

融合时间编码和时间分布注意力的推理模型

董文永^{1,2}, 梁智学², 周孟强², 贾亚洁³

(1. 新疆政法学院 信息网络安全学院, 新疆 图木舒克 843900;

2. 武汉大学 计算机学院, 湖北 武汉 430072;

3. 武汉大学 国家网络安全学院, 湖北 武汉 430072)

摘要:为了充分利用时序知识图谱数据的时间维度特征,解决时序知识图谱时间信息挖掘不充分的问题,提出了一种融合时间编码和时间分布注意力的推理模型 TETDM,该模型结合长期历史信息 and 短期历史信息来捕捉不同模式的历史事件对推理目标的影响。首先,模型将较长一段历史信息中对应实体或关系出现的频率作为推理模型的长期约束条件。其次,模型通过时间感知编码和时间分布注意力机制对短期历史信息进行编码,深入挖掘时序知识图谱数据的时间维度特征。其中,时间感知编码器能够精准捕捉时间信息对实体和关系的影响,从而生成带有时间特征的实体和关系表示。时间分布注意力机制通过在不同子图中建模每个重复事实的注意力,而不是仅仅学习它们的表示,来学习历史重复事件的变量分布。通过编码时间特征对实体和关系的影响以及编码不同时间戳下历史信息的贡献程度,可以获得在时间维度上更为精准的实体和关系的嵌入表示,从而提高模型的推理能力。在 ICEWS14、ICEWS05-15、ICEWS18 以及 GDELT 数据集上的结果进一步证实,TETDM 提出的时间编码模块和时间分布注意力模块均可以提升模型的推理性能。

关键词:时序知识图谱;图卷积神经网络;注意力机制;时间编码;知识推理

中图分类号:TP391.1

文献标识码:A

文章编号:1673-629X(2025)06-0182-07

doi:10.20165/j.cnki.ISSN1673-629X.2025.0014

A Reasoning Model Integrating Time Encoding and Time Distribution Attention

DONG Wen-yong^{1,2}, LIANG Zhi-xue², ZHOU Meng-qiang², JIA Ya-jie³

(1. School of Information Network Security, Xinjiang University of Political Science and Law, Tumushuke 843900, China;

2. School of Computer Science, Wuhan University, Wuhan 430072, China;

3. School of Cyber Science and Engineering, Wuhan University, Wuhan 430072, China)

Abstract: To fully leverage the temporal characteristics of temporal knowledge graph (TKG) data and address the issue of insufficient utilization of temporal information, we propose a reasoning model named TETDM (Temporal Encoding and Temporal Distribution Attention-based Reasoning Model), which integrates long-term and short-term historical information to capture the impact of different historical event patterns on inference targets. Specifically, the model incorporates the frequency of entities or relations over an extended historical period as a long-term constraint for inference. Furthermore, it encodes short-term historical information using a temporal-aware encoder and a temporal distribution attention mechanism to deeply explore the temporal characteristics of TKG data. The temporal-aware encoder accurately captures the influence of temporal information on entities and relations, generating temporal feature-enriched representations of entities and relations. The temporal distribution attention mechanism, instead of merely learning representations of repeated facts, models the attention of each repeated fact within different subgraphs to learn the variable distribution of historical repetitive events. By encoding the influence of temporal features on entities and relations, as well as the contribution of historical information across different timestamps, the model produces more precise temporal embeddings of entities and relations, thereby enhancing its reasoning capability. Experimental results on ICEWS14, ICEWS05-15, ICEWS18, and GDELT datasets further demonstrate that both the temporal encoding module and the temporal distribution attention module proposed by TETDM can improve the reasoning performance of the model.

Key words: temporal knowledge graphs; graph convolutional neural networks; attention mechanism; time encoding; knowledge reasoning

0 引言

时序知识图谱推理的原理是根据现有的时间信息和图结构信息预测新的事实^[1-3]。具体来说,准确预测未来的事实需要根据实体和关系的历史互动来推断它们表现出的行为模式和偏好^[4]。为了充分利用历史事实信息,研究者们分别对时序知识图谱的图结构和时间特征进行建模,该文致力于解决时序知识图谱时间信息挖掘不充分的问题。

现有的时序知识图谱推理方法在时间信息挖掘方面存在局限性。首先,大多数方法未能深入建模时间与实体和关系变化的具体关系,通常仅将时间视为附加特征融入实体和关系的表示中,忽略了时间如何直接或间接影响它们的动态语义特性。其次,大多数模型依赖循环神经网络(RNN)捕捉知识图谱整体的时间演化规律,然而这种方法容易受到短期记忆能力的限制,难以有效捕获长时间跨度的事件依赖性。此外,现有模型通常对不同时刻的知识图谱快照一视同仁,未能区分各时间点的信息对推理目标的实际贡献。这些局限性导致现有方法难以充分利用时序知识图谱的时间维度特征,从而限制了推理性能的提升。Deng等人^[5]利用基于组合的多关系图卷积网络^[6](CompGCN)捕捉邻近实体和事件类型的影响,并结合门控循环单元(GRU)来建模表示时间依赖性。RE-GCN^[7]使用RGCN^[8]来聚合邻近实体的消息,并利用自回归GRU来建模事件之间的时间依赖性。CyGNet^[9]通过复制生成机制来捕捉事实的全局重复频率。TiRGN^[10]结合了局部和全局的历史信息,通过带有双重循环机制的图神经网络(GNN)编码器来捕捉历史事实的序列性、重复性和周期性模式。HiSMATCH^[11]通过CompGCN形式的背景知识编码器将背景知识集成到TKGR模型中,补充捕捉实体之间的高阶关联。TANGO^[12]扩展了神经微分方程来建模时序知识图谱,将时间和结构信息编码进连续时间动态嵌入表示中。HGLS^[13]将时序知识图谱(TKG)序列转化为全局图,以明确关联不同时间步的历史实体。TPmod^[14]定义了关系的“优良度”和实体对的“亲密度”,利用注意力机制来聚合与实体相关的历史事件。EvoKG^[15]通过同时建模事件时间和网络结构的演化来捕捉时序知识图谱中的结构和时间动态。为了探索图随时间演化的本质,EvoExplore^[16]分别描述了图结构的形成过程和图从局部结构和全局结构的动态拓扑变换。HSAE^[17]和rGalT^[18]采用注意力机制来捕获实体和关系的结构信息。现有的全局历史编码器将所有历史时间戳的相关事实考虑在内,以避免在相邻时间戳中未出现的实体或关系丢失。这种处理方式存在潜在的问题,当全局历史信息较多时会使得全局编

码器部分的计算量很大。

在时序知识图谱中,实体及其关系的特征随着时间的推移而不断变化,同时时间对不同实体和关系的特征影响不同。受此启发,提出了TETDM模型,在HGLS^[13]模型利用长短期历史信息完成推理任务的启发下,引入了长期历史信息推理模块和短期历史信息推理模块。短期历史信息推理模块旨在利用时间轴上推理目标附近的历史信息;长期历史信息推理模块旨在利用时间轴上距离推理目标较远的历史信息。此外,受CyGNet^[9]的启发,TETDM认为过去经常发生的事件在将来有较大概率会再次发生,在基于事件存在性的全局信息编码器的基础上提出了基于事件发生频率的长期历史信息推理模块。该文的贡献如下:

(1)提出了时间编码模块来获得在特定时间下更准确的实体和关系的嵌入表示,在时间编码模块中利用时间感知机制对不同时间戳对不同实体和关系的影响进行编码。

(2)在时间序列上引入注意力机制来衡量不同时间戳对当前推理任务的影响程度,通过自主学习将适当的注意力分配给不同时间戳下的历史子图表示。

(3)通过对比实验和消融实验展示了该方法的优越性。

1 网络模型

1.1 模型架构

TETDM的结构如图1所示,其主要包括两部分:一是图1的左半部分即短期历史信息推理模块,二是图1的右半部分即长期历史信息推理模块。

短期历史信息推理模块的组成部分为时间编码模块、实体与关系聚合模块、历史信息聚合模块以及时间分布注意力模块。其中,时间编码模块和时间分布注意力模块是主要改进之处。具体来说,实体与关系聚合模块以及历史信息聚合模块都用了局部信息编码器,用于对每个时间戳下的历史子图中实体和关系进行编码。时间编码模块的作用则是将时间信息融入实体与关系的表示中。经过时间编码模块,可以获得包含时间信息的实体与关系嵌入表示。随后,实体聚合器和关系聚合器进一步提炼实体与关系的结构特征,形成富含语义信息的嵌入表示。在形成含有时间信息和语义信息的实体和关系表示后,使用门控循环单元(GRU)来编码相邻时间戳内相同实体或关系所包含的历史信息。经过时间信息学习模块与实体、关系聚合模块以及GRU网络的联合作用,可以获得 m 个时间戳下子图的实体嵌入和关系嵌入。通过这种方式,短期历史信息推理模块能够充分利用历史数据中的时间信息和图结构信息,为预测任务提供更为准确和全

面的支持。

长期历史信息推理模块旨在利用推理时刻前 l 个时间戳所积累的历史信息,深入统计分析目标实体 s 周围满足特定条件 $(s, r, ?)$ 的实体 o 的历史数据。通过挖掘这些历史数据中实体 o 出现的频率,该模块

能够预测当前时刻可能出现实体 o 的概率,这一过程可以增强预测任务的准确性和可靠性。

长期历史信息推理模块和短期历史信息推理模块分别得到候选实体的长短期预测概率,最后通过一个可变因子来衡量两者对推理目标的影响。

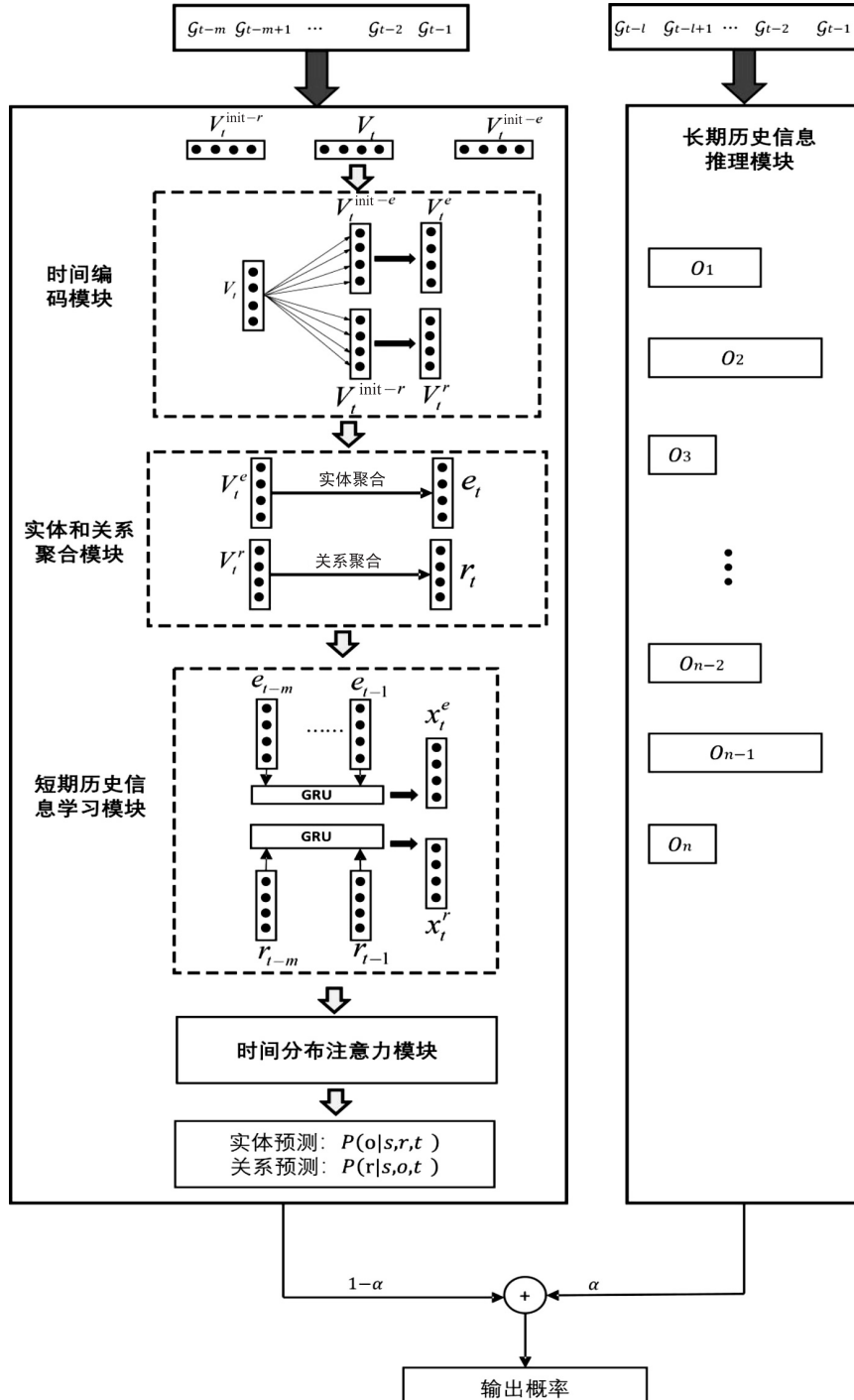


图 1 融合时间编码和时间分布注意力的推理模型

1.2 时间编码模块

在 TKG 中,实体及其关系的特征随着时间的推移而不断变化,并且时间对不同实体和关系的不同特征的影响也是不同的。受此启发,在时间编码模块中利用时间感知机制来对不同时间戳对不同实体和关系的

影响进行编码。对于时间戳 t 下的任何实体,其在特定时间的嵌入向量表示由两部分组成:一个随机初始化向量 V_i^{init-e} 及其相对于时间的变化 $\Delta V_i^e \in \mathbb{R}^d$ 。

$$V_i^e = V_i^{init-e} + \Delta V_i^e \tag{1}$$

$$\Delta V_i^e = f(I_i^e \odot (V_i^{init-e} * \Delta t)) \tag{2}$$

其中: $V_t^e \in \mathbb{R}^d$ 是在时间信息学习单元之后的时间戳 t 处一个实体的时间特定嵌入向量; $V_t^{\text{init}-e} \in \mathbb{R}^d$ 是实体 e 在时间戳 t 处随机初始化的嵌入向量; Δt 是标量, 并且 Δt 等于时间戳除以时间间隔加 1; f 是激活函数, 这里是 ReLU. $I_t^e \in \mathbb{R}^d$ 表示时间戳 t 对实体的影响因子向量, 该向量代表时间戳 t 对此实体的每个特征的影响; $*$ 和 \odot 分别表示向量乘法和阿达玛乘积 Hadamard product.

$$I_t^e = \text{softmax}(W_e \cdot V_t) \quad (3)$$

其中: $V_t \in \mathbb{R}^d$ 是时间戳 t 的随机初始化嵌入向量; $W_e \in \mathbb{R}^{d \times d}$ 是一个可学习的权重矩阵, 用于对 V_t 做线性变换, \cdot 代表矩阵乘法.

上述操作利用时间对一个实体的不同特征的影响进行编码。为了进一步编码时间对不同实体的影响, 引入了可学习的权重向量 $\omega_e \in \mathbb{R}^d$, 然后可以得到时间戳 t 下所有实体的嵌入矩阵。

$$E_t^{\text{ia}} = f(\omega_e \odot (I_t^e \odot (E_t * t))) \quad (4)$$

其中, $E_t \in \mathbb{R}^{N_e \times d}$ 是初始实体嵌入矩阵, $E_t^{\text{ia}} \in \mathbb{R}^{N_e \times d}$ 表示实体在时间戳 t 的输出嵌入矩阵, N_e 是实体的数量。

对于时间戳 t 处的关系, 执行相同的操作, 可以得到包含时间信息的关系嵌入表示。

$$I_t^r = \text{softmax}(W_r \cdot V_t) \quad (5)$$

$$R_t^{\text{ia}} = f(\omega_r \odot (I_t^r \odot (R_t * t))) \quad (6)$$

其中: $I_t^r \in \mathbb{R}^d$ 表示时间戳 t 对关系的影响因子向量, $W_r \in \mathbb{R}^{d \times d}$ 是一个可学习的权重矩阵, 用于对 V_t 做线性变换; $R_t \in \mathbb{R}^{N_r \times d}$ 是初始关系嵌入矩阵; $R_t^{\text{ia}} \in \mathbb{R}^{N_r \times d}$ 表示关系在时间戳 t 的输出嵌入矩阵, N_r 是关系的数量。

经过时间编码模块得到实体和关系的嵌入矩阵后, 将其作为到实体和关系信息聚合器的输入, 再进一步挖掘实体和关系的结构和语义信息。

1.3 时间分布注意力模块

该模块的目的是基于历史子图的分布表示来学习其在推理过程中的不同作用。在获得具有 m 长度历史的时间子图的分布式嵌入后, TETDM 通过注意力机制学习各种时间戳在推理任务中的不同作用, 图 2 展示了时间分布注意力模块的结构。对于查询 $(s, r, ?, t)$, 首先分别通过嵌入矩阵 E_t 和 R_t 获得其主体和关系在第 t 个时间戳的嵌入 x_t^e 和 x_t^r 。然后, 根据 Conv-TransE 的原理拼接 x_t^e 和 x_t^r , 并将它们输入到一维卷积网络中, 输入的通道数为 2, 卷积核的大小设置为 3×2 , 通道的数量设置为 50。

$$p = \text{Conv1D}([x_t^e, x_t^r]) \quad (7)$$

$$y_t = W_1 p + b_1 \quad (8)$$

其中, y_t 表示预测任务在特定历史时间戳 t 的中间输

出, W_1 和 b_1 表示可学习参数。

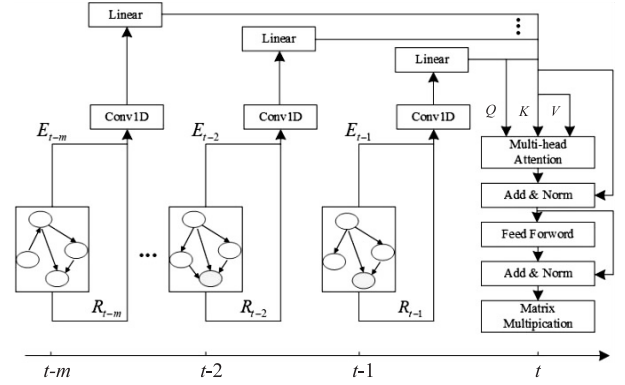


图 2 时间分布注意力网络结构

因此, 对于查询 $(s, r, ?, t)$, 在 m 长度的历史子图序列中, 预测的分布中间结果表示为 $\{y_{t-m}; \dots; y_{t-2}; y_{t-1}\}$ 。考虑到最新的历史时间戳在预测中起着最大的作用, 使 y_{t-1} 对所有历史时间戳施加可学习的注意力, 包括对其自身的注意力。然后, 通过 y_{t-1} 生成查询向量 Q , 通过 $\{y_{t-m}; \dots; y_{t-2}; y_{t-1}\}$ 生成键值向量 K 和 V 。

$$Q = W_q y_{t-1} \quad (9)$$

$$K = W_k [y_{t-m}; \dots; y_{t-2}; y_{t-1}] \quad (10)$$

$$V = W_v [y_{t-m}; \dots; y_{t-2}; y_{t-1}] \quad (11)$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (12)$$

其中, $Q \in \mathbb{R}^{d_q}$, $K \in \mathbb{R}^{d_k}$, $V \in \mathbb{R}^{d_v}$, $\sqrt{d_k}$ 是用于防止消失梯度问题的特定于问题的缩放因子。 W_q , W_k 和 W_v 是将可学习注意力权重分配给每个历史时间戳的可学习参数。为了预测未来事件, 矩阵 W_k 和 W_v 为每个历史事实分配了唯一的系数。因此, 通过学习, TETDM 可以为不同的历史重复事实分配不同的注意力权重。

为了提高模型学习能力, 从多个角度考虑历史重复事实的重要性, TETDM 在时间注意力层引入了多头注意力机制, 这里的“头”表示矩阵 W_q , W_k 和 W_v 的数量。接着, 使用前馈神经网络来引入深层语义信息。

$$g = \text{FFN}(z) = W_2(\text{ReLU}(W_3 z)) \quad (13)$$

其中, $z \in \mathbb{R}^d$ 是多头注意力机制的输出, $W_2 \in \mathbb{R}^{d \times d_{\text{hid}}}$ 和 $W_3 \in \mathbb{R}^{d \times d_{\text{hid}}}$ 是可学习的参数, d_{hid} 是前馈神经网络中隐藏单元的数量。

最后对输出 g 和实体在最新的历史时间戳的嵌入表示 E_{t-1} 执行矩阵乘法 (Matrix Multiplication, MM)。

$$p^{\text{short}} = \text{mm}(g, E_{t-1}) \quad (14)$$

其中, $g \in \mathbb{R}^d$ 以及 $E_{t-1} \in \mathbb{R}^{N_e \times d}$, p^{short} 是该模块最终得到的候选实体的概率分布。

1.4 长期历史信息推理模块

该模块旨在从频率的角度补充较远历史信息的影响。如图 3 所示, 对于推理任务 $(s, r, ?, t)$, 首先从不

同时间戳下的知识图谱快照中寻找满足条件 $(s, r, ?)$ 的客体。然后,使用一个 N 维的编码向量 $\mathbf{v}_i^{(s,r)}$, 记录在时间戳 t 的子图中满足条件 $(s, r, ?)$ 的对应实体。在时刻 t 的子图中,如果目标主体 s 与客体 o_i 之间存在关系 r ,则在向量 $\mathbf{v}_i^{(s,r)}$ 中将实体 o_i 对应元素置为 1;如果客体 o_i 与目标主体 s 之间不存在关系 r ,则将实体 o_i 在向量 $\mathbf{v}_i^{(s,r)}$ 中置为 0。TETDM 定义 $V_t^{(s,r)}$ 表示在 t 个时间步上满足 $(s, r, ?)$ 条件的实体 o 的统计信息,因此,提取的相关事实的历史频率信息可以表示如下:

$$V_t^{(s,r)} = V_{t-l}^{(s,r)} + V_{t-l+1}^{(s,r)} + \dots + V_{t-1}^{(s,r)} \quad (15)$$

为了最大限度地减少一些在历史上不与 s 和 r 形成已知事实的实体的概率,在使用统计到的频率信息 $V_t^{(s,r)}$ 之前,对其进行一定的预处理,具体做法是将 $V_t^{(s,r)}$ 中值为 0 的原始设置为一个较小的值,这里将其置为 -50。

最后,向量 $V_t^{(s,r)}$ 经过 Softmax 函数,可以得到候选实体对应的概率。以上推理过程可以表示如下:

$$p^{\text{long}} = \text{softmax}(V_t^{(s,r)}) \quad (16)$$

在分别得到推理目标的长期历史信息模块和短期历史信息模块的概率后,设置一个可变因子 α 来调整

两者比例以得到候选实体集中所有实体的最终概率。

$$p^{\text{final}} = (1 - \alpha)p^{\text{long}} + \alpha p^{\text{short}} \quad (17)$$

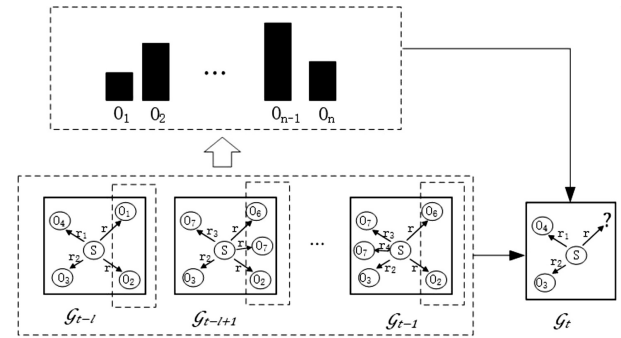


图3 基于事件频率的长期历史信息推理模块

2 实验设置

本节所有实验是在 Ubuntu 18.04 LTS 操作系统、Pytorch1.10 深度学习框架、DGL0.8.0 图机器学习库、python3.8 编程语言下,使用 RTX3090 显卡、Inter Core i9-11900K 的硬件环境下进行的。

实验在 4 个公开数据集上进行,分别是 ICEWS14、ICEWS05-15、ICEWS18 和 GDELT,表 1 展示了数据集的统计信息。

表 1 数据集信息统计

名称	实体数	关系数	训练集	验证集	测试集	时间戳数量	时间间隔
ICEWS14	7 128	230	63 685	13 823	13 222	365	1 天
ICEWS05-15	10 288	251	386 962	46 092	46 275	4 017	1 天
ICEWS18	23 033	256	373 018	45 995	49 545	7 272	15 分钟
GDELT	7 691	240	1 033 270	238 765	305 241	8 925	15 分钟

该文使用时序知识图谱推理模型常用的评价指标:平均倒数排名(MRR)与命中率(Hits@n)。

$$V_{\text{MRR}} = \frac{1}{|S|} \sum_{i=1}^{|S|} \frac{1}{V_{\text{rank}_i}} \quad (18)$$

$$V_{\text{Hits@n}} = \frac{1}{|S|} \sum_{i=1}^{|S|} \prod_{i=1}^n (V_{\text{rank}_i} \leq n) \quad (19)$$

其中, V_{MRR} 、 $V_{\text{Hits@n}}$ 和 V_{rank_i} 分别表示 MRR、Hits@n 和 rank_i 。

本节中 Hits@1、Hits@3、Hits@10 表示在推理中预测正确结果排名在前 1、3、10 中的比例。MRR 与 Hits@n 指标越大,表明模型的推理能力越强,且该文

对于 MRR 与 Hits@n 的结果展示是将真实值乘以 100 得到的。

3 实验与结果分析

3.1 关系预测结果分析

表 2 和表 3 展示了 ICEWS14、ICEWS05-15、ICEWS18 和 GDELT 数据集上的实体预测相关的实验结果。其中,最佳结果使用黑体加粗展示,次优结果使用下划线展示。值得注意的是,CluSTeR 和 HGLS 模型在原论文中没有进行 Hits@3 实体预测结果的评价,因此表格中没有相应结果。

表 2 ICEWS14 和 ICEWS05-15 数据集上实体预测结果

模型	ICEWS14				ICEWS05-15			
	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@3	Hits@10
CyGNet ^[9]	34.68	25.35	38.88	53.16	35.46	25.44	40.20	54.47
RE-GCN ^[7]	35.92	26.06	41.03	54.79	38.07	27.44	43.18	58.63
CluSTeR ^[19]	46.00	33.80	-	71.20	44.60	34.90	-	63.00
TiRGN ^[10]	44.04	33.83	48.95	63.84	50.03	39.25	56.13	70.71
HisMatch ^[11]	46.42	35.91	51.63	66.84	<u>52.85</u>	<u>42.01</u>	<u>59.05</u>	73.28

续表 2

模型	ICEWS14				ICEWS05-15			
	MRR	Hits@ 1	Hits@ 3	Hits@ 10	MRR	Hits@ 1	Hits@ 3	Hits@ 10
HGLS ^[13]	47.00	<u>35.06</u>	-	<u>70.41</u>	46.21	35.32	-	67.12
TETDM	<u>46.62</u>	37.67	<u>51.06</u>	68.82	53.26	43.52	59.90	<u>72.34</u>

表 3 ICEWS18 和 GDELT 数据集上实体预测结果

模型	ICEWS18				GDELT			
	MRR	Hits@ 1	Hits@ 3	Hits@ 10	MRR	Hits@ 1	Hits@ 3	Hits@ 10
CyGNet ^[9]	24.98	15.54	28.58	43.54	20.22	12.15	21.66	35.82
RE-GCN ^[7]	27.23	17.63	30.76	46.25	19.69	12.46	20.93	33.81
CluSTeR ^[19]	32.30	20.60	-	<u>55.90</u>	18.30	11.60	-	33.50
TiRGN ^[10]	33.66	23.19	<u>37.99</u>	54.22	21.67	13.63	23.27	<u>37.60</u>
HisMatch ^[11]	<u>33.99</u>	<u>23.91</u>	37.90	53.94	<u>22.01</u>	<u>14.47</u>	<u>23.80</u>	36.61
HGLS ^[13]	29.32	19.21	-	49.83	19.04	11.79	-	32.34
TETDM	36.08	26.16	39.87	57.18	23.47	17.13	28.32	40.90

实验结果表明,与其他时序知识图谱外推方法相比,TETDM 模型在所有评估指标都表现为最优或次优。具体来说,TETDM 在 ICEWS18 和 GDELT 数据集上的所有评估指标均优于其它对比模型,这可能是因为 TETDM 在实体和关系的嵌入表示中不仅考虑了图拓扑结构信息和时间的顺序性,还考虑了时间的具体影响,这在一定程度上展现了 TETDM 在推理任务上的优越性。

与近期较先进模型 HisMatch 和 HGLS 相比,TETDM 数据集 ICEWS14 和 ICEWS05-15 上的某些评估指标表现次优,这可能因为在长期历史信息的利用上,TETDM 只捕捉了事件频率信息,而 HisMatch 和 HGLS 使用了长期历史事件的并发结构信息以及事件时间顺序信息。更进一步地,对于评估指标 MRR,与模型 HisMatch 相比,TETDM 在 ICEWS14、ICEWS05-15、ICEWS18 和 GDELT 数据集上分别提高了 0.2、0.41、2.09 和 1.46,这在一定程度上表明时间编码模

块和时间分布注意力模块可以提升模型的推理能力。此外,还可以发现该模型在 ICEWS14 和 ICEWS05-15 数据集上的结果优于在 GDELT 和 ICEWS18 数据集上的结果。这可能是因为 ICEWS18 和 GDELT 数据集中实体和关系之间存在更复杂的动态交互,而这种复杂的动态事实关联难以准确建模。值得注意的是,从复杂度角度看,TETDM 的时间和空间复杂度与主流方法类似,但由于支持全局时序依赖建模,比 RE-GCN、CyGNet 等方法更高,但仍低于 HisMatch。

3.2 消融实验与结果分析

为了验证文中方法在时序知识图谱推理任务上的有效性,本节进行了实体预测任务的消融研究。为了分别验证时间编码模块和时间分布注意力模块的有效性,将原模型中去除时间编码模块的模型记为 TDM,将原模型中去除时间分布注意力模块的模型记为 TEM。分别在四个数据集上进行实验,所有结果通过平均排名倒数(MRR)进行评估。

表 4 ICEWS14 和 ICEWS05-15 数据集上消融实验结果

模型	ICEWS14				ICEWS05-15			
	MRR	Hits@ 1	Hits@ 3	Hits@ 10	MRR	Hits@ 1	Hits@ 3	Hits@ 10
TDM	46.39	36.70	50.39	67.37	53.05	42.37	58.12	72.08
TEM	45.56	36.18	50.17	66.29	52.91	41.98	58.26	71.96
TETDM	46.62	37.67	51.06	68.82	53.26	43.52	59.90	72.34

表 5 ICEWS18 和 GDELT 数据集上消融实验结果

模型	ICEWS18				GDELT			
	MRR	Hits@ 1	Hits@ 3	Hits@ 10	MRR	Hits@ 1	Hits@ 3	Hits@ 10
TDM	35.91	26.06	39.60	57.08	23.16	16.68	27.74	40.78
TEM	35.13	25.43	38.97	56.17	22.73	17.11	26.85	40.67
TETDM	36.08	26.16	39.87	57.18	23.47	17.13	28.32	40.90

根据表 4 和表 5 的实验数据可以发现,在移除时间编码模块后,ICEWS14、ICEWS05-15、ICEWS18 以及 GDELT 数据集上的推理结果的准确性均有所下降。这一结果有力地证明了,在实体结构聚合与关系语义聚合之前为实体和关系的嵌入表示增添时间编码模块,能够量化出实体和关系随着时间推移发生的特征变化,进而提升时序知识推理的准确性。进一步地,当移除时间分布注意力模块后,这四个数据集中的评价指标同样呈现下降态势。这表明,在时序知识图谱推理任务中,学习历史子图的分布表示及其在推理过程中的不同作用同样至关重要。消融实验的结果进一步证实,TETDM 提出的时间编码模块和时间分布注意力模块均可以提升模型的推理性能。

4 结束语

为了解决时序知识图谱推理模型对于 TKG 数据时间信息利用不充分的问题,提出了融合时间编码和时间分布注意力的推理模型 TETDM。通过提出时间编码模块来获得在特定时间下更准确的实体和关系的嵌入表示,在时间编码模块中利用时间感知机制来对不同时间戳对不同实体和关系的影响进行编码。在时间序列上引入注意力机制衡量不同时间戳对当前推理任务的影响程度,通过自主学习将适当的注意力分配给不同时间戳下的历史子图表示,通过对比实验和消融实验证明了该方法的优越性。

参考文献:

- [1] 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.
- [2] 孙水发, 李小龙, 李伟生, 等. 图神经网络应用于知识图谱推理的研究综述[J]. 计算机科学与探索, 2023, 17(1): 27-52.
- [3] 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.
- [4] LI Y, ZHANG X, ZHANG B, et al. SANe: space adaptation network for temporal knowledge graph completion[J]. Information Sciences, 2024, 667: 120430.
- [5] DENG S, RANGWALA H, NING Y. Dynamic knowledge graph based multi-event forecasting[C]//Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining. Virtual Event: ACM, 2020.
- [6] VASHISHTH S, SANYAL S, NITIN V, et al. Composition-based multi-relational graph convolutional networks[C]//Proceedings of the 8th international conference on learning representations. Addis Ababa: ICLR, 2020.
- [7] LI Z, JIN X, 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. Virtual Event: ACM, 2021.
- [8] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks[C]//15th international conference, ESWC 2018. Heraklion: Springer, 2018.
- [9] ZHU C, CHEN M, FAN C, et al. Learning from history: modeling temporal knowledge graphs with sequential copy generation networks[C]//Association for advancement of artificial intelligence 2021. [s.l.]: AAAI, 2021.
- [10] 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. Vienna: IJCAI, 2022.
- [11] LI Z, HOU Z, GUAN S, et al. Hismatch: historical structure matching based temporal knowledge graph reasoning[J]. arXiv:2210.09708, 2022.
- [12] HAN Z, DING Z, MA Y, 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. Punta Cana: EMNLP, 2021: 8352-8364.
- [13] ZHANG M, XIA Y, LIU Q, et al. Learning long-and short-term representations for temporal knowledge graph reasoning[C]//Proceedings of the ACM web conference. Austin: ACM, 2023.
- [14] BAI L, MA X, ZHANG M, et al. Tpmode: a tendency-guided prediction model for temporal knowledge graph completion[J]. ACM Transactions on Knowledge Discovery from Data, 2021, 15(3): 1-17.
- [15] 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 fifteenth ACM international conference on web search and data mining. Virtual Event: ACM, 2022.
- [16] KELLY J, JACOB C. EvoExplore: multiscale visualization of evolutionary histories in virtual reality[C]//Computational intelligence in music, sound, art and design. Parma: Springer, 2018.
- [17] REN X, BAI L, XIAO Q, et al. Hierarchical self-attention embedding for temporal knowledge graph completion[C]//The ACM web conference. Austin: ACM, 2023.
- [18] GAO Y, FENG L, KAN Z, et al. Modeling precursors for temporal knowledge graph reasoning via auto-encoder structure[C]//International joint conference on artificial intelligence. Vienna: IJCAI, 2022.
- [19] LI Z, JIN X, GUAN S, et al. Search from history and reason for future: two-stage reasoning on temporal knowledge graphs[J]. arXiv:2106.00327, 2021.