

采用消息聚合的演化感知时序知识图谱推理方法

朱可, 秦涛, 朱奎宇, 王晨旭

(西安交通大学智能网络与网络安全教育部重点实验室, 710049, 西安)

摘要: 针对现有时序知识图谱推理补全方法侧重图结构特征建模, 忽略其他重要演化依赖信息的问题, 提出一种基于消息聚合的演化感知时序知识图谱补全方法, 从局部和全局两个角度建模演化历史。基于邻近时间步内事实的演化具有依赖性的假设, 使用演化聚合单元提取邻近时间步内图谱的语义、结构和时序特征。基于全局历史具有重现性的假设, 提取全局历史中查询事实相关的交互实体或关系, 构建候选实体或关系编码矩阵。解码时采用时间增强的 ConvTransE 解码器聚合局部依赖性和全局重复性演化信息, 实现多粒度融合。训练时引入一种由易到难的课程学习策略, 充分捕获历史演化序列的长度多样性。实验结果表明, 该方法在 3 个标准数据集上均优于最优基线方法: 在 ICEWS14 数据集上, 平均倒数排名(M_{RR})、命中率 Hits@3 和 Hits@10 (正确答案出现在预测排名前 3、10 位的比例) 较时序归纳路径神经网络(TiPNN)分别提升了 1.13%、1.19% 和 2.45%; 在 ICEWS18 数据集上, M_{RR} 较基于图卷积网络的循环演化网络(RE-GCN)提升 0.98%, Hits@3 和 Hits@10 较 TiPNN 分别提升 0.73% 和 1.59%; 在 WIKI 数据集上, M_{RR} 较 RE-GCN 提升 0.13%, Hits@3 和 Hits@10 较 TiPNN 分别提升 2.97% 和 2.89%。在多个数据集上的性能提升验证了所提方法在全面建模演化依赖方面的有效性。

关键词: 时序知识图谱; 图卷积网络; ConvTransE 解码器; 课程学习

中图分类号: TP391 **文献标志码:** A

DOI: 10.7652/xjtuxb202511018 **文章编号:** 0253-987X(2025)11-0187-11

Evolution-Aware Temporal Knowledge Graph Reasoning Method Based on Message Aggregation

ZHU Ke, QIN Tao, ZHU Kuiyu, WANG Chenxu

(MOE Key Lab for Intelligent and Network Security, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: To address the issue that existing temporal knowledge graph reasoning and completion methods primarily focus on modeling graph structural features while neglecting other crucial evolutionary dependencies, this study proposes an evolution-aware temporal knowledge graph completion method based on message aggregation, which models evolutionary history from both local and global perspectives. Based on the assumption that the evolution of facts within adjacent timestamps exhibits dependencies, an evolution aggregation unit is employed to extract semantic, structural, and temporal features of the graph within neighboring timestamps. Furthermore, under the hypothesis that global history exhibits recurrence, interactive entities or relations relevant to the query fact are extracted from the global history to construct a candidate entity or

relation encoding matrix. During decoding, a time-enhanced ConvTransE decoder is adopted to aggregate local dependency and global recurrence-based evolutionary information, achieving multi-granularity fusion. A curriculum learning strategy, progressing from easy to difficult, is introduced during training to fully capture the length diversity of historical evolutionary sequences. Experimental results demonstrate that the proposed method outperforms state-of-the-art baselines on three standard datasets: On the ICEWS14 dataset, the mean reciprocal rank (M_{RR}), Hits@3, and Hits@10 (the proportion of correct answers appearing in the top 3 and 10 predicted rankings) improve by 1.13%, 1.19%, and 2.45%, respectively, compared to TiPNN. On the ICEWS18 dataset, M_{RR} increases by 0.98% over RE-GCN, while Hits@3 and Hits@10 improve by 0.73% and 1.59%, respectively, compared to TiPNN. On the WIKI dataset, M_{RR} rises by 0.13% over RE-GCN, and Hits@3 and Hits@10 increase by 2.97% and 2.89%, respectively, compared to TiPNN. The performance improvement on multiple datasets verifies the effectiveness of the proposed method in comprehensively modeling evolutionary dependencies.

Keywords: temporal knowledge graph; graph neural network; ConvTransE decoder; curriculum learning

知识图谱(KG)是人工智能发展过程中出现的一种能够承载和表示知识的技术和工具,它基于三元组(头实体、关系、尾实体)表示知识,以有向多关系图的方式将知识结构化。目前,许多大规模知识图谱,例如 FreeBase、DBPedia 等,已被构建并广泛应用于智能问答、垂直搜索、推荐系统等场景。然而,受限于半自动构建方式及数据来源,大多数图谱是不完整的^[1]。例如,FreeBase 中 71% 的人没有出生日期,75% 的人缺少国籍,这会影响到其下游应用^[2]。因此,大量研究致力于知识图谱推理以补全缺失事实。主流方法将实体和关系嵌入到连续向量空间,并设计三元组评分函数。例如:Bordes 等受 Word2Vec 启发,提出基于平移距离的知识图谱补全方法(TransE)^[3];Nickel 等提出关系学习张量分解方法(RESCAL),用三阶张量表示实体和关系为向量和矩阵,以关系矩阵捕捉实体间的潜在交互^[4];Yang 等简化 RESCAL 为对角关系矩阵,提出 DisMult 模型,缓解了参数过多的问题,但难以处理非对称关系^[5];CompIEx 方法则进一步扩展 DisMult 模型到复数域,使用复数运算的非对称性更好地表示非对称关系^[6]。

现实世界中的数据具有动态性及及时变性,传统静态 KG 主要聚焦于不变的常识概念型知识,难以表达知识的时效性。时序知识图谱(TKG)在其基础上添加时间维度,将事实扩展为头实体、关系、尾实体、时间步四元组,用于建模时效性事件。近年来,许多大规模 TKG 被构建并广泛应用,例如综合危机预警系统(ICEWS)^[7]和全球事件、语言和声调数据库(GDELT)^[8]。然而,与静态知识图谱相同,

TKG 也存在不完整性,并且其复杂的动态特性使得推理补全任务面临更大挑战。

为解决上述挑战,许多工作引入时间信息扩展静态知识图谱补全方法。例如:TTTransE 方法扩展了 TransE,将时间信息融入关系表征中对事实评分^[9];TA-DistMult 方法利用长短期记忆网络建模融合时间的关系表征,并使用 DistMult 评分^[10];TNTComplEx 方法在 ComplEx 方法的基础上,提出了四阶张量正则分解来获得实体的时间感知表征^[11]。这类方法通常利用浅层编码器获取实体、关系和时间表示,并引入时间敏感的解码函数,大多基于内推的假设。然而,在实际应用中,未来的事实尚未发生,基于未来信息的内推不符合现实推理需求。本文聚焦更符合认知的外推推理,即根据当前时间之前的历史事实,预测未来可能发生的事实。

时序知识图谱外推任务基于历史知识图谱预测未来事实,需全面理解历史演化规律。人类认知中,事实演化主要包括两类模式:一是相邻时间事件间的依赖演变,二是历史事件的重复性出现,现有研究多从这两个角度展开。对局部依赖性的捕捉主要聚焦于局部历史,例如:循环事件网络(RENET)基于查询事实固定长度内的子图序列捕捉局部演化信息^[12];基于图卷积网络的循环演化网络(RE-GCN)使用局部历史依赖建模每个时间步的实体和关系演化^[13];面向时序知识图谱推理的动态超图嵌入(DHE-TKG)引入动态超图嵌入获得局部历史中的高阶交互,并与低阶交互融合^[14];时序归纳路径神经网络(TiPNN)构建历史时序图取代独立子图以

建模查询相关路径^[15]。然而,此类方法主要关注固定长度历史中的实体演化,忽略了时间和关系演化。并且固定长度历史序列捕获到的演化信息有限,缺少建模演化序列长度多样性的有效策略。对全局重复性的捕捉主要聚焦于全局历史,例如时序复制生成网络(CyGNet)采用复制生成机制获取单跳事实的全局重复性用于实体推断^[16]。对比事件网络(CENET)学习全局历史和非历史依赖关系来识别最相关实体^[17]。部分方法尝试整合两类模式以提升效果。例如,时间引导的递归图网络(TiRGN)采用局部循环图结构编码邻近时间的依赖特征,并辅以全局历史信息编码模块挖掘图谱中反复出现的事实^[18]。类似地,历史信息传递网络(HIPNet)从时间、结构和重复角度预测未来事件^[19];重复-局部-全局历史网络(RLGNNet)从重复、局部和全局 3 个层次捕获和整合历史信息^[20]。然而,以上方法多聚焦图内结构和图间时序信息,忽略了图谱中丰富的语义信息。近年来,预训练语言模型(PLM)因具有强大的语义和上下文理解能力,被引入时序知识图谱推理任务中。例如:面向 TKG 补全的带提示的预训练语言模型(PPT)采样一系列四元组并用提示代替时间步之间的间隔,然后利用 PLM 掩码预测^[21];上下文学习(ICL)把历史事实构建成提示输入大语言模型 GPT-NeoX 来获得标记实体的概率从而得到排名预测^[22];带前缀调优的时序推理对比历史建模框架(ChapTER)转化 TKG 推理为查询-候选匹配问题,设计伪孪生网络并引入虚拟时间前缀进行高效微调^[23];历史事实学习(HFL)从长短期角度采样历史事实并转化实体、关系和时间间隔为特定标记进行掩码预测^[24]。此类方法虽然较好地捕捉了语义信息,但图谱中许多事件缺乏历史信息的稀疏特性,限制了事件关联信息的挖掘。

基于上述研究现状,为了全面建模历史演化信息用于推断未来事实,本文设计局部依赖性演化编码器和全局重复性演化编码器捕捉历史演化信息。其中,局部依赖性演化编码器基于邻近时间步内事实的演化具有依赖性的假设,分别使用预训练语言模型捕捉语义特征、图神经网络捕捉图内结构特征、门控机制捕捉图间时序特征,充分聚合邻近时间步内事实间的时序依赖。全局重复性演化编码器基于历史具有重现性,以及实体更容易与曾经交互过的实体或关系再次发生关联的假设,对全局历史中与查询实体交互过的实体或关系进行编码,构建候选实体或关系矩阵。采用时间增强的 ConvTransE 解

码器融合局部和全局的演化信息进行评分。使用一种基于课程学习的训练策略,使模型自适应地学习由短到长的局部历史序列演化模式,充分捕获历史演化序列的长度多样性。

1 问题定义

时序知识图谱可形式化为 $G = (E, R, T, F)$, 其中 E, R, T 和 F 分别表示实体、关系、时间步和四元组集合。在 G 中,事实以四元组 (s, r, o, t) 存储, $s, o \in E$ 分别为头实体和尾实体, $r \in R$ 为关系, $t \in T$ 为时间步, $(s, r, o, t) \in F$ 表示在时间 t 头实体 s 和尾实体 o 之间存在关系 r 。按时间步升序排列可以得到一系列 TKG 快照 $G = \{G_1, G_2, \dots, G_t, \dots, G_T\}$ 。

时序知识图谱补全旨在根据已知事实挖掘潜在事实从而补全图谱,分为内推和外推两类设置。对于 $G = \{G_1, G_2, \dots, G_T\}$,内推是指推理 $[0, T]$ 时间内的事实,外推是指预测 T 时间后的事实。其中,外推更符合现实场景中依赖历史预测未来的需求,本文聚焦于外推补全任务。依据补全目标,可划分为实体补全和关系补全。令 $?$ 表示缺失元素,实体补全指给定缺失四元组 $(s, r, ?, t)$ 或 $(?, r, o, t)$,从候选实体中选择使该事实四元组成立的实体进行补全;关系补全是指给定缺失四元组 $(s, ?, o, t)$,从候选关系中选择在时间步 t 头实体 s 和尾实体 o 之间最有可能存在的关系。本文主要以尾实体补全为例介绍模型。

2 模型设计

基于从历史推及未来这种直观的推理模式,建模历史知识图谱序列的演化特征对于补全当前缺失事实十分重要。为此,本文提出一种基于消息聚合的演化感知时序知识图谱补全方法。以尾实体补全为例,图 1 展示了模型的主要框架。

如图 1 所示,模型采用编码器-解码器框架,目标是给定历史知识图谱序列 $G = \{G_1, G_2, \dots, G_{t-1}\}$,补全知识图谱 G_t 中的缺失事实 $(s, r, ?, t)$ 。本模型从距查询最近的 k 个时间步提取邻近历史图谱序列 $\{G_{t-k}, \dots, G_{t-1}\}$ 的演化特征,并将其建模为时间步 t 的实体表征 $\mathbf{H}_t \in \mathbb{R}^{|E| \times d}$ 、关系表征 $\mathbf{R}_t \in \mathbb{R}^{|R| \times d}$ 和时间表征 $\mathbf{Z}_t \in \mathbb{R}^{1 \times d}$,其中, d 表示向量维度。同时,从全局历史知识图谱序列 $\{G_1, G_2, \dots, G_{t-1}\}$ 中提取候选实体矩阵 $\mathbf{M}_t$ 。据此,可将补全形式化定义为给定历史序列演化特征及缺失四元组进行实体预测的条件概率,可表示如下

$$p(o | s, r, G_{1:t-1}) = p(o | s, r, \mathbf{H}_t, \mathbf{R}_t, \mathbf{Z}_t, \mathbf{M}_t) \quad (1)$$

2.1 局部依赖性演化编码

2.1.1 结构特征聚合

在知识图谱中,实体和关系通过事实相互关联,为捕捉图谱的结构信息,本模型使用图神经网络进行消息聚合。对于时间步 t ,在每一层图注意力神经网络中,中心节点 o 的更新方式如下

$$\mathbf{h}_{o,t}^{l+1} = f\left(\mathbf{W}_1^l \mathbf{h}_{o,t}^l + \frac{1}{\lambda_o} \sum_{(s,r,o) \in G_t} \mathbf{W}_2^l (\mathbf{h}_{s,t}^l + \mathbf{r}_t^l)\right) \quad (2)$$

式中: $\mathbf{h}_{s,t}^l, \mathbf{h}_{o,t}^l, \mathbf{r}_t^l \in \mathbb{R}^d$ 分别表示时间步 t 第 l 层实体 s, o 和关系 r 的表征向量; $\mathbf{h}_{o,t}^{l+1}$ 表示第 $l+1$ 层实体 o 的表征向量, $l \in [0, a-1]$, a 为图神经网络的层数。 $\mathbf{h}_{s,t}^0, \mathbf{h}_{o,t}^0$ 和 $\mathbf{r}_t^0$ 由上一时间步的实体和关系向量进行初始化,否则随机初始化; $\mathbf{W}_1^l$ 和 $\mathbf{W}_2^l$ 分别表示 l 层自环消息和聚合消息的参数; λ_o 是实体 o 的入度归一化系数; $f(\cdot)$ 是 RReLU 激活函数。

关系表征的更新通过聚合该关系相关的实体和关系实现,表达式如下

$$\mathbf{r}_t^{l+1} = \frac{1}{\lambda_r} \sum_{(s,r,o) \in G_t} \mathbf{W}_3^l (\mathbf{h}_{s,t}^l + \mathbf{h}_{o,t}^l) + \mathbf{r}_t^l \quad (3)$$

式中: $\mathbf{r}_t^{l+1}$ 表示更新后第 $l+1$ 层关系 r 的表征; $\mathbf{W}_3^l$ 表示 l 层聚合消息的参数; λ_r 表示在时间步 t 具有关系 r 的边的数目。

对于每种关系,计算图中所有包含该关系的边的加和平局,以此作为更新后的关系的表征。

在 a 层消息聚合后,得到每一个实体的表征 $\mathbf{h}_{s,t}^a$ 或 $\mathbf{h}_{o,t}^a$ 以及每一个关系的表征 $\mathbf{r}_t^a$,组合后得到时间步 t 所有实体和所有关系的表征 $\mathbf{H}_t^w \in \mathbb{R}^{|E| \times d}$ 和 $\mathbf{R}_t^w \in \mathbb{R}^{|R| \times d}$,上标 w 表示结构。

2.1.2 语义特征聚合

事实本身的语义信息对于推理也十分重要,因此,本模型引入预训练语言模型——Transformer 双向编码表示模型(BERT),对事实四元组进行语义建模来增强实体和关系的语义表征。

由于时序知识图谱的规模通常很大,并且可能存在重复事实,将全部事实进行编码不仅计算开销大,而且可能掩盖关键信息,因此本文通过采样来获取部分代表性事实。具体地,对于时间步为 t 的图谱中的每种关系随机采样至少 m 个事实,若包含某种关系的事实数量少于 m ,则全部采样。在得到包含 n 个事实的采样子图 G'_t 后,将 G'_t 中的每个事实 (s, r, o) 的文本用 BERT 模型特殊标记符号[SEP](表示分隔符)与[CLS](表示分类符)拼接,得到 $T_{\text{ext}} =$

$\{[\text{CLS}]s[\text{SEP}]r[\text{SEP}]o\}$,然后输入 BERT 模型编码得到如下表征

$$\mathbf{E}^{s,r,o} = [\mathbf{E}_{[\text{CLS}]}, \mathbf{E}_s, \mathbf{E}_{[\text{SEP}]}, \mathbf{E}_r, \mathbf{E}_{[\text{SEP}]}, \mathbf{E}_o] = \text{BERT}(T_{\text{ext}}) \quad (4)$$

取 $[\text{CLS}]$ 的表征 $\mathbf{E}_{[\text{CLS}]}$ 作为事实的语义表征。对每个实体和关系,累加其所在事实的语义表征并求均值,计算过程如下

$$\mathbf{h}_t' = \frac{1}{n_s + n_o} \left(\sum_{(s,i,j) \in G'_t} \mathbf{E}_t^{s,i,j} + \sum_{(i,j,o) \in G'_t} \mathbf{E}_t^{i,j,o} \right) \quad (5)$$

$$\mathbf{r}_t' = \frac{1}{n_r} \sum_{(i,r,j) \in G'_t} \mathbf{E}_t^{i,r,j} \quad (6)$$

式中: n_s, n_o, n_r 分别表示实体和关系的出现频次; $\mathbf{E}_t^{s,i,j}, \mathbf{E}_t^{i,j,o}, \mathbf{E}_t^{i,r,j}$ 分别表示对应事实的语义表征。

将所有实体进行拼接得到 $\mathbf{H}_t' \in \mathbb{R}^{|E| \times q}$,将所有关系进行拼接得到 $\mathbf{R}_t' \in \mathbb{R}^{|R| \times q}$, q 表示 BERT 模型的输出向量维度。通过线性映射和激活函数后,得到实体和关系的语义表征

$$\mathbf{H}_t^c = \sigma(\mathbf{W}_1 \mathbf{H}_t' + \mathbf{b}_1) \quad (7)$$

$$\mathbf{R}_t^c = \sigma(\mathbf{W}_2 \mathbf{R}_t' + \mathbf{b}_2) \quad (8)$$

式中: $\mathbf{W}_1, \mathbf{W}_2 \in \mathbb{R}^{q \times d}$ 是线性映射参数; $\mathbf{b}_1, \mathbf{b}_2$ 是偏置参数; $\mathbf{H}_t^c \in \mathbb{R}^{|E| \times d}, \mathbf{R}_t^c \in \mathbb{R}^{|R| \times d}$,上标 c 表示语义; σ 为 Sigmoid 函数。

2.1.3 时序特征聚合

对于时间窗口长度为 k 的历史知识图谱 $\{G_{t-k}, \dots, G_{t-1}\}$,为了避免自回归迭代时序列长度增加引起的过度平滑及梯度消失问题,本模型施加门控机制聚合上一时间步和当前时间步信息。经过 k 次自回归迭代从而得到聚合 k 个时间步历史演化特征的实体表征、关系表征和时间表征 $\mathbf{H}_t, \mathbf{R}_t, \mathbf{Z}_t$ 。

(1)实体和关系表征。由图神经网络和 BERT 模型分别得到结构特征 $\mathbf{H}_{t-1}^w \in \mathbb{R}^{|E| \times d}, \mathbf{R}_{t-1}^w \in \mathbb{R}^{|R| \times d}$ 和语义特征 $\mathbf{H}_{t-1}^c \in \mathbb{R}^{|E| \times d}, \mathbf{R}_{t-1}^c \in \mathbb{R}^{|R| \times d}$,与上一时间步的实体表征 $\mathbf{H}_{t-1} \in \mathbb{R}^{|E| \times d}$ 、关系表征 $\mathbf{R}_{t-1} \in \mathbb{R}^{|R| \times d}$ 通过门控机制融合,得到当前步的实体和关系表征如下

$$\mathbf{H}_t = \mathbf{U}_t^h \odot (\mathbf{H}_{t-1}^w + \mathbf{H}_{t-1}^c) + (1 - \mathbf{U}_t^h) \odot \mathbf{H}_{t-1} \quad (9)$$

$$\mathbf{R}_t = \mathbf{U}_t^r \odot (\mathbf{R}_{t-1}^w + \mathbf{R}_{t-1}^c) + (1 - \mathbf{U}_t^r) \odot \mathbf{R}_{t-1} \quad (10)$$

式中: $\odot$ 表示逐元素相乘; $\mathbf{U}_t^h \in \mathbb{R}^{|E| \times d}$ 和 $\mathbf{U}_t^r \in \mathbb{R}^{|R| \times d}$ 分别为时间步 t 实体和关系更新的控制门,其计算方式如下

$$\mathbf{U}_t^h = \sigma(\mathbf{W}_h \mathbf{H}_{t-1} + \mathbf{b}_h) \quad (11)$$

$$\mathbf{U}_t^r = \sigma(\mathbf{W}_r \mathbf{R}_{t-1} + \mathbf{b}_r) \quad (12)$$

式中: $\mathbf{W}_H, \mathbf{W}_R \in \mathbb{R}^{d \times d}$ 为实体和关系的门控权重参数; $\mathbf{b}_H, \mathbf{b}_R$ 为实体和关系的门控偏重参数。

(2) 时间表征。为建模时间的独特性, 本模型引入随时间步更新的向量 $\mathbf{Z}_t \in \mathbb{R}^{1 \times d}$ 。首先, 基于位置信息计算时间步 $t-1$ 的位置嵌入 $\mathbf{P}_{t-1} \in \mathbb{R}^{1 \times d}$ 如下

$$\mathbf{P}_{t-1}(j) = \begin{cases} \sin\left(\frac{t-1}{10\,000^{j/d}}\right), & j = 2i \\ \cos\left(\frac{t-1}{10\,000^{j/d}}\right), & j = 2i+1 \end{cases} \quad (13)$$

式中: j 代表 d 中的奇偶位置。

然后, 对上一时间步的时间表征 $\mathbf{Z}_{t-1}$ 进行非线性映射后与位置嵌入相加, 获得具有上一时间步特征和当前位置特征的时间嵌入

$$\mathbf{Z}_t = \sigma(\mathbf{W}_Z \mathbf{Z}_{t-1} + \mathbf{b}_Z) + \mathbf{P}_{t-1} \quad (14)$$

式中: $\mathbf{W}_Z \in \mathbb{R}^{d \times d}$ 是时间表征的权重参数; $\mathbf{b}_Z$ 为偏重参数。若 $t-1$ 为首个时间步, 则随机初始化 $\mathbf{Z}_{t-1}$ 。

2.2 全局重复性演化编码器

历史上发生过的事实可能在未来会再次出现, 与查询曾经交互过的实体或关系, 可能会再次产生交互, 因此, 本模型设计全局重复性演化编码器, 捕捉全局历史中的重复性演化特征。对于实体查询 $(s, r, ?, t)$ 或 $(?, r, o, t)$, 遍历 t 之前的所有图谱 $\{G_1, G_2, \dots, G_{t-1}\}$, 提取查询实体在每个快照中的一跳邻域实体集合 $\{c_1^s, c_2^s, \dots, c_{t-1}^s\}$ 或 $\{c_1^o, c_2^o, \dots, c_{t-1}^o\}$, 对其取并集构成 $C_t^s = c_1^s \cup c_2^s \cup \dots \cup c_{t-1}^s$ 或 $C_t^o = c_1^o \cup c_2^o \cup \dots \cup c_{t-1}^o$ 。然后, 据此构建候选实体矩阵 $\mathbf{M}_t^s \in \mathbb{R}^{|E|}$ 或 $\mathbf{M}_t^o \in \mathbb{R}^{|E|}$, 其中, C_t^s 或 C_t^o 中存在的实体所在位置赋值 1, 其他位置赋值 0。同理, 对于关系查询 $(s, ?, o, t)$, 提取每个快照中实体 s 和 o 一跳子图关系集合 $\{c_1^s, c_2^s, \dots, c_{t-1}^s\}$, 取并集得到 $C_t^s = c_1^s \cup c_2^s \cup \dots \cup c_{t-1}^s$ 。然后, 据此构建候选关系矩阵 $\mathbf{M}_t^r \in \mathbb{R}^{|R|}$, C_t^s 中存在的关系所在的位置赋值 1, 其他位置赋值 0。

2.3 时间增强的 ConvTransE 解码器

ConvTransE^[25] 是常用于 TKG 补全模型 (如 RE-GCN^[13]、TiRGN^[18] 等) 中的解码器, 它将实体和关系拼接后进行卷积操作, 具备较好的准确性和可解释性。本文在其基础上添加时间表征的拼接, 设计了时间增强的 ConvTransE 解码器。对于缺失四元组 $(s, r, ?, t)$, 首先将头实体 $\mathbf{h}_{s,t}$ 、关系 $\mathbf{r}_t$ 和时间表征 $\mathbf{z}_t$ 拼接成 $3 \times d$ 的矩阵, $\mathbf{z}_t \in \mathbb{R}^d$ 为当前时间步下时间表征 $\mathbf{Z}_t \in \mathbb{R}^{1 \times d}$ 的首个行向量, 然后使用 C 个 $3 \times K$ 的卷积核对拼接的矩阵做卷积操作, 计算过程如下

$$m_c(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t, n) =$$

$$\sum_{\tau=0}^{K-1} \omega_c(\tau, 0) \hat{\mathbf{h}}_{s,t}(n+\tau) + \omega_c(\tau, 1) \hat{\mathbf{r}}_t(n+\tau) + \omega_c(\tau, 2) \hat{\mathbf{z}}_t(n+\tau) \quad (15)$$

式中: $m_c(\cdot)$ 为卷积操作; c 为不同的卷积核, 共 C 个; K 是核的宽度; $n \in [0, d-1]$; ω_c 是可学习核参数; $\hat{\mathbf{h}}_{s,t}, \hat{\mathbf{r}}_t, \hat{\mathbf{z}}_t$ 分别是 $\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t$ 的填充结果, 使得卷积输出维度为 d , 填充长度为 $\lceil K/2 \rceil, \lceil \cdot \rceil$ 为向上取整符号。

对填充后的拼接矩阵用单个卷积核进行卷积操作可以得到

$$M_c(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t) = [m_c(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t, 0), \dots, m_c(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t, d-1)] \quad (16)$$

将所有通道的输出 $M_c(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t)$ 拼接得到卷积结果 $\mathbf{M}(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t) \in \mathbb{R}^{C \times d}$, 然后经向量化、线性变换和 ReLU 激活, 与所有候选实体内积得到候选实体与缺失四元组匹配程度的得分, 计算过程如下

$$\psi(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t) = \text{ReLU}(\text{vec}(\mathbf{M}(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t)) \mathbf{W}_s) \mathbf{H}_t \quad (17)$$

式中: $\text{vec}(\cdot)$ 表示向量化; $\mathbf{W}_s \in \mathbb{R}^{Cd \times d}$ 表示线性转移的权重参数; ReLU 为激活函数; $\psi(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t) \in \mathbb{R}^{|E|}$ 表示所有候选实体的匹配得分。

将得分结果与候选实体矩阵 $\mathbf{M}_t^i$ 逐元素相乘进行过滤, 并使用加权求和的方式计算候选实体与缺失四元组最终的概率得分, 计算过程如下

$$p(o | s, r, G_{1:t-1}) = \lambda \psi(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t) + \psi(\mathbf{h}_{s,t}, \mathbf{r}_t, \mathbf{z}_t) \odot \mathbf{M}_t^i \quad (18)$$

式中: λ 为可学习的参数。

在得到所有候选实体的评分后, 选择评分最高的实体作为补全结果。模型训练时采用交叉熵损失函数。令 $\mathbf{y}_t^o \in \mathbb{R}^{|E|}$ 表示时间步 t 的标签向量, 若候选四元组为真, 则标签值为 1, 否则为 0。损失函数计算如下

$$L_o = - \sum_{(s,r,o,t) \in D_{\text{train}}} \sum_{i=0}^{|E|-1} y_{t,i}^o \ln p_i(o | s, r, G_{1:t-1}) \quad (19)$$

式中: $y_{t,i}^o$ 是 $\mathbf{y}_t^o$ 中的第 i 个元素; $p_i(o | s, r, G_{1:t-1})$ 表示目标实体是第 i 个元素的概率; D_{train} 为训练集。

2.4 基于课程学习的训练策略

对于时序模型来说, 较长的历史序列虽然包含更多的演化历史信息, 但对模型的学习能力提出了更高的挑战, 而较短的历史序列则存在演化历史信息不足的问题。并且, 固定长度的历史序列会限制模型学习固定的依赖模式。因此, 本模型借鉴人类由易到难的课程学习过程, 从短到长

逐步增加历史序列长度进行模型训练,使模型自适应地学习不同长度的演化依赖。图 1 展示了具体的课程学习过程。

如图 1 所示,设置 k^1 为最小历史序列长度, k^n 为最大历史序列长度。 M^n 表示使用长度为 n 的历史图谱序列训练后得到的模型, M^0 表示初始模型。从 k^1 开始训练模型,然后逐步地增加序列长度在已有模型的基础上进行训练,当模型的性能指标平均倒数排名不再提升或达到最大历史序列长度 k^n 时停止训练得到最终模型。

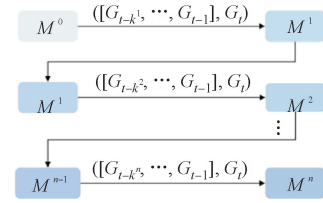


图 1 基于课程学习的训练策略

Fig. 1 Training strategy based on curriculum learning

以尾实体补全为例,图 2 展示了模型的主要框架。

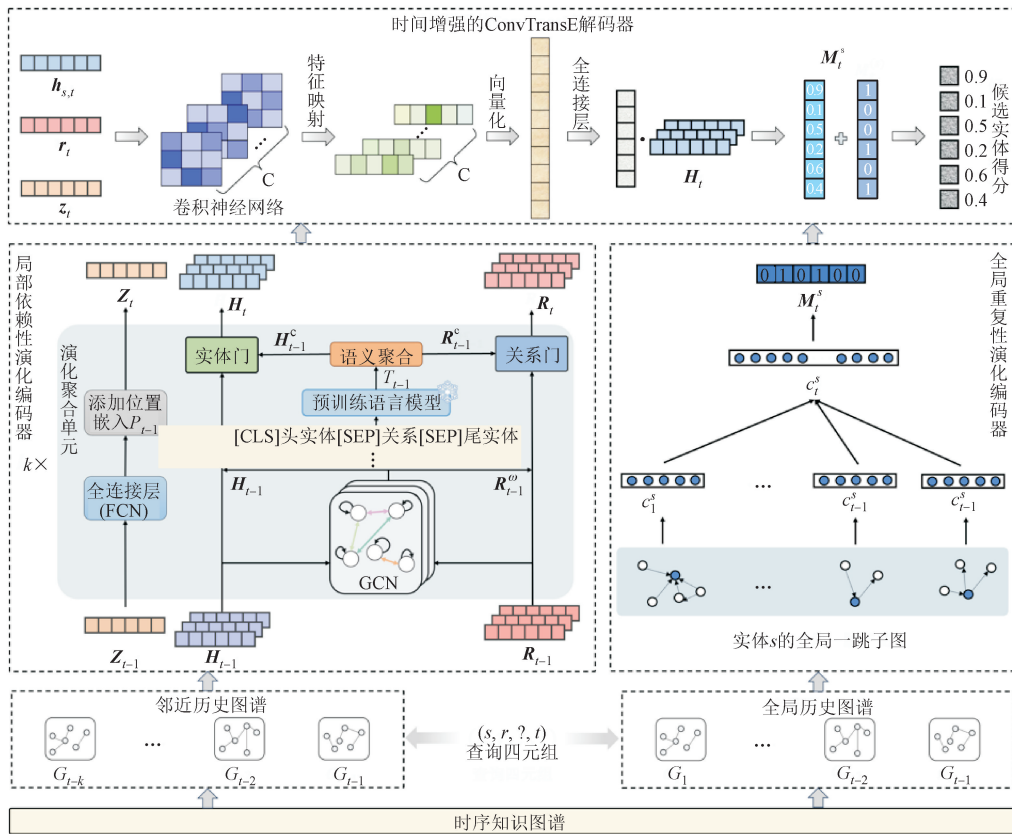


图 2 基于消息聚合的演化感知时序知识图谱推理模型框架图

Fig. 2 Framework of the evolution-aware reasoning model over temporal knowledge graph based on message aggregation

3 模型性能评估

3.1 数据集

实验使用 ICEWS14^[13]、ICEWS18^[12] 和 WIKI^[9] 3 个标准数据集。其中,ICEWS14 和 ICEWS18 均来源于综合危机预警系统(ICEWS)数据库,该数据库整合了 100 多个数据源,涵盖全球 250 多个国家和地区的实时新闻报道,并以每日更新的频率提供最新事件信息,支持政策制定中的危机干预与资源分配。ICEWS14 和 ICEWS18 分别是

ICEWS 中包含 2014 年和 2018 年事实的子集,时间标注为事件发生的精确时间点,更真实地反映了事件的动态演化过程。WIKI 是从 Wikidata 中提取的事件子集。Wikidata 是维基媒体基金会支持的项目之一,为维基百科、维基文库等项目提供结构化和半结构化数据支撑,涵盖人物、组织、事件等信息。其时间表示方式为时间段(事实开始时间和终止时间),适用于建模具有较长持续周期的事件演化过程。表 1 总结了 3 个数据集的统计信息。

表 1 数据集统计信息
Table 1 Statistics of the datasets

数据集	实体数	关系数	四元组数			时间步数	时间间隔/d
			训练集	验证集	测试集		
ICEWS14	7 128	230	74 845	8 514	7 371	365	1
ICEWS18	23 033	256	373 018	45 995	49 545	304	1
WIKI	12 554	24	539 286	67 538	63 110	232	365

3.2 评估指标

平均倒数排名(M_{RR})表达式为

$$M_{RR} = \frac{1}{|D_{test}|} \sum_{(s,r,o,t) \in D_{test}} \frac{1}{\text{rank}(s,r,o,t)} \quad (20)$$

式中: $\text{rank}(s,r,o,t)$ 表示真实四元组在所有候选四元组中的排名; D_{test} 表示测试集中的真实四元组。

命中率 $\text{Hits}@n$ (n 取 3, 10) 表达式为

$$\text{Hits}@n = \frac{1}{|D_{test}|} \sum_{(s,r,o,t) \in D_{test}} I(\text{rank}(s,r,o,t) \leq n) \quad (21)$$

式中: $I(\cdot)$ 表示计数函数。

3.3 基线方法

为验证模型性能,与时序知识图谱内推方法和外推方法共 18 种代表性方法进行对比。内推方法包括 TTransE^[9]、TA-DistMult^[10]、DE-Simple^[26]、TNTComplEx^[11]。外推方法包括 RE-NET^[12]、CyGNet^[16]、xERTE^[27]、TANGO-Tucker^[28]、TANGO-DistMult^[28]、RE-GCN^[13]、CENET^[17]、PPT^[21]、ICL^[22]、ChapTER^[23]、RLGNet^[20]、DHE-TKG^[14]、TiPNN^[15]、HFL^[24]。

3.4 实验设置

模型基于 Pytorch 实现,在搭载 GeForce RTX 3090 GPU 的平台上训练。对于所有数据集,实体、关系、时间向量的维度为 200,训练轮数为 100,采用 Adam 进行参数学习,BERT 模型的学习率为 2×10^{-5} ,其他模块的学习率为 1×10^{-3} 。GCN 层数为 2,丢弃率为 0.2。CNN 的卷积核宽为 3,卷积核数为 50。在 ICEWS14、ICEWS18 和 WIKI 数据集中,每种关系的采样数分别为 12、85、1,最小历史序列长度 k^1 分别为 1、1、4,最大历史序列长度 k^n 均为 12。采用与 TANGO-Tucker 方法相同的时间感知过滤策略,仅过滤与查询事实同时间出现的答案。

3.5 实验结果

表 2 展示了不同方法在 ICEWS14、ICEWS18 和 WIKI 数据集上的对比实验结果。

基于该实验结果,得到如下结论。

(1) 本文方法在 3 个数据集上整体优于其他方

法,验证了其有效性和优越性。具体地,相比于表现最优的内推方法, M_{RR} 平均提升 16.95%, 命中率 Hit@3 平均提升 18.25%, Hit@10 平均提升 20.38%; 相比于表现最优的外推方法, M_{RR} 平均提升 0.75%, 命中率 Hit@3 平均提升 1.63%, Hit@10 平均提升 2.31%。

(2) 从整体表现看,外推方法的性能普遍优于内推方法。这是因为内推方法大多扩展了静态补全方法,建模实体和关系演化的能力有限,并且由于更关注历史事实的查漏补全,缺乏对未来事实的预测能力。

(3) 由于同时建模局部依赖性和全局重复性,本文方法在各项指标上优于其他外推方法。相比于仅捕获全局重复模式的 CyGNet 和 CENET,本方法不仅考虑了全局重复性,还考虑了邻近历史的依赖性,使得性能大幅提升。相比于建模邻近时间历史信息的 RE-NET、xERTE、RE-GCN、TANGO-Tucker、TANGO-DistMult、RE-GCN、DHE-TKG 和 TiPNN 以及捕捉多层次信息的 RLGNet,本文方法在语义、结构和时序特征上的捕捉更加全面,且更关注实体、关系和时间三者的演化建模,因而效果更优。相比于 PPT、ICL、ChapTER 和 HFL 等基于语言模型的方法,本文方法在增强语义的同时,也利用了结构和时序信息提升表示能力,从而在补全时表现出了更佳的性能。

(4) 在不同数据集上,各个方法都表现出了明显差异。多数方法的推理效果在以事件为中心的知识图谱 ICEWS14 和 ICEWS18 上明显不如语义知识图谱 WIKI,这可能是由于以事件为中心的图谱中事实之间存在更复杂的依赖演变模式,实体和关系之间动态交互更复杂,使得建模难度加大。

3.6 消融研究

为了验证模型各个模块的合理性与贡献,本小节进行了消融实验。主要的消融模块包括:演化聚合单元中的结构特征聚合模块(GE)、语义特征聚合模块(SE)和时序特征聚合模块(TE),以及时间增强的 ConvTransE 解码器(TD)和基于课程学习的训练策略(CL)。表 3 展示了 3 个数据集上各模块消融后的性能表现。

表 2 不同方法在 3 个数据集上实体补全任务对比结果
Table 2 Comparison results of entity prediction on different datasets

方法	M_{RR}			命中率 Hits@3			命中率 Hits@10		
	ICEWS14	ICEWS18	WIKI	ICEWS14	ICEWS18	WIKI	ICEWS14	ICEWS18	WIKI
TTransE(2018)	13.72	8.31	29.27	17.70	8.56	34.43	35.74	21.89	42.39
TA-DistMult(2018)	25.80	16.75	44.53	29.74	18.41	48.73	45.99	33.59	51.71
DE-SimplE(2020)	33.36	19.30	45.43	37.15	21.86	47.71	49.82	34.80	49.55
TNTComplEx(2020)	34.05	21.23	45.03	38.50	24.02	49.31	50.92	36.91	52.03
RE-NET(2020)	36.93	28.81	49.66	39.51	32.44	51.19	54.78	47.51	53.48
CyGNet(2021)	35.05	24.93	33.89	39.01	28.28	36.10	53.55	42.61	41.86
xERTE(2021)	40.02	29.31	71.14	44.63	33.51	76.11	56.17	46.48	79.01
TANGO-Tucker(2021)		28.68	50.43		32.17	51.47		47.04	53.35
TANGO-DistMult(2021)		26.65	51.15		30.08	52.16		44.09	53.58
RE-GCN(2021)	40.39	<u>30.58</u>	<u>77.55</u>	44.96	34.34	80.38	59.21	48.75	83.68
CENET(2023)	38.24	27.75	62.67	42.14	32.08	65.31	56.82	46.19	67.90
PPT(2023)	38.42	26.63	53.95	42.50	30.64	57.28	57.01	45.43	60.56
ICL(2023)				46.00	31.30	<u>80.40</u>	56.50	41.40	<u>84.40</u>
ChapTER(2024)	33.80	24.40	53.40	38.00	27.60	57.00	52.70	41.20	65.30
RLGNet(2024)	39.06	29.90	64.34	42.03	33.64	66.71	58.12	49.08	69.51
DHE-TKG(2024)	40.02	29.23	51.20	44.99	33.31	69.25			
TiPNN(2024)	<u>41.20</u>	30.32	73.99	<u>46.23</u>	<u>35.06</u>	76.82	<u>59.60</u>	<u>50.08</u>	80.67
HFL(2025)	37.90	27.10		42.70	30.40		57.30	45.50	
本文方法	42.33	31.56	77.68	47.42	35.79	83.37	62.05	51.67	87.29

注:黑体表示最优结果;下划线表示次优结果。

表 3 本文方法在 3 个数据集上对各模块消融实验结果
Table 3 Ablation study results on different datasets

方法	M_{RR}			命中率 Hits@3			命中率 Hits@10		
	ICEWS14	ICEWS18	WIKI	ICEWS14	ICEWS18	WIKI	ICEWS14	ICEWS18	WIKI
本文方法	42.33	31.56	77.68	47.42	35.79	83.37	62.05	51.67	87.29
-w/o GE	39.32	29.16	68.62	44.19	33.09	74.14	58.76	48.36	82.09
-w/o SE	41.48	31.38	77.50	46.70	35.67	83.02	61.10	51.43	87.42
-w/o TE	39.84	29.52	77.02	44.78	33.59	82.22	60.02	49.95	87.43
-w/o TD	40.56	29.67	75.91	45.48	33.90	81.54	60.11	49.68	86.75
-w/o CL	40.78	30.05	76.57	45.84	34.30	81.93	60.30	50.05	86.66

注:“-w/o”表示去除模块。

如表 3 所示,去除各个模块后,模型的性能均有所下降。在去除演化聚合单元中的结构、语义和时序特征聚合模块后, M_{RR} 在 3 个数据集上平均下降 4.82%、0.40% 和 1.73%。其中,结构特征聚合模块对模型的性能影响最大,其次是时序特征聚合模块,表明这二者涵盖了大量的演化信息,语义特征聚

合模块对历史演化信息的捕捉起到了一定的积极作用。将时间增强的 ConvTransE 解码器替换成线性转移层时, M_{RR} 在 3 个数据集上平均下降 1.81%,说明卷积神经网络在聚合实体、关系和时间特征方面更具优势。当去除课程学习训练策略后, M_{RR} 在 3 个数据集上平均下降 1.39%,表明固定长度的历史

序列限制了模型的演化模式学习,而课程学习策略能够有效地捕捉到不同长度历史序列的演化模式。

3.7 计算复杂度分析

演化聚合单元中,结构、语义和时序特征聚合的时间复杂度分别为 $O(\omega(|E| + |R|))$ 、 $O(n)$ 和 $O(|E| + |R| + 1)$ 。当局部历史序列长度为 k 时,局部性依赖演化编码器的时间复杂度为 $O(k((\omega + 1)(|E| + |R|) + n + 1))$ 。当全局历史序列长度为 T 时,全局重复性演化编码器的时间复杂度为 $O(Td_h)$,其中 d_h 表示实体连接边数平均值。时间增强的 ConvTransE 解码器的时间复杂度为 $O(C + |E|)$,由于实体数远大于卷积核数,可简化为 $O(|E|)$ 。课程学习策略的迭代次数为 g 时,所提方法的整体时间复杂度为 $O(l(k((\omega + 1)(|E| + |R|) + n + 1) + Td_h + |E|))$ 。由于关键参数(如 g, k, ω, d_h)较小,模型整体具有较高的计算效率。

3.8 参数敏感度分析

3.8.1 每种关系的采样事实数

为确定语义特征提取时每种关系的最佳采样数 m ,以 M_{RR} 作为评价指标,在采样事实数占总事实数的百分比为 0% 到 100% 的区间内每隔 5%~10% 选择数据点进行实验。图 3 展示了 ICEWS14 数据集上参数 m 的敏感性分析实验结果。

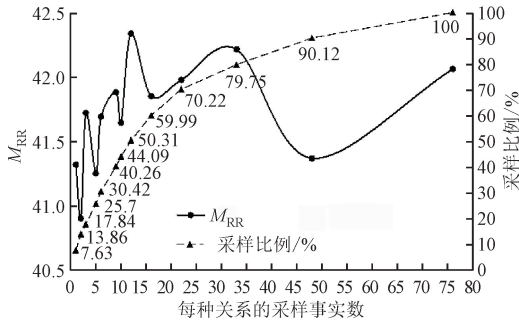


图 3 ICEWS14 数据集参数 m 敏感性分析

Fig. 3 Sensitivity analysis of parameter m on ICEWS14

如图 3 所示,在 ICEWS14 中,当采样比例随采样事实数 m 增长时, M_{RR} 随之波动,但是没有呈现出明显的规律。当采样比例为 50.31% (对应 $m=12$) 时,模型性能最佳。ICEWS18 和 WIKI 数据集上的实验结果也呈现出了类似的波动,没有呈现统一的规律。

3.8.2 课程学习训练参数

课程学习参数主要包含最小历史序列长度 k^1 和最大历史序列长度 k^n ,由于 k^n 显然会限制模型性能上升,通常不对其进行约束,因此主要探索 k^1 对

模型性能的影响。图 4 展示了 ICEWS14 数据集上 k^1 的敏感性分析实验结果。

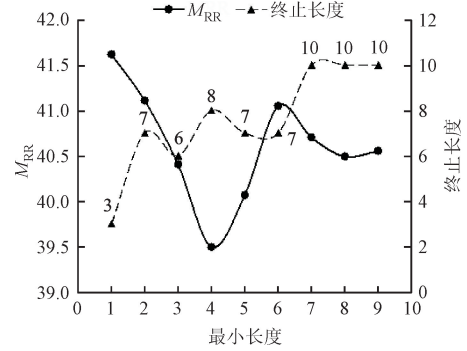


图 4 ICEWS14 数据集最小长度 k^1 敏感性分析结果

Fig. 4 Sensitivity analysis of minimum length k^1 on ICEWS14

由图 4 可以看出,在 ICEWS14 数据集中,当最小长度 k^1 为 1、课程学习自适应训练终止长度达到 3 时, M_{RR} 表现最佳。当 k^1 逐渐增加时, M_{RR} 出现一定波动,但没有呈现出明显规律。ICEWS18 和 WIKI 数据集上的表现也类似,这说明模型对 k^1 有一定的敏感性,但没有明显的变化趋势。

4 结 论

本文提出一种基于消息聚合的演化感知时序知识图谱推理方法,综合建模查询邻近历史图谱中的结构、语义和时序演化依赖信息,并对全局历史图谱中的交互实体和关系进行建模,引入时间增强的 ConvTransE 解码器融合全局和局部的演化历史信息。基于课程学习训练策略引导模型从易到难地捕捉不同长度的历史演化特征,学习演化历史长度的多样性。在 3 个标准数据集上的实验结果表明,所提方法在 M_{RR} 上平均提升了 0.75%,在命中率 Hits@3 上平均提升了 1.63%,在 Hits@10 上平均提升了 2.31%,证明了其优势和有效性。

参考文献:

- [1] SHI Baoxu, WENINGER T. Open-world knowledge graph completion [C]//Proceedings of the 32nd AAAI Conference on Artificial Intelligence and 30th Innovative Applications of Artificial Intelligence Conference and 8th AAAI Symposium on Educational Advances in Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2018: 1957-1964.
- [2] DONG Xin, GABRILOVICH E, HEITZ G, et al. Knowledge vault: a web-scale approach to probabilis-

- tic knowledge fusion [C]//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM, 2014: 601-610.
- [3] BORDES A, USUNIER N, GARCIA-DURÁN A, et al. Translating embeddings for modeling multi-relational data [C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Vancouver, Canada: Neural Information Processing Systems Foundation, 2013: 2787-2795.
- [4] NICKEL M, TRESP V, KRIEDEL H. A three-way model for collective learning on multi-relational data [C]//Proceedings of the 28th International Conference on Machine Learning. Madison, WI, USA: Omnipress, 2011: 809-816.
- [5] YANG B, YIH S, HE X, et al. Embedding entities and relations for learning and inference in knowledge bases [C]//Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015). San Diego, CA, USA: ICLR, 2015: 20193407343615.
- [6] TROUILLON T, WELBL J, RIEDEL S, et al. Complex embeddings for simple link prediction [C]// Proceedings of the 3rd International Conference on Machine Learning. New York, USA: JMLR, 2016: 2071-2080.
- [7] BOSCHÉ E, LAUTENSCHLAGER J, O'BRIEN S, et al. ICEWS coded event data [EB/OL]. (2025-01-01)[2025-01-10]. <http://doi.org/10.7910/DVN/28075>.
- [8] LEETARU K, SCHRODT P. GDELT: global data on events, location, and tone, 1979—2012 [C]//ISA Annual Convention. Los Angeles, CA, USA: ISA, 2013: 1-49.
- [9] LEBLAY J, CHEKOL M W. Deriving validity time in knowledge graph [C]//Companion Proceedings of the Web Conference 2018. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2018: 1771-1776.
- [10] GARCÍA-DURÁN A, DUMANČIĆ S, NIEPERT M. Learning sequence encoders for temporal knowledge graph completion [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA, USA: ACL, 2018: 4816-4821.
- [11] LACROIX T, OBOZINSKI G, USUNIER N. Tensor decompositions for temporal knowledge base completion [C]//Proceedings of the 8th International Conference on Learning Representations. San Diego, CA, USA: ICLR, 2020: 20231313796168.
- [12] JIN W, QU Meng, JIN Xisen, et al. Recurrent event network: autoregressive structure inference over temporal knowledge graphs [C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Stroudsburg, PA, USA: ACL, 2020: 6669-6683.
- [13] LI Zixuan, JIN Xiaolong, LI Wei, et al. Temporal knowledge graph reasoning based on evolutionary representation learning [C]//Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval. New York, USA: ACM, 2021: 408-417.
- [14] LIU Xinyue, ZHANG Jianan, MA Chi, et al. Temporal knowledge graph reasoning with dynamic hypergraph embedding [C]//Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024). Stroudsburg, PA, USA: ACL, 2024: 15742-15751.
- [15] DONG Hao, WANG Pengyang, XIAO Meng, et al. Temporal inductive path neural network for temporal knowledge graph reasoning [J]. Artificial Intelligence, 2024, 329: 104085.
- [16] ZHU Cunchao, CHEN Muhao, FAN Changjun, et al. Learning from history: modeling temporal knowledge graphs with sequential copy-generation networks [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2021: 4732-4740.
- [17] XU Yi, OU Junjie, XU Hui, et al. Temporal knowledge graph reasoning with historical contrastive learning [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2023: 4765-4773.
- [18] LI Yujia, SUN Shiliang, ZHAO Jing. TiRGN: time-guided recurrent graph network with local-global historical patterns for temporal knowledge graph reasoning [C]//Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence. San Mateo, CA, USA: IJCAI, 2022: 2152-2158.
- [19] HE Yongquan, ZHANG Peng, LIU Luchen, et al. HIP network: historical information passing network for extrapolation reasoning on temporal knowledge graph [C]//Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence. San Mateo, CA, USA: IJCAI, 2021: 1915-1921.
- [20] LV Ao, HUANG Yongzhong, OUYANG Guige, et

- al. RLGNet: repeating-local-global history network for temporal knowledge graph reasoning [EB/OL]. (2024-03-31) [2025-1-10]. <https://arxiv.org/abs/2404.00586v1>.
- [21] XU Wenjie, LIU Ben, PENG Miao, et al. Pre-trained language model with prompts for temporal knowledge graph completion [C]//Findings of the Association for Computational Linguistics: ACL 2023. Stroudsburg, PA, USA: ACL, 2023: 7790-7803.
- [22] LEE D H, AHRABIAN K, JIN W, et al. Temporal knowledge graph forecasting without knowledge using in-context learning [C]//Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA, USA: ACL, 2023: 544-557.
- [23] PENG Miao, LIU Ben, XU Wenjie, et al. Deja vu: contrastive historical modeling with prefix-tuning for temporal knowledge graph reasoning [C]//Findings of the Association for Computational Linguistics: NAACL 2024. Stroudsburg, PA, USA: ACL, 2024: 1178-1191.
- [24] XU Wenjie, LIU Ben, PENG Miao, et al. Historical facts learning from long-short terms with language model for temporal knowledge graph reasoning [J]. Information Processing & Management, 2025, 62(3): 104047.
- [25] SHANG Chao, TANG Yun, HUANG Jing, et al. End-to-end structure-aware convolutional networks for knowledge base completion [C]//Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2019: 3060-3067.
- [26] 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. Palo Alto, CA, USA: AAAI Press, 2020: 3988-3995.
- [27] HAN Zhen, CHEN Peng, MA Yunpu, et al. Explainable subgraph reasoning for forecasting on temporal knowledge graphs [C]//Proceedings of the 9th International Conference on Learning Representations. New York, USA: ICLR, 2021: 1-24.
- [28] HAN Zhen, DING Zifeng, MA Yunpu, et al. Learning neural ordinary equations for forecasting future links on temporal knowledge graphs [C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA, USA: ACL, 2021: 8352-8364.

(编辑 武红江)