉了长距离依赖关系; 再次, 动作评分器利用双向门控循环单元进行动作预测, 提升了预测准确性. 此外, 针对奖励稀疏性问题, 所提模型引入了一种新型奖励函数, 综合考虑时间塑形奖励、路径长度奖励以及路径多样性奖励, 提供更细致的反馈以优化路径选择. 本文在 4 个公开数据集上与现有先进方法进行比较, 结果表明所提模型在评估指标 MRR 和 Hits@k 上较基线方法均有所提升.

关键词: 时序知识图谱补全; 强化学习; Transformer; 奖励函数

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2026)05-1108-09

Temporal Knowledge Graph Completion Method Based on Transformer and Reinforcement Learning

ZHANG Yaru¹, XIE Jun², LÜ Jiaqi², LUO Xiongyan², CHEN Guijun²

¹(College of Integrated Circuits, Taiyuan University of Technology, Jinzhong 030600, China)

²(College of Electronic Information Engineering, Taiyuan University of Technology, Jinzhong 030600, China)

Abstract: Aiming at the problem that the existing temporal knowledge graph reasoning has insufficient ability to capture long-distance dependencies and lack of interpretability, a temporal knowledge graph completion model combining Transformer and reinforcement learning is proposed. This model uses reinforcement learning to design a new highly interpretable strategy network, which is composed of three core components: time-aware encoder, path context encoder and action scoring device. First, the time-aware encoder uses the self attention mechanism to embed the time information into the relational representation, which enhances the ability to deal with the time dynamics; Secondly, the path context encoder uses Transformer to efficiently encode historical event sequences, capturing long-distance dependencies; Thirdly, the action scorer uses the two-way gated cycle unit to predict actions, which improves the accuracy of prediction. In addition, for the problem of reward sparsity, the proposed model introduces a new reward function, which comprehensively considers time shaping reward, path length reward and path diversity reward, and provides more detailed feedback to optimize path selection. This paper compares the proposed model with the existing advanced methods on four public datasets, and the results show that the proposed model is effective in evaluating the MRR and Hits@k. Compared with the baseline method, the above method has improved.

Keywords: temporal knowledge graph completion; reinforcement learning; Transformer; reward function

0 引言

知识图谱存储了丰富的知识和事实, 被广泛应用于许多下游人工智能应用^[1], 如推荐系统, 问答系统, 对话生成系统等^[2]. 一般来说, 知识图谱是以三元组的形式 (e_s, r, e_o) 来存储事实的, 其中 e_s, e_o 分别表示头实体和尾实体, r 表示关系,

例如(屠呦呦, 荣获, 诺贝尔奖). 为了反映事实的时间属性, 传统的知识图谱被扩展为时序知识图谱(Temporal Knowledge Graph, TKG), 在三元组的基础上引入时间信息形成四元组 (e_s, r, e_o, t) , 例如(屠呦呦, 荣获, 诺贝尔奖, 2015年10月5日). 通过引入时间维度^[3], TKG不仅增强了对信息全面和精确的表达能力, 还为那些高度依赖时间背景的应用场景

收稿日期: 2025-04-30 收修改稿日期: 2025-06-03 基金项目: 国家自然科学基金青年项目(62201377)资助; 虚拟现实技术与系统国家重点实验室(北京航空航天大学)开放课题基金项目(VRLAB2022C1)资助; 山西省重点研究计划项目(202102020101004)资助; 山西省回国留学人员科研基金项目(2024-61)资助. 作者简介: 张亚睿, 女, 1999年生, 硕士研究生, CCF会员, 研究方向为时序知识图谱补全、图神经网络等; 谢璐, 女, 1979年生, 博士, 副教授, CCF会员, 研究方向为自然语言处理、粒计算、粗糙集、数据挖掘、机器学习; 吕佳琪, 女, 2002年生, 硕士研究生, CCF会员, 研究方向为知识图谱、人工智能等; 雒雄艳, 女, 2002年生, 硕士研究生, CCF会员, 研究方向为知识图谱、推荐系统等; 陈桂军, 男, 1987年生, 博士, 副教授, 研究方向为智能信息处理.

提供了强有力的支持^[4]。

考虑到 TKG 的不完整性,时序知识图谱补全(Temporal Knowledge Graph Completion, TKGC)研究有助于从已知事实中推断缺失事件,例如在四元组查询(屠呦呦,荣获,?, 2015 年 10 月 5 日)中缺失实体可以通过知识图谱推理的方法预测得出。给定时间跨度为 $[t_0, t_T]$ 的知识图谱,TKG 可以分为内插和外推两种类型^[5],内插方法旨在推断出满足 $t_0 \leq t \leq t_T$ 下缺失的事实,即对已知时间范围内的事实进行补全。相反,外推方法则专注于对未来时间 $t > t_T$ 下的事实进行预测,本文重点研究 TKGC 外推任务。

在 TKGC 外推任务领域,有翻译模型^[6],张量分解模型^[7],基于神经网络^[8]的方法等,尽管这些研究利用时间信息来预测 TKG 中缺失的元素,但它们本质上是黑箱系统,缺乏足够的解释能力^[9,10]。近年来,在传统知识图谱补全领域表明基于强化学习的多跳推理方法取得了令人鼓舞并且具备解释性的成果,这成功启发了在外推任务中应用强化学习预测未知时间的事实。将强化学习应用于 TKGC 任务,一个关键的挑战在于构建合适的奖励函数,TAgent^[11]首次将强化学习引入外推任务,但是模型仅用基本命中奖励作为模型的最终奖励,限制了模型的性能。TPath^[12]引入了路径多样性奖励,而 TITer^[13]则融入了时间塑形奖励来指导模型学习。尽管这些方法在一定程度上缓解了 TAgent 存在的问题,但仍存在奖励矩阵稀疏的问题。第 2 个关键挑战在于如何捕获时间演变和长距离依赖关系,一些强化学习方法无法捕获历史事件的动态影响。

针对上述问题,本文提出了一种基于 Transformer 和强化学习的时序知识图谱补全模型 TeRL。主要贡献如下:

1) 针对时间动态性和长距离依赖建模困难的问题,设计了一种两阶段动态编码器,第 1 阶段采用时序感知编码器将时间信息融入关系表示中,以捕捉实体间关系的时序演化特性;第 2 阶段通过路径上下文编码器对历史事件序列进行建模,从而更好地刻画长期依赖关系。

2) 针对现有方法在动作预测中仅利用单向序列信息导致预测准确性受限的问题,设计了一个基于双向门控循环单元的动作评分器。该评分器通过双向感知机制,在动作预测过程中提供更全面的视角提升预测性能。

3) 针对多跳推理过程中存在的奖励稀疏问题,提出了一种新的综合奖励函数,该奖励函数融合了时间塑形奖励、路径长度奖励以及路径多样性奖励,使得模型在推理过程中能够更有效地识别与预测结果一致且具有可解释性的路径。

1 相关研究

TKG 中的外推推理已经吸引了大量关注,主要涵盖了基于神经网络的补全方法和基于强化学习的补全方法^[14]。这些方法各自探索了不同的技术路径以解决 TKG 中实体与关系随时间变化的预测问题,为动态环境中知识图谱的补全和未来事实的预测提供了多样化的解决方案。基于神经网络的方法利用深度学习模型捕捉复杂的模式和潜在的关系;基于强化学习的方法通过与环境交互学习最优策略进行预测。

1.1 基于图神经网络的 TKGC 方法

图神经网络(Graph Neural Network, GNN)在 TKGC 中被广泛用于捕捉实体之间或实体与关系之间的内在拓扑相关性,从而获得高质量的嵌入。TeMP^[15]利用 RGCN 建模邻居实体的影响,并引入一个基于频率的门控 GRU 来捕获非活跃事件之间的时间依赖关系;RE-NET^[16]创新性地将推理过程定义为概率推断,同时利用 RNN 捕捉时序交互作用;CyGNet^[8]提出一种复制生成机制捕捉事实的全局重复频率,有效增强了模型对 TKG 中频繁出现的事实或事件的理解能力。然而这些方法难以同时考虑到图谱的结构特性和时间演化特性。GHT^[17]提出一种基于 Transformer 的 TKG 外推推理模型同时捕获了图谱结构信息和时间信息;EvoKG^[18]通过结合事件发生的时间信息与网络结构的演变共同建模,该方法致力于同时反映 TKG 内的结构特点及其随时间演化的趋势;TiR-GN^[19]结合 GNN 与门控循环单元构建局部循环编码器,以探索知识图谱的结构特征和历史依赖;同时,采用全局历史编码器记录事件历史出现情况。尽管这些模型都取得了不错的进展,但是由于神经网络本质上是黑箱性质,这为推理的解释带来了挑战。此外,尽管上述部分模型在一定程度上改善了时间依赖的捕获能力,但在全面、准确地识别和利用长距离时间依赖方面仍显不足。

1.2 基于强化学习的 TKGC 方法

为了提升知识推理过程中的可解释性和透明度,研究者们引入了基于强化学习^[20]的多跳推理机制。基于强化学习的 TKGC 方法可以视为一个马尔可夫决策过程^[21],这种方法在生成可解释性预测方面具有很强的优势。具有代表性的是,TAgent^[11]通过学习识别关键关系路径以指导搜索策略的形成,并采用门控机制筛选候选动作。TPath^[12]将长短期记忆网络与强化学习框架相结合,用来丰富路径奖励的多样性。此外,TITer^[13]通过遍历历史子图和时间链,采用时间奖励函数优化模型。DREAM^[22]提出基于注意力机制的自适应强化学习模型用来支持多跳推理过程;HierarT^[23]通过解析推理层次并融合混合时间编码技术和 K 均值奖励塑形方法,以增强推理过程的效率与准确性。上述方法大多仅侧重于基本奖励命中奖励,使得智能体获得的奖励矩阵是包含大量 0,少量 1 的稀疏矩阵,这种局限性直接导致了奖励机制中出现的稀疏性问题,在一定程度上影响了模型的预测效果。

2 模型描述

2.1 任务定义

1) 基本概念:TKG 是一种特殊的知识表示形式,它通过带有时间戳事实形成的有向多关系图来捕捉相关信息。在本文中,将 TKGC 看作是随时间变化的一系列图快照,它可以表示为 $G_{1:T} = \{G_1, G_2, \dots, G_T\}$,其中 $G_t = (E, R, F_t)$ 表示时刻 t 下的静态知识图谱快照, E 和 R 分别表示实体集和关系集, F_t 表示发生在同一时刻的四元组 (e_s, r, e_o, t) ,其中 $e_s, e_o \in E$, $r \in R$ 。由于 TKG 快照之间没有边的连接,智能体(agent)没有办法从一个快照转移到另一个快照,因此添加了 3 种类型的边^[15]。

时序边:在本文中,TKG 被抽象为一系列由时序信息组成的知识图谱快照,为了使得模型具有对相关历史事件的访

问能力,模型引入了时序边.时序边的连接关系如图1所示,在图1中可以看到实体 e_q 与它的历史尾实体 e_1, e_2, e_3 通过时序边相连接;逆向边,对于每一个TKG四元组 (e_s, r, e_o, t) 都添加 (e_o, r^{-1}, e_s, t) 到知识图谱中,其中 r^{-1} 表示关系 r 的逆向边,因此可以通过将 $(?, r, e_o, t)$ 转换成 $(e_o, r^{-1}, ?, t)$ 来预测头实体;自循环边,在固定步数的多条推理框架下,若智能体提前完成有效路径的推理,缺乏自循环边可能会导致模型进行无意义的动作,从而引入噪声影响预测结果.为此,模型引入了自循环边,当推理路径达到预设的最大步数、智能体已找到目标实体,或当前实体缺乏有效的出边时,允许智能体停留在某一位置充当停止动作.

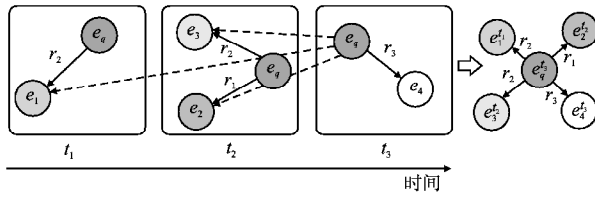


图1 带有时序边的TKG示意图

Fig. 1 Illustration of the TKG with temporal edges

2)推理任务:基于历史事实 $G_{1:t_q-1} = \{(e_i, r, e_j, t_i) | t_i < t_q\}$,TKGC外推的任务是通过回答形式为 $(e_q, r_q, ?, t_q)$ 或 $(?, r_q, e_q, t_q)$ 的查询来推断未来事实.

2.2 模型框架介绍

本节将详细介绍TeRL模型的具体实现方法,模型的整体架构如图2所示.

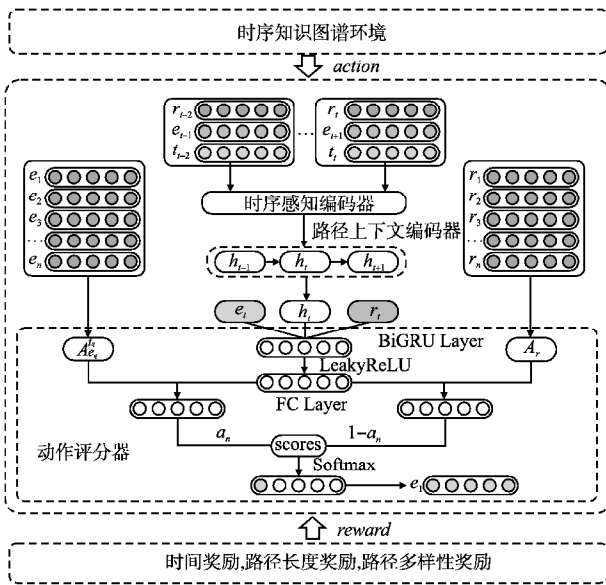


图2 TeRL模型整体结构图

Fig. 2 TeRL model architecture diagram

在强化学习框架内,TeRL模型通过3个关键组件及创新的奖励机制实现其目标.首先,TeRL使用时序感知编码器在关系表达中嵌入时序信息,其次,利用Transformer架构对历史事件序列进行高效编码,捕捉复杂的模式和长距离依赖关系,为动作的选择提供更多的上下文支撑.再次,采用双向门控循环单元(BiGRU)对序列信息进行分析,提高了路径探索

的精确性.此外,在此基础上,为了解决奖励不充分的问题,TeRL提出了一种综合考量时间特征、路径长度及路径多样性的新型奖励函数,为模型提供更精确的反馈.

3 时序知识图谱推理建模

3.1 强化学习框架

强化学习系统由环境(Environment)和智能体(Agent)两部分组成.本文所提模型引入马尔可夫决策过程(MDP)探索TKG环境中的路径规划问题,其目标是选择一系列最优动作.这一过程可以直观地理解为“在图上行走”,从而为TKGC提供可解释的结果.这样模型可以通过与TKG的交互来更新其推理策略,即对缺失元素的预测.马尔可夫决策过程有以下几个组成部分:

状态:状态空间 S 可以看作是所有可能状态的集合.在TKG中状态由TKG四元组中的元素构成.具体来说,在推理步骤 l 中的每个状态可以表示为 $s_l = (e_i, t_l, e_q, t_q, r_q) \in S$,其中 (e_i, t_l) 表示当前推理步骤 l 下的实体以及对应的时间戳,而 (e_q, t_q, r_q) 则对应查询中的实体,时间戳和关系,并在整个推理过程中保持不变.由于基于强化学习的智能体从查询的源节点开始,因此初始状态为 $s_0 = (e_q, t_q, e_q, t_q, r_q)$.考虑到关系随时间的演化特性,在状态表示中引入了自注意力机制用来捕获时间的演化特性.

动作:设 A 表示动作空间, A_l 表示在第 l 步的可选动作集合, A_l 被定义为 $A_l = \{(r', e', t') | (e_i, r', e', t') \in A, t' \leq t_l, t' \leq t_q\}$,且 $A_l \subseteq A$,即所有从实体 e_i 出发,时间戳 t' 满足 $t' \leq t_l$ 且 $t' \leq t_q$ 的关系和目标实体对.此外,在动作空间 A 中添加了自循环边,当推理到最大的推理步数时,自循环边的作用类似于停止动作用来避免推理过程无限展开.

状态转移:新状态 s_{l+1} 由前一个状态 s_l 和动作 A_l 共同更新,转移函数定义为 $P(s_l, A_l) = s_{l+1} = (e_{l+1}, t_{l+1}, e_q, t_q, r_q)$.具体来说,智能体在当前状态 s_l 下执行动作 A_l 后,会根据转移函数进入下一个新的状态 s_{l+1} .这个新状态包含了更新后的实体 e_{l+1} 和时间戳 t_{l+1} ,同时保留查询相关信息 (e_q, t_q, r_q) 不变.

奖励:在强化学习模型中,当智能体经过一系列推理或搜索操作后,若最终结果与预期目标匹配,则给予正向奖励(奖励值为1);反之,若未能找到正确目标实体,则不给予任何奖励(奖励值为0).这种机制通过简单的二进制反馈指导智能体逐步优化其行为策略.具体来说,如果 $s_L = (e_L, t_L, e_q, t_q, r_q)$ 是最终状态, (e_q, r_q, e_{gt}, t_q) 表示知识图谱中的真实事实,基本奖励 $R_{base}(s_L)$ 公式如公式(1)所示:

$$R_{base}(s_L) = I\{e_L = e_{gt}\} \quad (1)$$

其中, $I\{\cdot\}$ 是一个指示函数,用来表示条件是否满足,若条件为真函数值为1,否则为0; e_{gt} 表示实体预测中的正确答案; e_L 表示在状态 s_L 下预测出的结果.

通常,具有相同实体的四元组倾向于集中在特定的时间段内,这种现象导致了时间上的分布特性和时间稀疏性.基于这一特性,查询的目标实体在时间维度上也呈现出特定的分布模式.为了利用这一先验知识,引入了一种基于狄利克雷分布的时间感知奖励机制,并将其融入到奖励函数中,以引导智能体在决策时更加关注那些可能包含答案的时间快照,从而

提升搜索效率.

具体来说,对于一个给定的状态 s_L ,其调整后的奖励值 $R_{time}(s_L)$ 通过原始奖励 $R_{base}(s_L)$ 和一个基于时间差的 Δt_L 的调整因子来计算,参数 P 是从狄利克雷分布 $Dirichlet(\alpha^q)$ 中抽取的一组参数 (p_1, p_2, \dots, p_K) ,如公式(2)~公式(4)所示:

$$R_{time}(s_L) = (1 + p\Delta t_L)R_{base}(s_L) \quad (2)$$

$$\Delta t_L = t_q - t_L \quad (3)$$

$$(p_1, p_2, \dots, p_K) \sim Dirichlet(\alpha^q) \quad (4)$$

对于多跳路径推理任务,较短的路径有助于提高路径的合理性,并且具有较高的准确率.为了限制智能体倾向于找到较短的路径,模型会给较短的路径分配较高的分数,本文定义的路径奖励函数 $R_{path}(s_L)$ 如公式(5)所示.其中, L 表示当前路径的长度.

$$R_{path}(s_L) = \frac{1}{L} \quad (5)$$

通常来说,智能体倾向于选择得分高的路径导致模型容易选择重复路径,为了激励智能体选择不同的路径,本文定义了路径多样性奖励 $R_{diversity}(s_L)$,如公式(6)所示:

$$R_{diversity}(s_L) = \frac{\sum_1^{number_p} p \cos(p, p_i)}{number_p} \quad (6)$$

其中, $number_p$ 表示在头实体和尾实体之间找到路径的数量, p 和 p_i 分别表示当前路径关系组合的嵌入表示和历史路径关系组合的嵌入表示.

综上所述,总奖励 $R_{total}(s_L)$ 是路径多样性奖励 $R_{diversity}(s_L)$ 、时间塑形奖励 $R_{time}(s_L)$ 和路径长度奖励 $R_{path}(s_L)$ 的线性组合,如公式(7)所示:

$$R_{total}(s_L) = \lambda_1 R_{time}(s_L) + \lambda_2 R_{path}(s_L) + \lambda_3 R_{diversity}(s_L) \quad (7)$$

其中, λ_i 是超参数, $\lambda_1, \lambda_2, \lambda_3 \in [0, 1]$ 且 $\lambda_1 + \lambda_2 + \lambda_3 = 1$.

3.2 策略网络

在本文中设计了一个策略网络 $\pi_\theta(a_t | s_t) = p(a_t | s_t; \theta)$ 对连续空间中的智能体进行建模,其中 $a_t \in A_t$, θ 是模型中训练的可变参数,策略网络由以下3部分构成,时序感知编码器,路径上下文编码器以及动作评分器.

3.2.1 时序感知编码器

在TKG中,为了捕捉关系随时间的变化特征,为此提出利用自注意力机制来生成时间信息的关系嵌入,从而增强模型对于时间动态的捕捉能力.对于 G_t 中的每个关系 $r_t = [r_t, t]$ 使用时序感知编码器来将时间信息融入到关系中.

首先查询 Q , 键 K 和值 V 被视为时间表示向量,通过线性投影矩阵 W_q, W_k, W_v 分别对查询、键和值进行变换,自注意力得分可以通过公式(8)计算:

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (8)$$

其中, d_k 表示键向量的维度,用于缩放点积防止梯度消失或爆炸问题, softmax 函数将相似度分数转换为概率分布,接着乘以值向量 V 得到加权的结果.本文采用多头注意力机制来增强模型的表达能力:

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (9)$$

其中, W_i^Q, W_i^K, W_i^V 分别是第 i 个注意力头的参数,最终的多头注意力输出通过拼接所有注意力头的结果并再次线性变换得到:

$$MultiHead(Q, K, V) = (head_1 \oplus head_2, \dots, \oplus head_h) W^O \quad (10)$$

其中, h 表示头的数量, W^O 是输出变换的参数矩阵.最后生成融合时间信息的关系嵌入向量 r_t .

$$r_t = Norm(r_t + MultiHead(Q, K, V)) \quad (11)$$

3.2.2 路径上下文编码器

根据TKG多跳路径推理的特点,由于Transformer能够有效捕捉长距离依赖关系,可以更高效、并行地处理历史信息,提升对复杂路径模式的学习能力.因此选择用Transformer来对历史路径进行编码,路径上下文编码器结构如图3所示.

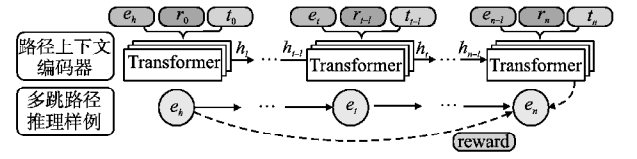


图3 路径上下文编码器示意图

Fig. 3 Path context encoder diagram

具体来说,对于搜索过程中的历史序列,它由一系列动作组成: $h_t = ((e_q, t_q), r_1, (e_1, t_1), \dots, r_t, (e_t, t_t))$, 每个动作包含了一个关系 r 和一个实体-时间对 (e, t) , 也可以用 e^t 来表示. Transformer 采用了多头注意力机制处理历史输入序列,以生成每个时间步骤的状态表示,直到达到路径长度限制或找到正确的尾实体,计算方法如公式(12)所示:

$$h_t = Transformer(r_{t-1}; e_{t-1}^t) \quad (12)$$

$$h_0 = Transformer(r_0; e_q^t) \quad (12)$$

其中, r_0 表示从查询出发 e_q^t 的初始连接关系,当最后一个动作是自循环时,Transformer 会保持状态不发生变化.

3.2.3 动作评分器

动作评分器对每个可选动作进行评分并且计算状态转移概率,设 $a_n = \langle e_n, t_n, r_n \rangle \in A_t$ 表示步骤 t 处的一个可选动作,鉴于未来的事件往往充满不确定性,对于某些查询而言,很难找到明确的因果关系链条.因此,实体与查询间的关联度显得尤为重要.为了解决这个问题,引入了一种加权的动作评分机制.这种机制旨在帮助智能体更加重视目标节点特性或者连接边类型,而不是仅仅依赖于那些具有高度不确定性的预测结果.

在此背景下,首先使用 BiGRU 来编码状态信息,输出预测动作的实体特征嵌入表示 \bar{e} 和关系的嵌入表示 \bar{r} ,如公式(13)和公式(14)所示.其次,智能体进一步计算预测动作与候选动作之间的实体相似度和关系相似度,对两个相似度进行加权求和,获得最终的候选动作评分,评分函数 $\varphi(a_n, s_t)$ 和加权因子 α_n 的计算分别如公式(15)和公式(16)所示:

$$\bar{e} = BiGRU(h_t \oplus e_q^t \oplus r_q) \quad (13)$$

$$\bar{r} = BiGRU(h_t \oplus e_q^t \oplus r_q) \quad (14)$$

$$\varphi(a_n, s_t) = \alpha_n \langle \bar{e}, e_n^t \rangle + (1 - \alpha_n) \langle \bar{r}, r_n \rangle \quad (15)$$

$$\alpha_n = \delta(W_\alpha [h_t; e_q^t; r_q; e_n^t; r_n]) \quad (16)$$

其中, \oplus 表示拼接操作, $\delta(\cdot)$ 表示 LeakyReLU 激活函数, $\langle \cdot \rangle$ 表示相似性计算操作. e_n^t 和 r_n 表示由模型计算出来的候选动作实体和关系的特征嵌入表示.

3.3 训练和优化

模型将搜索长度固定为 L , 从策略网络 $\pi_\theta: \{a_1, a_2, \dots, a_L\}$ 中生成一条长为 L 的路径, 通过最大化所有训练样本 F_{train} 的期望回报来训练策略网络, 目标函数如公式(17)所示:

$$J(\theta) = E_{(e_s, r, e_o, t) \sim F_{train}} [E_{a_1, a_2, \dots, a_L \sim \pi_\theta} [\tilde{R}(s_L | e_s, r, t)]] \quad (17)$$

接着利用策略梯度方法 REINFORCE 算法来对策略进行优化, 该算法会遍历 F_{train} 中的所有四元组, 并通过随机梯度来更新参数 θ , 具体的计算方法如公式(18)所示:

$$\nabla_\theta J(\theta) \approx \nabla_\theta \sum_{m \in [1, L]} \tilde{R}(s_L | e_s, r, t) \log \pi_\theta(a_i | s_i) \quad (18)$$

4 实验分析

本文所有实验均在同一台服务器上完成计算, 具体实验环境为: 操作系统 Ubuntu5. 4. 0-150-generic; GPU NVIDIA GeForce RTX 3090 Ti; CPU Intel (R) Xeon (R) silver 4210R CPU 2. 40GHz; PyTorch 版本为 1. 12. 0; Python 版本为 3. 9. 0; CUDA 版本为 11. 3.

4.1 实验数据集

本文选取 4 个公开数据集 ICEWS14^[24]、ICEWS18^[25]、YAGO11k^[26] 和 Wikidata12k^[27] 进行实验. 其中 ICEWS14 和 ICEWS18 是综合危机预警系统(ICEWS)数据集的子集, 其中 ICEWS14 指的是 2014 年发生的事件, ICEWS18 指的是 2018 年期间发生的事件 YAGO11k 和 Wikidata12k 分别是 YAGO3 和 Wikidata 的子集, 属于开放领域的数据集. 本文采用与文献[16]相同的数据集分割策略, 通过时间戳将数据集划分为训练集、验证集和测试集, 并且设置其时间条件满足: 训练集中数据的时间 < 验证集中数据的时间 < 测试集中数据的时间. 4 个公开数据集的详细信息如表 1 所示.

表 1 时序知识图谱数据集

Table 1 Temporal knowledge graph datasets					
数据集	实体	关系	训练集	验证集	测试集
ICEWS14	6869	230	72826	8941	8963
ICEWS18	23033	256	373018	45995	49545
YAGO11k	10623	10	16408	2050	2051
Wikidata12k	12554	24	32497	4062	4062

4.2 实验设置

4.2.1 实验评价指标

为了衡量知识图谱补全模型的性能, 实验采用平均倒数排序(MRR)值、以及 Hits@k 值(k 值一般为 1、3、10)来评估模型在所有测试数据上的性能. MRR 表示正确实体在候选实体中的平均排名的倒数, 如公式(19)所示:

$$\text{MRR} = \frac{1}{|F_{test}|} \sum_{(e_s, r, e_o, t) \in F_{test}} \frac{1}{2 \left(\frac{1}{\text{rank}_{r, e_o, t}(e_s)} + \frac{1}{\text{rank}_{e_s, r, t}(e_o)} \right)} \quad (19)$$

其中, F_{test} 表示测试集中的所有四元组的集合, $|F_{test}|$ 表示测试集中四元组的个数, $\text{rank}_{r, e_o, t}(e_s)$ 表示 TKGC 模型预测头实体的排序, $\text{rank}_{e_s, r, t}(e_o)$ 表示 TKGC 模型预测尾实体的排序. MRR 指标的值越大, 说明模型的性能越好.

Hits@k 表示正确实体在所有候选实体集中排名小于或

等于 k 的百分比, k 值一般取 1、3、10, 具体计算方法如公式(20)所示. Hits@k 指标的值越大, 说明模型的性能越好.

$$\text{Hits@k} = \frac{1}{|F_{test}|} \sum_{(e_s, r, e_o, t) \in F_{test}} \frac{1}{2} \left(\frac{|\text{rank}_{r, e_o, t}(e_s) \leq k|}{|\text{rank}_{e_s, r, t}(e_o) \leq k|} \right) \quad (20)$$

4.2.2 实验参数设置

由于 TeRL 模型的部分功能受到 TITer 模型启发而实现的, 因此公共参数设置与文献中保持一致. 具体来说, 实体嵌入维度设置为 100, 关系嵌入维度设置为 100, 时间嵌入维度设置为 20. 在每个步骤中, 本文选择最新的 N 条边作为 TeRL 的候选动作, 对于不同的数据集, N 的值分别是: ICEWS14 和 ICEWS18 为 50, Wikidata12k 为 60, YAGO11k 为 30. REINFORCE 算法的折扣因子 γ 设为 0. 95. 推理路径长度的 p 设置为 3. 本文使用 Adam 优化器来优化参数, 学习率设置为 0. 001. 训练过程中, 批量大小设置为 256. 在推理阶段, 本文采用束搜索, 束大小设置为 100.

本文新型奖励函数由 3 部分构成, 分别是时间塑形奖励, 路径长度奖励以及路径多样性奖励. 本文通过线性变换实现了不同奖励函数之间的有效整合, $\lambda_1, \lambda_2, \lambda_3$ 每个参数的取值都会影响到这 3 种奖励在总奖励中所占的比例, 因此 $\lambda_1, \lambda_2, \lambda_3$ 参数的合理设置可以使得智能体获得更加合理的奖励, 实现 TKG 更加准确的多跳路径推理, 经过实验发现, 当 $\lambda_1 = 0. 6, \lambda_2 = 0. 2, \lambda_3 = 0. 2$ 的时候, TeRL 模型达到了最优的性能.

4.2.3 基准方法

为了验证 TeRL 模型的有效性, 本文使用了以下几种最新的基准方法与之进行比较.

- 1) TTransE^[6]. 该方法于 2018 年基于 TransE 框架, 将时间信息整合进关系嵌入中来增强知识图谱中的时间敏感性.
- 2) DE-SimpleE^[28]. 该方法于 2020 年创新性地运用历史实体嵌入, 为 TKG 任意时间点的实体赋予动态表征.
- 3) RE-NET^[16]. 该方法于 2020 年提出利用循环神经网络对事实序列进行建模.
- 4) CyGNet^[8]. 该方法于 2021 年提出了一个复制生成机制预测周期性演化事实.
- 5) GHT^[17]. 该方法于 2022 年提出图霍克斯过程与 Transformer 架构结合的方法建模知识图谱中的时序依赖.
- 6) TPath^[12]. 该方法于 2021 年提出了使用 LSTM 和强化学习框架来提高路径的准确性.
- 7) TAgent^[11]. 该方法于 2021 年利用强化学习识别关键关系路径来指导搜索策略的形成, 并采用了一种门控机制来筛选候选动作.
- 8) TIter^[13]. 该方法于 2021 年提出利用强化学习来对历史事实进行建模.
- 9) HierarT^[23]. 该方法于 2024 年通过层次化强化学习将推理过程解耦为关系推理和实体推理两层.
- 10) MPNet^[29]. 该方法于 2024 年分别从实体和关系两个方面构建策略网络更全面评估候选动作.

4.3 消融实验

为了研究 TeRL 模型的每个组件对性能的贡献程度, 进行了消融实验. 以 ICEWS18 数据集为基准, 本文设计了 6 组

消融实验,实验结果如表 2 所示.

表 2 消融实验结果(%)

方法	MRR	Hits@1	Hits@3	Hits@10
TeRL w/o TAE	28.05	19.72	31.47	44.05
TeRL w/o PCE	30.11	22.01	33.78	45.27
TeRL w/o ASL	29.94	21.95	33.49	44.78
TeRL w/o TSR	28.89	21.36	32.21	43.10
TeRL w/o PLR	29.69	21.59	33.24	44.72
TeRL w/o PDR	30.28	22.24	34.12	45.21
TeRL	33.45	23.85	36.87	50.05

1) TeRL w/o TAE:表示移除了时序感知编码器,仅对关系使用静态表示的变体.

2) TeRL w/o PCE:表示移除了路径上下文编码器,仅使用简单的多层感知机来捕获历史序列信息的变体.

3) TeRL w/o ASL:表示移除了动作评分器,仅使用简单的 LSTM 代替 BiGRU 进行动作预测的变体.

4) TeRL w/o TSR:表示将时间塑形奖励从新型奖励函数中移除的变体.

5) TeRL w/o PLR:表示将路径长度奖励从新型奖励函数中移除的变体.

6) TeRL w/o PDR:表示将路径多样性奖励从新型奖励函数中移除的变体.

从表 2 中可以看出,去除 TAE、PCE 和 ASL 后的模型在 MRR 和 Hits@k 指标上均显著低于 TeRL 模型.去除 TAE 后的模型导致 ICEWS18 中 Hits@10 减少了 6.0%,这表明利用自注意力机制生成的时序感知关系嵌入确实能够提高 TKGC 模型的学习能力,有助于 TeRL 模型更准确地选择正确答案.去除 PCE 后的模型导致 ICEWS18 中 Hits@10 减少了 4.78%,这表明与简单多层感知机相比,使用 Transformer 的路径上下文编码器能更有效地捕获和利用历史序列信息.去除 ASL 后的模型在 MRR 和 Hits@k 指标上显著低于完整 TeRL 模型的表现.这表明相比于 LSTM,使用 BiGRU 的动作

评分器能够更有效地捕捉序列特征及提升动作预测的准确性.

去除 TSR 后的模型导致 ICEWS18 中 Hits@10 减少了 6.95%,这是因为去掉时间塑形奖励后,无法获取出现在整个时间跨度内的知识影响了 TKG 的补全能力.去除 PLR 后的模型导致 ICEWS18 中 Hits@10 减少了 5.33%,这是因为去除路径长度奖励后,路径长度不再受到限制,过长的路径引入了误导性信息从而降低了推理的准确性.去除 PDR 后的模型导致 ICEWS18 中 Hits@10 减少了 4.84%,这是因为去除路径多样性奖励后智能体倾向于选择得分较高的路径,忽视了探索更具潜力的替代路径,导致智能体陷入局部最优路径,从而影响了推理的准确性.

综上所述,TeRL 模型中时序感知编码器、路径上下文编码器、动作评分器以及新型奖励函数中的每一项都对模型性能的提升起到了不可或缺的作用.

4.4 对比实验

为了验证 TeRL 模型的优越性,本文与其他先进模型在 4 个公开数据集 ICEWS14、ICEWS18、Wikidata12k、YAGO11k 和上进行对比,这些对比基线模型整体上可以分为 3 组:基于内插的 TKGC 模型、基于神经网络的 TKGC 外推模型以及基于强化学习的 TKGC 外推模型,实验结果如表 3 和表 4 所示.本文对 TeRL 模型 5 次运行的结果取平均值.如表 3 和表 4 所示,TeRL 在 ICEWS14、ICEWS18、Wikidata12k 和 YAGO11k4 个数据集上的 MRR 值分别达到了 44.74%,33.45%,83.18%,89.88%,相比于多个强化学习基准模型的平均 MRR 值平均提高了 2.11,2.96,5.34,1.66 个百分点.这一结果表明,TeRL 在整体补全准确性方面优于现有主流的强化学习方法,尤其在 Wikidata12k 数据集上展现出显著优势.

从表 3 和表 4 的分组情况来看,可以推断出以下 4 个结论:

1) 整体上,外推的 TKGC 模型效果基本上优于第 1 组基于内插的 TKGC 模型,这是因为内插的模型忽略了时间历史对于预测未来事件的重要性.

表 3 ICEWS14 和 ICEWS18 数据集上对比实验结果(%)

数据集 方法	ICEWS14				ICEWS18			
	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@3	Hits@10
TTransE	13.43	9.4	17.32	34.55	8.31	1.92	8.56	21.89
DE-SimpleE	32.67	24.43	35.69	49.11	19.30	11.53	21.86	34.80
RE-NET	38.28	28.68	41.34	54.52	28.81	19.05	32.44	47.51
CyGNet	32.73	23.69	36.31	50.67	24.93	15.90	28.28	42.61
GHT	37.40	27.77	41.66	56.19	27.40	18.08	30.76	45.76
Tpath	38.9	29.4	42.5	54.6	27.4	21.1	31.3	42.8
TAgent	39.1	30.7	43.2	55.2	27.2	20.1	30.7	42.3
TIter	40.9	32.1	45.5	57.6	29.98	<u>22.05</u>	33.46	44.83
HierarT	43.24	34.14	48.27	<u>60.66</u>	<u>31.00</u>	21.92	<u>34.96</u>	<u>48.85</u>
MPNet	<u>43.76</u>	<u>34.85</u>	<u>48.32</u>	60.61	-	-	-	-
TeRL(ours)	44.74	34.99	50.38	61.83	33.45	23.85	36.87	50.05

注:本文模型实验结果及其提升效果加粗显示,次优结果用下划线显示,“-”代表原文中缺少实验结果

2) 本文提出的模型 TeRL 在所有评价指标上都优于第 2 组基于神经网络的外推基线模型,具体而言,本文所提模型在

预测未来事件方面优于 RE-NET 模型,这是因为利用启发式演化捕捉事件序列中的动态变化时不可避免会将以前的知识

遗忘,进一步表明本文所提模型能够更好地处理长距离依赖关系.

3) 本文提出的模型 TeRL 在所有评价指标上都优于第 3 组基于强化学习的外推基线模型,这是因为 TeRL 模型利用新型的奖励机制优化路径,使得模型能在推理过程中选择最优路径,提高了模型的补全效果.

4) 本文提出的模型在 Wikidata12k 和 YAGO11k 数据集上的补全效果更为明显.这可能是因为 Wikidata12k 和 YAGO11k 是大规模的知识图谱数据集,包含了丰富的实体和关系,从实验结果来看,TeRL 模型能够更好地适应这种复杂的

数据环境,展示了其出色的泛化能力.尽管本文提出的模型在多数指标上表现出色,但在 YAGO11k 数据集的 Hits@3 指标方面,其性能略逊于 MPNet 模型.这一现象归因于 MPNet 模型分别从实体和关系两个维度建模策略网络,这种机制在 Hits@3 这类强调次优路径识别的指标上具有优势,尤其在 YAGO11k 这类实体关系稀疏、知识图谱结构较为松散的数据集中,表明 MPNet 更擅长生成多样化的候选答案.相比之下,本文所提模型在 Hits@1 和 Hits@10 指标上表现更优,表明模型在准确定位最佳答案以及搜索候选答案广度方面优于 MPNet.

表 4 Wikidata12k 和 YAGO11k 数据集上对比实验结果(%)

数据集 方法	ICEWS14				ICEWS18			
	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@3	Hits@10
TTransE	29.27	21.67	34.43	42.39	31.19	18.12	40.91	51.21
DE-SimpleE	45.43	42.6	47.71	49.55	54.91	51.61	57.30	60.17
RE-NET	49.66	46.88	51.19	53.48	58.02	53.06	61.08	66.29
CyGNet	33.89	29.06	36.10	41.86	52.07	45.36	56.12	63.77
GHT	-	-	-	-	-	-	-	-
Tpath	-	-	-	-	-	-	-	-
TAgent	27.66	17.51	-	46.68	-	-	-	-
TIter	75.50	72.96	77.49	79.02	87.47	84.89	89.96	90.27
HierarT	78.50	76.14	80.87	81.51	87.55	85.14	90.07	90.32
MPNet	<u>79.54</u>	<u>76.14</u>	<u>82.45</u>	<u>84.13</u>	<u>89.66</u>	87.00	<u>92.23</u>	<u>92.59</u>
TcRL (ours)	83.18	82.03	84.70	86.22	89.88	88.02	<u>91.19</u>	92.61

注:本文模型实验结果及其提升效果加粗显示,次优结果用下划线显示,“-”代表原论文中缺少实验结果

4.5 敏感性实验分析

1) 推理路径长度,在强化学习中,决策路径的优化是一个至关重要的课题.尤其是在处理复杂 TKG 时,如何有效地找到最优或接近最优的决策路径显得尤为重要.因此本文选择推理路径长度 $p = \{2, 3, 4, 5, 6, 7, 8\}$ 作为敏感性分析中的参数测试范围来探讨不同的推理路径长度 p 在 ICEWS14 和 Wikidata12k 数据集上对于模型性能的影响,实验结果如图 4 所示.

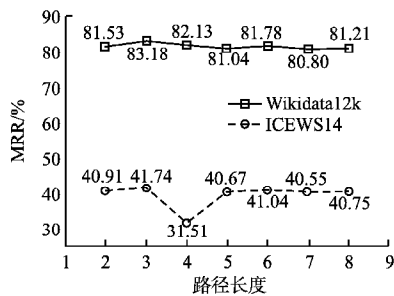


图 4 路径长度对模型性能的影响

Fig. 4 Effect of path length on model performance

从图 4 中可以看出,在 ICEWS14 数据集上当 p 为 3 时,模型表现出最佳的性能,进一步增加 p 并不会带来预期中的提升,反而在 $p=4$ 时出现大幅下降,随后呈现波动趋势.这是因为过长的推理路径会导致错误路径的数量会急剧增加,从而使决策更加困难.为了增强模型的鲁棒性,本文在 Wikidata12k 数据集上进行了重复实验,结果表明在语义关系更为丰

富和图结构更稠密的数据集中,模型依然在 p 为 3 时表现最优.

2) 候选动作数量 N ,在基于强化学习的模型中,动作空间的大小是一个显著影响模型性能的重要因素.为了深入研究候选动作数量对模型表现的具体影响,本文系统地选择了不同的候选动作数量在 ICEWS14 数据集上进行敏感性分析.具体来说,候选动作数量 N 被设置为 $\{10, 20, 30, 40, 50, 60, 70\}$ 以观察其如何影响模型的整体性能.

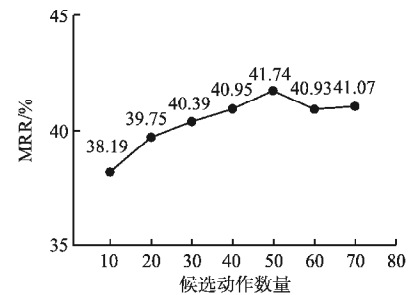


图 5 候选动作数量对模型性能的影响

Fig. 5 Effect of the number of candidate actions on model performance

根据图 5 展示的实验结果可以看出,随着候选动作数量的增加,模型的性能呈现出先上升后趋于稳定的趋势.当 $N=50$ 时,模型性能达到了峰值,随后继续增加 N ,模型性能呈现小幅度下降并且稳定在一定的区间范围之内.这是因为过大

的动作空间引入了过多的选择和噪声,使得模型难以高效地识别出最优的动作路径,从而导致了一定程度上的性能损失。

4.6 样例分析

TeRL 采用强化学习技术,将 TKGK 定义为一个马尔科夫决策过程,并基于训练好的策略网络执行知识推理。为了更直观地展示 TeRL 模型如何利用路径进行推理解释,本文选取了 ICEWS14 数据集中的 3 个实际案例,具体呈现了由 TeRL 模型生成的推理路径,如表 5 所示。

1) TeRL 模型会优先考虑与查询事实相邻的历史事件。第 1 个例子中对于测试集中的四元组 (Citizen (Yemen), Use unconventional violence, Armed Rebel (Yemen), 2014-06-28), 推测目标实体叙利亚, TeRL 模型给出其中的一条推理路径如表 5 所示,可以看出 TeRL 选择了 2014 年 6 月 21 日附近的事件作为路径的一部分,这表明算法在寻找答案时优先考虑了靠近查询时间点的事件。

2) TeRL 模型能够利用时间信息区分相似查询。从表 5

中可以看出前两个查询具有相同的主体实体 Citizen (Yemen) 和关系 Use unconventional violence,唯一不同的是它们对应的时间戳,分别是 2014 年 6 月 28 日和 2014 年 11 月 11 日。对于第 1 个查询,路径通过 Use unconventional violence-1 连接到 Armed Rebel (Yemen),而对于第 2 个查询,则是通过 Carry out suicide bombing 连接到了 Combatant (Houthis),这说明 TeRL 模型能够有效区分并利用时间信息找到正确答案。

3) TeRL 模型能够自适应地调整节点与边的重要性。如公式 (15) 所示,当 TeRL 更加重视邻近节点而非边时,参数 α_n 会增大。从表 5 中可以看出,对于第 3 个查询 (Kazakhstan, Express intent to cooperate, China, 2014-12-14), 其参数 α_n 为 0.95,远高于前两个查询的加权因子(分别为 0.35 和 0.45)。这意味着在这个案例中,TeRL 更多地依赖于实体间的关系而不是具体的时间序列事件来直接找到答案。这也反映了经过训练后,实体表示积累了足够的语义信息,使得模型能够更直接地识别出正确答案。

表 5 ICEWS14 中路径推理可视化
Table 5 Path visualization on ICEWS14

测试集四元组	推理路径	α_n
(Citizen (Yemen), Use unconventional violence, Armed Rebel (Yemen), 2014-06-28)	Citizen (Yemen) $\xrightarrow[t = 2014-06-21]{Use\ unconventional\ violence^{-1}}$ ArmedRebel (Yemen)	0.35
(Citizen (Yemen), Use unconventional violence, Combatant (Houthis), 2014-11-11)	Citizen (Yemen) $\xrightarrow[t = 2014-10-27]{Carry\ out\ suicide\ bombing}$ Combatant (Houthis)	0.45
(Kazakhstan, Express intent to cooperate, China, 2014-12-14)	Kazakhstan $\xrightarrow[t = 2014-11-30]{Express\ intent\ to\ cooperate}$ France $\xrightarrow[t = 2014-12-05]{Engage\ in\ diplomatic\ cooperation}$ China	0.95

5 结论

本文针对当前时序知识图谱补全模型中存在的语义依赖捕获不足、可解释性缺乏以及奖励不充分的问题,创新性地提出了一种基于 Transformer 和强化学习的时序知识图谱补全模型 (TeRL),该模型由时序感知编码器、路径上下文编码器、动作评分器三大核心部件以及新型奖励函数构成。首先时序感知编码器在关系表达中嵌入时间维度的信息,其次利用路径上下文编码器对历史信息进行高效编码,最后利用动作评分器来选择下一步的动作,同时针对奖励稀疏问题,创新性地提出了一种新型奖励函数,综合考虑了时间塑形、路径长度以及路径多样性,为模型提供更加准确的反馈。实验结果显示 TeRL 模型在 4 个公开的时序知识图谱补全数据集上表现优越,表明本文提出的模型在实体预测任务中具有显著的优势。

在当前大数据资源日趋多元化的背景下,融合文本和图像等多模态信息的时序知识图谱已成为发展的趋势。通过对多源异质数据的有效融合,既能丰富图谱的内涵,又能为其提供多维信息支撑,提高预测精度和稳健性。其中,针对具体的应用场景,研究具有针对性的信息融合机制,是今后研究的重点。

References:

- [1] MA H K, QI Y S, WU Y B. Knowledge graph completion method based on disentangled neighborhood information aggregation [J]. Applied Research of Computers, 2024, 41 (3) : 772-778.

- [2] LI C Y, WU Y Q, TANG Z K, et al. Learner personalized resource recommendation based on knowledge graph [J]. Journal of Chinese Computer Systems, 2024, 45 (2) : 285-292.
- [3] XIAO L, LI Q. Survey on time series knowledge graph completion methods [J]. Computer Engineering and Application, 2024, 60 (6) : 43-54.
- [4] SHEN Y H, JIANG X H, WANG Y Z, et al. Review of reasoning on temporal knowledge graphs [J]. Chinese Journal of Computer, 2023, 46 (6) : 1272-1301.
- [5] WANG Y H, CHEN Z Y, ZHAO X, et al. Research progress and trend of temporal knowledge graph representation and reasoning [J]. Journal of Software, 2024, 35 (8) : 3923-3951.
- [6] Leblay J, Chekol M W. Deriving validity time in knowledge graph [C] // Companion Proceedings of the Web Conference, 2018: 1771-1776.
- [7] Ma Y, Tresp V, Daxberger E A. Embedding models for episodic knowledge graphs [J]. Journal of Web Semantics, 2019, 59: 100490, doi:10.1016/j.websem.2018.12.008.
- [8] Zhu C, Chen M, Fan C, et al. Learning from history: modeling temporal knowledge graphs with sequential copy-generation networks [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2021: 4732-4740.
- [9] Wan G, Pan S, Gong C, et al. Reasoning like human: hierarchical reinforcement learning for knowledge graph reasoning [C] // Inter-

- national Joint Conference on Artificial Intelligence, International Joint Conference on Artificial Intelligence, 2021; 1926-1932.
- [10] Zheng S, Chen W, Zhao P, et al. When hardness makes a difference; multi-hop knowledge graph reasoning over few-shot relations [C]//Proceedings of the 30th ACM International Conference on Information & Knowledge Management, 2021; 2688-2697.
- [11] Tao Y, Li Y, Wu Z. Temporal link prediction via reinforcement learning [C]//IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021; 3470-3474.
- [12] Bai L, Yu W, Chen M, et al. Multi-hop reasoning over paths in temporal knowledge graphs using reinforcement learning [J]. Applied Soft Computing, 2021, 103; 107144, doi: 10.1016/j.asoc.2021.107144.
- [13] Sun H, Zhong J, Ma Y, et al. Timetraveler: reinforcement learning for temporal knowledge graph forecasting [C]//Empirical Methods in Natural Language Processing, 2021; 8306-8319.
- [14] LI Z L, CHEN Q, SHI L, et al. Dynamic knowledge graph completion based on time-aware combination [J]. Journal of Zhejiang University: Engineering Science, 2024, 58(8); 1738-1747.
- [15] Wu J, Cao M, Cheung J C K, et al. Temp: temporal message passing for temporal knowledge graph completion [C]//Empirical Methods in Natural Language Processing, 2020; 5730-5746.
- [16] Jin W, Qu M, Jin X, et al. Recurrent event network; autoregressive structure inference over temporal knowledge graphs [C]//Empirical Methods in Natural Language Processing, 2020; 6669-6683.
- [17] Sun H, Geng S, Zhong J, et al. Graph hawks transformer for extrapolated reasoning on temporal knowledge graphs [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2022; 7481-7493.
- [18] Park N, Liu F, Mehta P, et al. Evokg: jointly modeling event time and network structure for reasoning over temporal knowledge graphs [C]//Proceedings of the 15th ACM International Conference on Web Search and Data Mining, 2022; 794-803.
- [19] Li Y, Sun S, Zhao J. TiRGN: time-guided recurrent graph network with local-global historical patterns for temporal knowledge graph reasoning [C]//International Joint Conference on Artificial Intelligence (IJCAI), 2022; 2152-2158.
- [20] Li C, Qiu M. Reinforcement learning for cyber-physical systems: with cybersecurity case studies [M]. Chapman and Hall/CRC, 2019.
- [21] Yi X, Junyong L, Mingjing L, et al. Reason more like human; incorporating meta information into hierarchical reinforcement learning for knowledge graph reasoning [J]. Applied Intelligence, 2022, 53(11); 13293-13308.
- [22] Zheng S, Yin H, Chen T, et al. DREAM: adaptive reinforcement learning based on attention mechanism for temporal knowledge graph reasoning [C]//Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2023; 1578-1588.
- [23] Luo X, Zhu A, Zhang J, et al. HierarT: multi-hop temporal knowledge graph forecasting with hierarchical reinforcement learning [J]. Knowledge-Based Systems, 2024, 300; 112164, doi: 10.1016/j.knosys.2024.112164.
- [24] Liu W, Wang P, Zhang Z, et al. Multi-scale convolutional neural network for temporal knowledge graph completion [J]. Cognitive Computation, 2023, 15(3); 1016-1022.
- [25] Zhao X, Miao J, Yang F, et al. Geometry interaction embeddings for interpolation temporal knowledge graph completion [J]. Mathematics, 2024, 12(13); 2022, doi: 10.3390/math12132022.
- [26] Gao Y, He Y, Kan Z, et al. Learning joint structural and temporal contextualized knowledge embeddings for temporal knowledge graph completion [C]//Findings of the Association for Computational Linguistics (ACL), 2023; 417-430.
- [27] Zhang M, Xia Y, Liu Q, et al. Learning latent relations for temporal knowledge graph reasoning [C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1; Long Papers), 2023; 12617-12631.
- [28] Goel R, Kazemi S M, Brubaker M, et al. Diachronic embedding for temporal knowledge graph completion [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020; 3988-3995.
- [29] Wang J, Wu R F, Wu Y W, et al. MPNet: temporal knowledge graph completion based on a multi-policy network [J]. Applied Intelligence, 2024, 54(3); 2491-2507.

附中文参考文献:

- [1] 马浩凯, 祁云嵩, 吴宇斌. 解纠缠邻域信息聚合的知识图谱补全方法 [J]. 计算机应用研究, 2024, 41(3); 772-778.
- [2] 李春英, 武毓琦, 汤志康, 等. 融合知识图谱的学习者个性化学习资源推荐 [J]. 小型微型计算机系统, 2024, 45(2); 285-292.
- [3] 肖 蕾, 李 琪. 时序知识图谱补全方法研究综述 [J]. 计算机工程与应用, 2024, 60(6); 43-54.
- [4] 沈英汉, 江旭晖, 王元卓, 等. 时态知识图谱的推理研究综述 [J]. 计算机学报, 2023, 46(6); 1272-1301.
- [5] 王俞涵, 陈子阳, 赵 翔, 等. 时序知识图谱表示与推理的研究进展与趋势 [J]. 软件学报, 2024, 35(8); 3923-3951.
- [14] 李忠良, 陈 麒, 石 琳, 等. 时间感知组合的动态知识图谱补全 [J]. 浙江大学学报(工学版), 2024, 58(8); 1738-1747.