

融合位置编码的层次图神经网络表示方法

闫晨曦, 武志学

(成都信息工程大学 人工智能学院, 四川 成都 610225)

摘要: 时序知识图谱(TKG)推理旨在通过观察到的事实预测未来时间戳的事实。然而,大多数现有方法忽略了节点的规范位置信息,限制了图神经网络(GNNs)区分同构节点及图对称性等问题的能力。为此,提出一种融合位置编码的层次图神经网络(HGLS-PE)表示方法以提升节点表示能力。具体而言,为了显式关联不同时间戳的实体,首先将TKG转换成全局图,基于构建的全局图模型,从子图级别对每个时间戳内子图中并发出现的事实进行建模,以捕获其潜在的语义关系,从全局图捕获不同时间戳下实体之间的时间依赖性。此外,在子图和全局图两个层级的学习过程中,将位置编码融入节点的消息传播机制中。该方式能够有效区分不同时刻节点和关系对目标节点表示的影响,从而显著增强模型的区分能力和推理能力。在两个公开TKG数据集上进行了广泛实验,结果证明了HGLS-PE的有效性。

关键词: 时序知识图谱;图神经网络;位置编码

DOI: 10.11907/rjdk.242026

中图分类号: TP183

文献标识码: A

文章编号: 1672-7800(2026)004-0132-08

扫描二维码阅读全文:



Hierarchical Graph Neural Network Representation Method with Integrated Positional Encoding

YAN Chenxi, WU Zhixue

(School of Artificial Intelligence, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: Temporal Knowledge Graph (TKG) reasoning aims to predict future facts based on observed ones. However, most existing methods overlook the normalized positional information of nodes, which limits the ability of Graph Neural Networks (GNNs) to address challenges such as isomorphic nodes and graph symmetry. To tackle these issues, we propose a Hierarchical Relational Graph Neural Network with Structural and Positional Encodings (HGLS-PE) to enhance node representation. Specifically, to explicitly associate entities across different time-stamps, we first convert the TKG into a global graph. The proposed model captures the semantic dependencies of concurrent facts within sub-graphs at each timestamp and temporal dependencies between entities across timestamps in the global graph. Furthermore, during the learning process at both subgraph and global graph levels, positional encodings are integrated into the message-passing mechanism of nodes. This integration effectively distinguishes the influence of nodes and relations at different timestamps on the target node representation, significantly improving the model's discriminative and reasoning capabilities. Extensive experiments conducted on two public TKG datasets demonstrate the effectiveness of HGLS-PE.

Key Words: temporal knowledge graph; graph neural networks; positional encoding

0 引言

时序知识图谱(Temporal Knowledge Graph, TKG)作为传统知识图谱(Knowledge Graph, KG)的扩展,旨在捕捉和表示动态演化的知识^[1-2]。与静态 KG 不同,TKG 不仅包含

实体及其关系,而且引入了时间维度,用于刻画实体和关系随时间的变化。传统 KG 已在多个领域展现了其应用价值。文献[3]通过构建药品知识图谱并引入药品间的语义信息,显著提高了推荐精度,并降低了数据稀疏性;文献[4]提出一种基于知识图谱的零样本文档检索伪查询生成方法,通过对伪查询进行增强以引入外部知识,有效提高

收稿日期:2025-03-04

基金项目:成都信息工程大学引进人才科研启动项目(376/376604)

作者简介:闫晨曦(1999-),男,CCF 学生会会员,成都信息工程大学人工智能学院硕士研究生,研究方向为时序知识图谱推理;武志学(1960-),男,博士,成都信息工程大学人工智能学院教授、硕士生导师,研究方向为知识图谱。本文通讯作者:武志学。

了文档检索性能。文献[5]提出一种面向税收条例的知识建模方法,通过定义条例知识子图,设计基于BERT的税收条例知识要素抽取模型,实现了税收条例知识图谱的自动构建,为智慧税务应用提供了重要的语义支撑。TKG推理已广泛应用于时序预测^[6-8]、未来事件推断^[9-11]等实际场景中,因此成为当前研究的重点方向。TKG推理主要分为插值和外推两种方式:插值任务关注在已知时间范围内推断未知事件,外推任务则面向未来事件的预测^[12-13]。本文在外推设置下进行TKG推理研究。为了促进TKG推理研究的发展,本文方法所涉及的代码及数据已公开,研究人员可以通过访问项目页面(<https://github.com/todochenxi/HGLS-PE>)获取相关资源。

目前,一些推理方法取得了显著的研究进展。L2TKG通过结果编码器和潜在关系学习模块挖掘 intra-time 与 inter-time 的潜在关系,有效识别了实体之间的重要关联^[14]。HiSMatch通过建模并发事实之间的结构依赖、历史事实的时间数值信息以及事实之间的时间顺序,全面捕捉两个历史结构中的信息,并通过建模近期背景知识,补充性地捕捉实体之间的高阶关联^[15]。RPC通过关系对应单元和周期对应单元挖掘关系相关性与周期模式背后的信息,有效利用了关系之间的语义相关性和周期性时间模式^[16]。针对现有时序知识图谱推理模型无法充分挖掘并发事实的结构依赖和潜在关系、时间编码方式单一的问题,文献[17]提出一种基于联合卷积的模型,通过联合聚合器的图卷积挖掘节点特征,并结合向量编码和事件属性编码捕获时间信息,增强了模型的时间敏感性。针对现有时序知识图谱推理中外推方法未充分利用时间信息的问题,文献[18]提出将关系嵌入分为静态和动态两部分,并通过头实体嵌入、关系嵌入和所有实体嵌入之间的双线性评分函数,计算得到对象实体的概率,从而预测对象实体。针对传统变分自编码器方法在时序知识图谱推理中易过拟合,且难以捕捉实体语义动态演化的问题,文献[19]提出基于扩散概率分布的模型,通过正向加噪与反向高斯采样联合建模,得到更精确的建模方法。TiPNN从实体无关的视角通过构建历史时序图和查询相关时序路径,实现对时序知识图谱的归纳推理^[20]。CyGNet结合生成模式和复制模式,生成模式在整个实体空间上进行概率预测,以建模潜在分布,同时利用复制模式识别历史事实中的重复模式。然而,其对并发事实的结构依赖性关注不足^[21]。SimFy则通过结构相似性匹配和历史频率学习两个核心模块实现推理,但简化的模型结构难以捕获复杂的时序依赖关系^[22]。RE-NET采用多关系图聚合器结合RNN,通过聚合器捕获图结构信息,同时利用RNN建模时序依赖^[23-24]。RE-GCN通过关系感知图卷积网络捕获演化单元中的结构依赖,并利用门控递归组件实现并行的时序建模,同时引入静态图约束整合实体属性^[25]。然而,该方法对长期历史信息利用不足,更多关注实体的局部结构依赖与静态

特性。

为了更好地学习实体和关系的长短期表示,HGLS提出一种层次关系图神经网络(Graph Neural Networks, GNNs)结构^[26]。该方法在子图级别捕获并发事实之间的语义依赖性,在全局图级别建模实体的时序依赖,并通过门控循环单元(Gated Recurrent Unit, GRU)融合长短期表示^[27]。然而,现有方法普遍忽视了子图结构位置编码(Positional Encoding, PE)的依赖关系,从而限制了图神经网络在区分同构节点或图对称性方面的能力。此外,HGLS构建全局图时仅关注当前时间点之后的时间序列,而忽略了先前的序列信息。这些限制削弱了模型对复杂时序依赖的捕获能力。

为了解决上述问题,本文提出融合位置编码的层次图神经网络(Hierarchical Relational Graph Neural Network with Structural and Positional Encodings, HGLS-PE)。该方法是对HGLS方法的改进,主要包括以下两点:①使用当前时刻之前的KG序列构建全局图;②将位置编码融合到层次图神经网络的消息传播中。改进后的方法基于构建的图关联不同时刻的相关实体,并从子图和全局图两个层次捕获实体的时序依赖关系。在子图级别,学习并发事实间的语义依赖;在全局图级别,显式关联不同时间戳的实体。同时将位置编码融合到子图和全局图中实体的消息传播过程中,强化了模型对节点与关系的事件区分能力。

实验结果表明,HGLS-PE在2个TKG基准数据集上的外推任务中的表现显著优于现有方法,能够更好地捕获实体的时间依赖与语义关系。

1 相关概念

时序知识图谱(TKG)是在传统知识图谱(KG)基础上引入时间维度的扩展形式,其核心是按时间戳组织的KG序列,能够反映实体关系的动态演化。通俗来讲,时序知识图谱在传统的静态知识图谱三元组的基础上增加了时间信息,比如(黑神话悟空,发售,Steam,2024年8月20日)表示在2024年8月20日这个时间点,实体“黑神话悟空”与“Steam”之间具有“发售”的关系。对于时序知识图谱的规范化定义如下:

定义1(时序知识图谱) 时序知识图谱可形式化定义为带有时间戳的有向图 $G = (E, R, T)$ 。其中, E 为实体集合,表示为 $E = \{e_1, e_2, \dots, e_n\}$, 包含 $|E|$ 种不同实体; R 为关系集合,表示为 $R = \{r_1, r_2, \dots, r_m\}$, 包含 $|R|$ 种不同关系; T 为时间戳集合,表示为 $T = \{t_1, t_2, t_3, \dots, t_k\}$, 包含 $|T|$ 种不同时间戳。时序知识图谱也可表示为每个时间戳的静态知识图谱集合,即 $G = (G_1, G_2, G_3, \dots, G_t)$ 。其中,每个时间切片的子图 $G_t = (E, R)$ 表示在时间 t 存在的事实集合。TKG 中一个具体的事实 α 可表示为四元组 (e_s, r, e_o, t) , 其中 $e_s \in E$

为源实体, $r \in R$ 为实体间关系, $e_o \in E$ 为目标实体, $t \in T$ 为事实发生的时间戳。整个 TKG 可表示为一个时间序列 $\{G_t\}_{t=1}^T$, 用于描述实体及其关系在时间上的变化规律。

时序知识图谱上的推理任务主要包括实体推理任务和关系预测任务, 两个推理任务的定义如下:

定义 2 (实体预测) 对于时序知识图谱 $G = (E, R, T)$, 给定实体 $e_s \in E$ (或 $e_o \in E$)、关系 $r \in R$ 和时间 $t + 1 \in T$, 求解缺失实体的过程被称为实体预测, 即 $(?, r, e_o, t + 1)$ 或 $(e_s, r, ?, t + 1)$ 。

定义 3 (关系预测) 对于时序知识图谱 $G = (E, R, T)$, 给定实体 $e_s, e_o \in E$ 、关系 $r \in R$ 和时间 $t + 1 \in T$, 求解实体间关系的过程被称为关系预测, 即 $(e_s, ?, e_o, t + 1)$ 。

2 模型构建方法

2.1 全局图构建

同一时间戳的 KG 序列表现出子图结构依赖性, 之前的方法使用 GCN 获取邻居节点的信息, 然后使用 RNN 递归建模 KG 序列来学习每个时间戳的实体和关系表示, 但该方法并不能显式地关联整个 KG 序列中不同时间间隔的相同实体之间的语义关系。构建全局图旨在显式地关联不同时间间隔的相同实体, 通过两个层级的消息传播, 实体表示融合了局部和全局信息。

不同于 HGLS 考虑当前时刻之后的 KG 序列, 本文方法考虑了当前时刻之前的所有 KG 序列。如图 1 所示, 对于实体预测任务 $(e_s, r, ?, t + 1)$, 图谱构建旨在显式关联时刻 $t + 1$ 前和 e_s 相同的实体。

对于时间集合 T 中的任意时刻 $t + 1 \in T$, 考虑之前的任意两个 KG 序列 G_{t_i} 和 G_{t_j} , 其中 $(0 \leq i < j \leq t)$ 。如果存在共同实体 e , 则在实体 e_{t_i} 和 e_{t_j} 之间添加一条虚拟边 $r_{t_{ij}}$ (见图

1 中的“图谱构建”部分), 使每个实体不仅与同时刻的邻接实体存在结构语义关系, 而且与不同时刻的相同实体存在时间相关关系。

同样的, 也可以使用除当前时刻外的所有 KG 序列来构建全局图。使用当前时刻之前的所有 KG 序列构建全局图记为 HGLS-PE, 使用当前时刻之后的所有 KG 序列构建全局图记为 HGLS-After, 使用除当前时刻外的所有 KG 序列构建全局图记为 HGLS-All, 后文将比较 3 种构建方式对实体预测性能的影响。构建的全局图记为 P 。

2.2 融合位置编码的层次图神经网络

2.2.1 融合位置编码的图神经网络

Transformer 中子图结构位置编码 (PE) 的主要目的是在保留原始词嵌入语义的同时, 为模型提供序列中每个元素的唯一位置向量, 使其能够学习序列依赖关系, 从而有效区分不同位置的相同词, 并捕捉输入序列的结构信息^[28]。构建的全局图 P 中每个时刻的 KG 序列包含来自不同时刻的实体关系。将 PE 同节点特征连接起来有助于目标节点区分不同节点和关系的影响。随机游走位置编码 (Random Walk Positional Encoding, RWPE) 已被证明可以在分子数据集上取得成功, 这得益于其能够有效捕捉图结构中的拓扑信息和节点关系特征, 并提供唯一的节点表示^[29-30]。相比传统位置编码, RWPE 通过随机游走算法提取节点的结构特征, 为图神经网络提供了更加丰富和具有语义意义的位置表示。RWPE 的定义如式 (1) 所示:

$$p_i^{RWSE} = [RW_{ii}, RW_{ii}^2, \dots, RW_{ii}^k] \in R^k \quad (1)$$

其中, $RW = AD^{-1}$ 是随机游走算子, A 是图的邻接矩阵, D 是图的度矩阵。融合位置编码的图神经网络更新方程如式 (2)、式 (3) 所示, 输入特征的初始化方程如式 (4)、式 (5) 所示。

$$h_i^{l+1} = f_h \left(h_i^l, \{h_j^l\}_{j \in N_i}, e_{ij}^l \right), h_i^{l+1}, h_i^l \in R^d \quad (2)$$

$$e_i^{l+1} = f_e \left(h_i^l, h_j^l, e_{ij}^l \right), e_i^{l+1}, e_i^l \in R^d \quad (3)$$

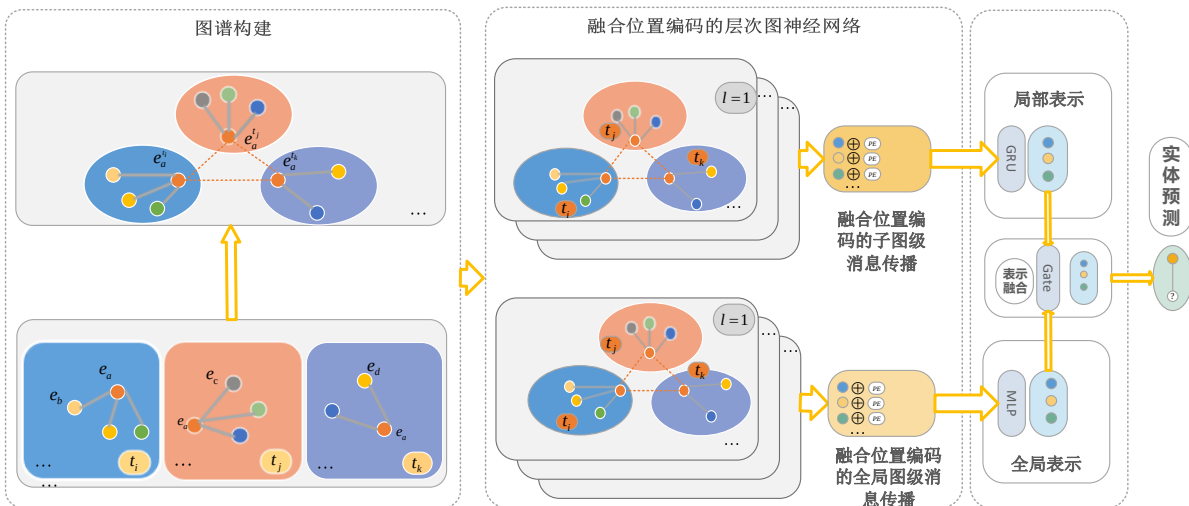


Fig. 1 The structure of HGLS-PE model

图1 HGLS-PE 模型结构

$$h_i^{l=0} = LL_h \left(\begin{bmatrix} h_i^{in} \\ p_i^{in} \end{bmatrix} \right) = D^0 \begin{bmatrix} h_i^{in} \\ p_i^{in} \end{bmatrix} + d^0 \in R^d \quad (4)$$

$$e_{ij}^{l=0} = LL_e \left(e_{ij}^{in} \right) = B^0 e_{ij}^{in} + b^0 \in R^d \quad (5)$$

其中, $p_i^{in} \in R^k$ 是节点 e_i 的输入 PE, $D^0 \in R^{d \times (d_s + k)}$ 、 $d^0 \in R^d$ 、 $B^0 \in R^{d \times (d_s + k)}$ 、 $b^0 \in R^d$ 是可学习的线性变换参数。

2.2.2 子图层

子图层的设计旨在捕获同一时间戳下 KG 序列中并发事实的结构依赖关系,从而为每个子图 G_i^t 中的节点 h_{e_i, t_i} 生成有效的嵌入。这些嵌入能够反映实体间强烈的语义相关性,尤其是在多个事实同时发生时,实体与其邻居之间通常会展现出显著的关联性。通过将节点嵌入与 PE 相结合,可在消息传播过程中增强目标节点对其邻居节点的感知能力,从而更好地捕捉同一时间戳下并发事实之间的结构和语义依赖。

这种结合不仅提升了嵌入的表达能力,而且帮助模型更准确地理解节点在图中的相对位置及其关系。尤其是在多个事实同时发生的情况下,能够有效捕捉节点之间的潜在关联性,从而增强模型的推理能力。

具体来说,本文使用一个关系图卷积神经网络作为语义聚合器来计算子图 G_i^t 中每个节点的嵌入,同时通过 RWPE 计算每个节点的 PE,并将其与来自邻居节点传播的消息进行连接。融合位置编码的子图聚合器定义如(6)所示:

$$h_{s, t_i}^{l+1} = f \left(\frac{1}{|N_{e_s, t_i}|} \sum_{e_o, t_o \in N_{e_s, t_i}} W_1^l \left[\left(h_{o, t_o}^l + x_r \right) \parallel p_{o, t_o}^l \right] + W_2^l h_{s, t_i}^l \right) \quad (6)$$

其中, N_{e_s, t_i} 是节点 e_s^t 在子图 G_i^t 中的邻居集合, $f(\cdot)$ 是激活函数 $RReLU$, $W_1^l, W_2^l \in R^d$ 是用于聚合融合位置编码的邻居信息和第 l 层自环对应的可训练权重矩阵。 $\parallel$ 是连接操作, p_i^l 是源节点 e_o^t 由式(1)计算的局部 PE。将初始化实体嵌入 $h_{s, t_i}^0, h_{o, t_i}^0$ 设置成静态嵌入 x_s 和 x_o 。经过 w 层卷积后,实体表示 h_{s, t_i}^w 捕获了来自邻居节点相对位置的依赖关系, h_{s, t_i}^w 可简写成 $h_{s, t_i} \circ$ 。

2.2.3 全局图层

子图层的建模是为每个实体在其对应的子图中生成局部嵌入,这些嵌入充分考虑了子图结构中事实的结构依赖关系。为进一步捕捉跨时间戳的实体之间的时间相关性,本文基于子图层的输出在全局图中作进一步处理。具体来说,本文将在全局图中执行融合位置编码的消息传播和聚合操作,以捕捉实体之间的时间依赖关系。

全局图中存在不同时间间隔的实体,因此将时间差引入到实体 e_s^t 和 e_o^t 之间关系 r 的表示中,以更好地捕捉跨时间戳的时间依赖性。具体做法是通过计算时间戳之间的差异,将该差异作为附加信息融入关系表示中,如式(7)所示:

$$z_r^{i,j} = x_r + \mathcal{O}(|t_i - t_j|) \quad (7)$$

其中, $|t_i - t_j|$ 是时间差的绝对值。时间编码函数的定义如式(8)所示:

$$\mathcal{O}(t) := \sqrt{\frac{1}{d}} \left[\cos(w_1 t + p_1), \dots, \cos(w_d t + p_d) \right] \quad (8)$$

其中, $w, p \in R^d$ 是可学习的参数向量。

为更好地捕捉相邻节点之间的时间和语义依赖,本文采用注意力机制来评估相邻节点之间的影响力。该影响力的系数计算方式如式(9)所示:

$$a_{s,o}^{i,j} = \frac{\exp \left(g \left(a^T W_3^l \left[z_{s, t_i}^l \parallel z_{o, t_j}^l \parallel z_r^{i,j} \right] \right) \right)}{\sum_{e_o^t \in \tilde{N}_{e_s, t_i}} \exp \left(g \left(a^T W_3^l \left[z_{s, t_i}^l \parallel z_{o, t_j}^l \parallel z_r^{i,j} \right] \right) \right)} \quad (9)$$

其中,实体嵌入的初始化表示 z_{s, t_i}^0 为子图级实体嵌入表示 h_{s, t_i} 的输出, $\tilde{N}_{e_s, t_i}$ 是节点 e_s^t 在全局图中的邻居节点集合, $a \in R^{3d}$ 和 $W_3^l \in R^{3d \times 3d}$ 是每一层的可学习权重参数, $g(\cdot)$ 是激活函数 $LeakyReLU$, $\cdot^T$ 是转置操作, $\parallel$ 是连接操作。

接下来可以通过融合全局 PE 进行消息传播,为全局图中每个实体聚合来自邻居节点的嵌入。该过程的计算方式如式(10)所示:

$$z_{s, t_i}^{l+1} = h \left(\sum_{e_o^t \in \tilde{N}_{e_s, t_i}} a_{s,o}^{i,k} W_4^l \left(\left(z_{o, t_o}^l + z_r^{i,k} \right) \parallel p_{o, t_o}^l \right) + W_5^l z_{s, t_i}^l \right) \quad (10)$$

其中, $h(\cdot)$ 是激活函数 $ReLU$, W_4^l, W_5^l 是每一层中聚合和自环的权重参数矩阵, p_{o, t_o}^l 是源节点 z_{o, t_o}^l 根据式(1)计算得到的全局位置编码。在全局图上经过 β 层的卷积操作可获得每个实体在全局图上的输出 z_{s, t_i}^β , 可简写成 $z_{s, t_i} \circ$ 。

2.3 层次图神经网络表示

2.3.1 子图层表示

子图层表示反映了在子图结构中实体和关系的局部依赖。为了捕获实体的局部表示,本文使用门控循环单元(GRU),根据子图层的输出对每个实体最近 τ 个时间戳的表示进行编码。每个实体的局部表示计算公式如式(11)所示:

$$e_{s, t+1}^S = GRU_E(e_{s, t}^S, h_{s, t}) \quad (11)$$

其中,所有实体共享 GRU_E 的参数, $h_{s, t}$ 是节点 e_s^t 在子图层的输出。类似地,关系的局部表示计算公式如(12)所示:

$$e_{r, t+1}^S = GRU_R(e_{r, t}^S, h_{r, t}) \quad (12)$$

其中, $h_{r, t}$ 是关系在时刻 t 的表示。通过对交互的实体表示进行平均池化得到 $h_{r, t}$, 计算公式如(13)所示:

$$h_{r, t} = \text{Meanpooling}(h_{s, t}), \forall e_s \in N_r^t \quad (13)$$

其中, N_r^t 是关系 r 在时刻 t 相关联的实体集合。

2.3.2 全局图表示

全局图表示反映了实体和关系在全局图中的长时间依赖性,本文将全局图的输出经过一个非线性转换获得实体的全局表示,计算公式如式(14)所示:

$$e_{s,t+1}^L = \sigma(W_6 z_{s,t} + b) \quad (14)$$

其中, $\sigma(\cdot)$ 是 tanh 函数, $W_6 \in R^{d \times d}$ 是权重矩阵, $b \in R^d$ 是偏置向量。关系的全局表示相对稳定, 本文用其静态嵌入作为全局表示, 计算公式如式(15)所示:

$$e_{r,t+1}^L = x_r, r \in R \quad (15)$$

2.3.3 层次图神经网络表示融合

为了获得实体和关系在子图与全局图上的统一表示, 本文采用可学习的门控函数来融合两个层级的嵌入。实体和关系表示的计算公式如式(16)、式(17)所示:

$$e_{s,t+1} = \sigma(g_e) \odot h_{s,t+1}^L + (1 - \sigma(g_e)) \odot h_{s,t+1}^S \quad (16)$$

$$e_{r,t+1} = \sigma(g_r) \odot h_{r,t+1}^L + (1 - \sigma(g_r)) \odot h_{r,t+1}^S \quad (17)$$

其中, $g_e, g_r \in R^d$ 是融合每个实体和关系层级表示的门控向量参数; $\sigma(\cdot)$ 是 Sigmoid 函数, 将变量映射到 $[0, 1]$; $\odot$ 表示逐元素乘法。

2.4 参数学习

本节概述了参数学习的核心步骤, 包括如何通过 ConvTransE 解码器计算实体和关系的交互概率, 以及实体预测和关系预测任务的学习目标。目标函数整合了两个子任务的损失, 并加入正则化项, 以确保模型的鲁棒性。该设计能够有效地捕捉动态表示, 并优化模型的整体性能。

2.4.1 分数函数

本文使用 ConvTransE 作为解码器, 用于计算在时刻 $t+1$ 且关系为 r 条件下实体 e_s 和 e_o 的交互概率^[31]。公式如式(18)所示:

$$p_{t+1}(ols, r, G_{<t+1}) = \sigma(e_{r,t+1} \text{ConvTransE}(e_{s,t+1}, e_{o,t+1})) \quad (18)$$

其中, $\sigma(\cdot)$ 是 Sigmoid 函数, $e_{s,t+1}, e_{r,t+1}$ 和 $e_{o,t+1}$ 是融合局部与全局信息的动态表示。

2.4.2 学习目标

本文考虑了实体预测和关系预测, 两个学习任务的计算公式如式(19)、式(20)所示:

$$L_e = -\sum_{t=0}^T \sum_{(e_s, r, e_o, t+1) \in G_{t+1}} \log p_{t+1}(ols, r, G_{<t+1}) \quad (19)$$

$$L_r = -\sum_{t=0}^T \sum_{(e_s, r, e_o, t+1) \in G_{t+1}} \log p_{t+1}(rls, o, G_{<t+1}) \quad (20)$$

目标函数计算公式如式(21)所示:

$$L = \lambda_1 L_e + (1 - \lambda_1) L_r + \lambda_2 \|\theta\|_2 \quad (21)$$

其中, λ_1 是调节两个子任务权重的超参数, $\|\cdot\|_2$ 是 L_2 范数, λ_2 是控制正则化强度的超参数。

3 实验与结果分析

通过在两个典型的时序知识图谱数据集上进行实验, 以评估本文提出的模型。目标是通过实验回答以下问题:

Q1: HGLS-PE 和现有预测方法在实体预测任务上的性能比较。

Q2: 3 种构建全局图的方式对 HGLS-PE 性能的影响。

Q3: 位置编码对 HGLS-PE 性能的影响。

Q4: HGLS-PE 对不同超参数的敏感性。

3.1 实验设置

3.1.1 实验环境

本文使用的 GPU 为 3090-24GB, 数量为 1; CPU 为 12th Gen Intel(R) Core(TM) i5-12600K, PyTorch 版本为 2.4.1, CUDA 版本为 12.1, DGL 版本为 2.4.0, Python 版本为 3.10。实验主要的超参数设置如表 1 所示。

Table 1 Main hyperparameter setting of the experiment

表 1 实验主要超参数设置

超参数	中文名称	数值
dropout	随机失活	0.2
learning_rate	学习率	0.001
epoch	训练轮次	30
batch_size	训练批次	1
embed_dim	嵌入维度	200

3.1.2 数据集

本文使用两个典型的 TKG 数据集进行实验, 包括 ICEWS14 和 ICEWS18, 记录了 2014 年和 2018 年的全球事件。按照时间顺序以 80%、10%、10% 的比例划分训练集、验证集和测试集。数据集的详细信息如表 2 所示。

Table 2 Dataset information

表 2 数据集信息

数据集	实体数	关系数	训练集	验证集	测试集	时间间隔/h
ICEWS14	6 869	230	74 845	8 514	7 371	24
ICEWS18	23 033	256	373 018	45 995	49 545	24

3.1.3 基线模型与评估指标

本文将 HGLS-PE 模型在两个数据集上的实体预测性能与先前提出的 TKG 推理模型进行比较, 包括 CyGNet、RE-Net、RE-GCN、TITer^[32]、SiMFy、HGLS。对于 SiMFy、RE-GCN 和 HGLS, 本文使用其开源代码和默认参数设置在两个数据集上重新运行模型, 其他模型的结果引用自文献[16]。实验中采用广泛使用的指标平均倒数排名 (Mean Reciprocal Rank, MRR) 和命中率 HITs@{1, 10} 在 raw 设置下评估模型在两个数据集上的实体预测性能。

3.2 性能比较

所有模型在实体预测任务上的指标如表 3 所示, 可以看到 HGLS-PE 在两个数据集的 3 个指标全面优于其他 TKG 模型, 验证了模型的有效性。具体来说, CyGNet 忽略了并发事实的结构依赖; RE-NET 只考虑了目标实体的一些历史交互进行预测; RE-GCN 忽略了跨时间戳的事件依赖性; SiMFy 结构简单, 无法捕捉复杂的时序依赖; TITer 局限于路径结构, 难以捕获全局长期依赖; 尽管 HGLS 能够从短期和长期建模实体与关系的表示, 但构建的图仅考虑了当前时间点之后的 KG 序列, 而忽略了先前的依赖关系, 并且在两个层次的消息传播时忽略了源节点的相对位置依赖。

3.3 消融实验

为了研究不同构图方式(Q2)和位置编码(Q3)对模型实体预测性能的影响, 本文比较了 HGLS-PE 的不同变体

Table 3 Performance comparison of the models on two datasets

表 3 模型在两个数据集上的性能比较

Model	14-MRR	14-Hit@1	14-Hit@10	18-MRR	18-Hit@1	18-Hit@10
CyGNet	36.51	27.42	54.44	26.82	17.13	45.72
RE-NET	38.91	29.32	57.51	28.42	18.41	47.92
RE-GCN	39.48	29.52	58.64	29.11	19.11	48.64
SiMFy	39.54	29.56	59.18	28.65	18.99	47.62
TITer	41.18	32.13	57.94	28.09	21.13	44.01
HGLS	47.00	35.06	70.41	29.32	19.21	49.83
HGLS-PE	49.38	37.17	73.20	33.41	21.62	56.84

在 MRR 指标上的性能表现。其中,HGLS-All 构建全图时考虑除当前时间点外的所有序列,HGLS-After 只考虑当前时间点之后的序列,HGLS-NoPE 不使用位置编码的模型变体。变体模型的消融实验结果如表 4 所示。具体来说,HGLS-After 只考虑当前时间点之后的序列,但先前的实体关系被忽略,无法捕捉事件演变的依赖;HGLS-All 虽然考虑了全局序列,但当前时间点之后的序列可能引入不相关的噪声数据;HGLS-PE 仅考虑当前时间点之前的序列,可以有效捕捉当前时间点之前的事件依赖关系;HGLS-NoPE 在两个层级的消息传播中不使用位置编码,因此实

体在聚合时无法区分邻居信息和全局信息,从而降低了全局信息对节点表示的影响。

Table 4 The MRR metric results of entity prediction tasks in ablation experiments

表 4 实体预测任务在消融实验中的 MRR 指标结果 (%)

Model	ICEWS14	ICEWS18
HGLS-All	48.84	32.35
HGLS-After	48.56	32.12
HGLS-NoPE	48.79	32.46
HGLS-PE	49.38	33.41

3.4 敏感性分析

为了进一步探索 HGLS-PE 对重要超参数的敏感性(Q4),本文研究了位置编码(PE)长度对 HGLS-PE 性能的影响。通过调整位置编码长度可以分析该超参数在不同长度下对模型性能的具体影响,进而评估模型对位置编码长度变化的适应能力和优化效果。

因此,本文在 ICEWS14 和 ICEWS8 两个数据集上使用不同长度的位置编码进行了实验。图 2 展示了局部 PE 的实验结果,图 3 展示了全局 PE 的实验结果。调整一位 PE 的长度时,另一位 PE 保持最优配置。

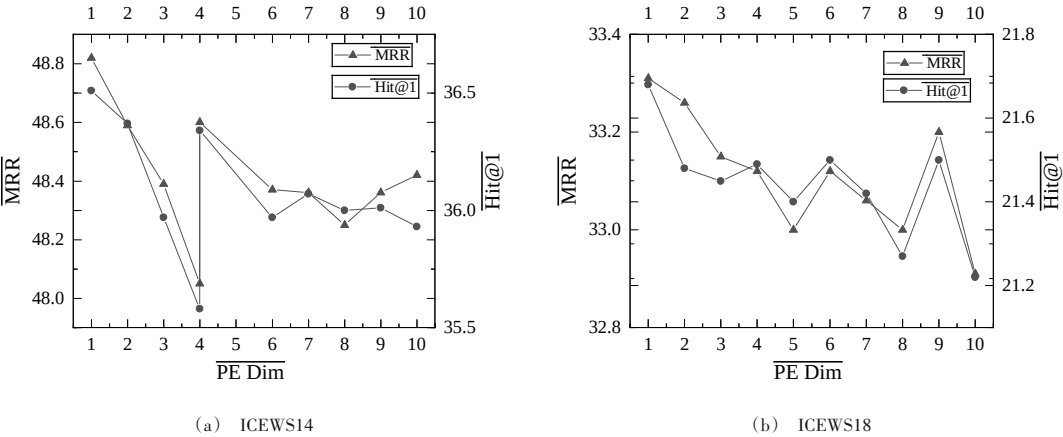


Fig. 2 Performance of HGLS-PE with different lengths of local PE

图 2 HGLS-PE 在不同局部 PE 长度下的性能

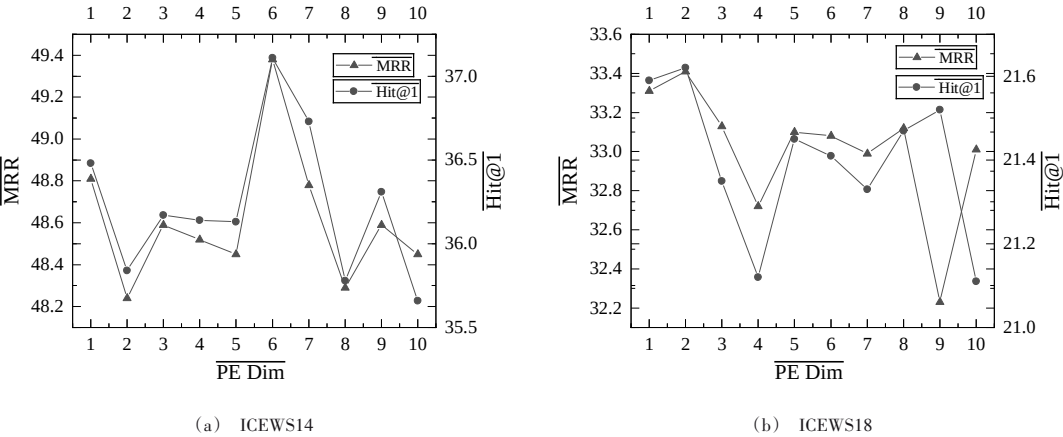


Fig. 3 Performance of HGLS-PE with different lengths of global PE

图 3 HGLS-PE 在不同全局 PE 长度下的性能

实验结果证明,局部 PE 在捕捉短期依赖时具有明显优势。较短的局部 PE 能够更高效地学习时间接近的事件之间的依赖关系,特别是在短期内事件时间间隔较小的情况下。此时,子图层能更精确地捕捉这些细微的依赖,从而提高了模型的预测精度。

与局部 PE 相比,全局 PE 则更适合学习长期依赖关系。长时间跨度的 PE 能够帮助模型理解历史事件与当前事件之间的深层关联,尤其是当事件之间时间间隔较长时。较长的全局 PE 可以有效捕捉这些长期演化的模式和关系,从而增强对复杂依赖关系的建模能力。

此外,实验结果还证明,模型在 ICEWS18 数据集上展现出良好的鲁棒性,进一步验证了 HGLS-PE 在处理不同时间依赖特性和数据集时的适应能力与稳定性。

4 结语

本文提出一种融合位置编码的层次图神经网络表示方法,通过构建全局图显式关联当前时间点之前的 KG 序列,有效捕捉了事件的依赖关系。此外,本文将位置编码融入关系感知图卷积神经网络的消息传播过程中,大幅增强了目标节点对邻居节点的区分能力,从而进一步提升了目标节点嵌入的表示能力。

在两个基准数据集上的实验结果表明,本文提出的图构建方式以及位置编码在 TKG 实体预测任务中具有显著效果。消融实验进一步证明了各个模块的有效性,敏感性分析则探讨了不同长度的位置编码对模型性能的影响。

然而,当前的研究仅考虑了源节点的位置编码以及由子图计算的位置编码。未来的研究将重点探索由整个时序知识图谱计算的位置编码、由源节点和目标节点计算的相对位置编码以及不同位置编码的融合。此外,更好的位置编码方式也是下一步研究的方向。

参考文献:

- [1] CAI B, XIANG Y, GAO L X, et al. Temporal knowledge graph completion: a survey [C]//Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, 2023: 6545–6553.
- [2] JI S X, PAN S R, CAMBRIA E, et al. A survey on knowledge graphs: representation, acquisition, and applications [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(2): 494–514.
- [3] ZHANG W Z, SHEN Y G, LI Y. Drug recommendation algorithm integrating knowledge graph and collaborative filtering [J]. Software Guide, 2023, 22(8): 42–47.
张文卓, 申艳光, 李焰. 融合知识图谱与协同过滤的药品推荐算法 [J]. 软件导刊, 2023, 22(8): 42–47.
- [4] LIU J P, SUN Y G, ZHU Q, et al. Knowledge graph aware pseudo query generation for zero-shot document retrieval [J]. Software Guide, 2024, 23(11): 47–52.
刘军平, 孙医贵, 朱强, 等. 基于知识图谱的零样本文档检索伪查询生成 [J]. 软件导刊, 2024, 23(11): 47–52.
- [5] ZOU A Q, CHEN Y P. Knowledge graph construction method for tax regulations [J]. Software Guide, 2023, 22(4): 32–37.
邹安琪, 陈艳平. 面向税收条例的知识图谱构建方法 [J]. 软件导刊, 2023, 22(4): 32–37.
- [6] YANG F Q, ZHANG Y, ZHAO X C, et al. Unleashing the power of decoders: temporal knowledge graph extrapolation with householder transformation [J]. Symmetry, 2024, 16(9): 1166.
- [7] LIU X L, LIU S Q. Improved temporal knowledge graph prediction model based on RE-GCN [J]. Journal of Heilongjiang University of Science and Technology, 2024, 34(5): 812–816.
刘兴丽, 柳始群. 基于 RE-GCN 改进的时序知识图谱预测模型 [J]. 黑龙江科技大学学报, 2024, 34(5): 812–816.
- [8] CHEN J, ZHAO X C, SUI J Y, et al. Story-inspired large language models for temporal knowledge graph prediction [J]. Pattern Recognition and Artificial Intelligence, 2024, 37(8): 715–728.
陈娟, 赵新潮, 隋京言, 等. 故事启发大语言模型的时序知识图谱预测 [J]. 模式识别与人工智能, 2024, 37(8): 715–728.
- [9] LI Z X, GUAN S P, JIN X L, et al. Two-stage temporal knowledge graph inference based on multi-historical sequence joint evolution modeling [J]. Journal of Chinese Information Processing, 2024, 38(2): 46–53.
李紫宣, 官赛萍, 靳小龙, 等. 基于多历史序列联合演化建模的两阶段时序知识图谱推理 [J]. 中文信息学报, 2024, 38(2): 46–53.
- [10] LI Z X, GUAN S P, JIN X L, et al. Complex evolutionary pattern learning for temporal knowledge graph reasoning [C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 2022: 290–296.
- [11] XU Y, OU J J, XU H, et al. Temporal knowledge graph reasoning with historical contrastive learning [C]//Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence, 2023: 4765–4773.
- [12] SHEN Y H, JIANG X H, WANG Y Z, et al. A survey on inference in temporal knowledge graphs [J]. Chinese Journal of Computers, 2023, 46(6): 1272–1301.
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.
- [13] 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.
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] ZHANG M Q, XIA Y W, 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, 2023: 12617–12631.
- [15] LIU Z X, HOU Z N, GUAN S P, et al. HiSMATCH: historical structure matching based temporal knowledge graph reasoning [C]//Findings of the Association for Computational Linguistics: EMNLP 2022, 2022: 7328–7338.
- [16] LIANG K, MENG L Y, LIU M, et al. Learn from relational correlations and periodic events for temporal knowledge graph reasoning [C]//Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2023: 1559–1568.
- [17] ZHANG C K, MA H D. Temporal knowledge graph inference based on joint convolution [J]. Computer Engineering and Applications, 2025, 61(18): 166–174.
张成坤, 马汉达. 基于联合卷积的时序知识图谱推理 [J]. 计算机工程

- 与应用, 2025, 61(18): 166–174.
- [18] LIU W, XIE L J, ZHANG Z H, et al. Temporal knowledge graph inference based on tensor decomposition embedding[J]. Journal of Beijing Information Science & Technology University, 2024, 39(1): 49–54.
刘伟, 谢璐钧, 张智慧, 等. 基于张量分解嵌入的时序知识图谱推理[J]. 北京信息科技大学学报, 2024, 39(1): 49–54.
- [19] ZHOU G Y, LI P F, XIE P H, et al. Temporal knowledge graph inference based on diffusion probability distribution[J]. Journal of Software, 2024, 35(11): 5083–5097.
周光有, 李鹏飞, 谢鹏辉, 等. 基于扩散概率分布的时序知识图谱推理[J]. 软件学报, 2024, 35(11): 5083–5097.
- [20] DONG H, WANG P Y, XIAO M, et al. Temporal inductive path neural network for temporal knowledge graph reasoning[J]. Artificial Intelligence, 2024, 329: 104085.
- [21] ZHU C C, CHEN M H, FAN C J, 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.
- [22] LIU Z T, TAN L, LI M F, et al. SiMFy: a simple yet effective approach for temporal knowledge graph reasoning[C]//Findings of the Association for Computational Linguistics: EMNLP 2023, 2023: 3825–3836.
- [23] JIN W, QU M, JIN X S, 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, 2020: 6669–6683.
- [24] ELMAN J L. Finding structure in time[J]. Cognitive Science, 1990, 14(2): 179–211.
- [25] LI Z X, JIN X L, LI W, 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, 2021: 408–417.
- [26] ZHANG M Q, XIA Y W, LIU Q, et al. Learning long-and short-term representations for temporal knowledge graph reasoning[C]//Proceedings of the ACM Web Conference 2023, 2023: 2412–2422.
- [27] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[DB/OL]. <https://arxiv.org/pdf/1412.3555>.
- [28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 6000–6010.
- [29] DWIVEDI V P, LUU A T, LAURENT T, et al. Graph neural networks with learnable structural and positional representations[DB/OL]. <https://arxiv.org/pdf/2110.07875>.
- [30] RAMPÁŠEK L, GALKIN M, DWIVEDI V P, et al. Recipe for a general, powerful, scalable graph transformer[C]//Proceedings of the 36th International Conference on Neural Information Processing Systems, 2022: 14501–14515.
- [31] SHANG C, TANG Y, HUANG J, 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, 2019: 3060–3067.
- [32] SUN H H, ZHONG J L, MA Y P, et al. TimeTraveler: reinforcement learning for temporal knowledge graph forecasting[C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 2021: 8306–8319.

(责任编辑: 黄健)