

基于混合注意力机制的时间旋转知识图谱补全*

王璐璐

(大连大学 信息工程学院, 辽宁 大连 116000)

摘要: 针对现有时序知识图谱补全中捕捉动态关系模式, 处理非对称、临时和自反关系方面的局限性, 提出了一种新颖的融合混合注意力机制与时间旋转的模型。一方面, 通过引入时间旋转, 利用复数空间中的向量来表示随时间演化的实体与关系, 特别是处理时间区间内的关系变化, 采用双复数嵌入方案显著增强了对时态特性的表达能力; 另一方面, 通过对知识图谱引入空间注意力和通道注意力两个维度分析, 能够更好地聚焦于时序序列中对预测最为关键的实体和关系特征, 从而在复杂的时间序列中挖掘时序关联信息。通过在 ICEWS14、ICEWS18、YAGO11k 和 WIKI12k 数据集上的实验评估, 模型在 MRR、Hits@1、Hits@3 和 Hits@10 上普遍优于基线模型, 体现出算法的优越性和强鲁棒性。

关键词: 时序知识图谱; 时间旋转; 混合注意力机制; 链接预测

中图分类号: TP391.1

文献标识码: A

DOI: 10.19358/j.issn.2097-1788.2024.10.007

引用格式: 王璐璐. 基于混合注意力机制的时间旋转知识图谱补全 [J]. 网络安全与数据治理, 2024, 43(10): 42-48.

Temporal rotation knowledge graph completion based on hybrid attention mechanism

Wang Lulu

(School of Information Engineering, Dalian University, Dalian 116000, China)

Abstract: A novel model that integrates a hybrid attention mechanism with temporal rotation is proposed to address the limitations of capturing dynamic relation patterns, handling asymmetric, temporary, and reflexive relations in existing temporal knowledge graph completion. On one hand, by introducing temporal rotation, we leverage vectors in complex spaces to represent entities and relations evolving over time, especially to handle relation changes within temporal intervals. The adoption of a dual-complex embedding scheme significantly enhances the expressive power for temporal characteristics. On the other hand, by introducing spatial attention and channel attention to analyze the knowledge graph from two dimensions, the model can better focus on the most crucial entity and relation features in the temporal sequence for prediction, thus mining temporal correlation information from complex time series. Through experimental evaluations on the ICEWS14, ICEWS18, YAGO11k, and WIKI12k datasets, the model outperforms baseline models in terms of MRR, Hits@1, Hits@3, and Hits@10, demonstrating the superiority and strong robustness of the proposed algorithm.

Key words: temporal knowledge graph; temporal rotation; hybrid attention mechanism; link prediction

0 引言

随着人工智能时代的到来, 互联网中的数据规模也随之爆发式增长, 这些大量的数据中往往包含很多有价值的信息, 而知识图谱^[1]作为一种新的组织和存储海量数据的技术, 受到越来越多的关注。知识图谱早期主要用于搜索结果优化, 随后在系统推荐^[2]、智能问答^[3]、

知识搜索^[4]等领域发挥了重要作用。知识图谱通常是静态的, 以其当前状态存储事实。知识图谱以三元组的形式表示事实: $(s; r; o)$, 然而, 现有知识图谱大都采用半自动化结合人工的方式构建, 具有不完整性^[5]。并且在现实中, 实体之间的关系经常随着时间而变化, 即知识具有时效性。为此, 时序知识图谱被引入^[6], 时序知识图谱将时间事实表示为四元组: $(s; r; o; t)$, 即通过扩展时间为 t 的静态三元组, 描述此事实发生在时间为 t 时有

* 基金项目: 辽宁省科学研究项目 (LJKZ1180)

效。时序知识图谱可以不断提供和完善时间维度的知识，更具有研究价值。

近年来，尽管静态知识图谱补全技术取得了显著进展，如基于翻译的模型（如 TransE^[7]）、基于图神经网络的方法（如 RE-GCN^[8]）等，但这些方法大多忽视了时间维度对于关系动态变化的深刻影响。为应对这一挑战，时序知识图谱补全（Temporal Knowledge Graph Completion, TKGC）应运而生，旨在结合时间序列分析与深度学习技术，挖掘时序模式，捕捉动态关系的演化规律。现有的时序知识图谱补全模型大多数是 TransE 和 DistMult 的扩展，例如 TTransE^[9]、TA-DisMult^[10]、ChronoR^[11] 等，因此不能完全表达如今的某些关系模式。最新的模型还存在计算资源的分配问题。

本文提出一种新的时序知识图谱补全模型 YiTX，该模型创新性地融合了时间旋转嵌入与混合注意力机制，旨在更精准地建模实体间随时间变化的关系，并有效预测图谱中缺失的实体与关系。混合注意力机制在聚合信息上的优势和注意力权重的分配，在面对数据的多样性和复杂性时，仍能实现高效学习和快速推理。通过在四个公开时序数据集上的广泛实验，验证了所提模型在链接预测任务上的优越性能。本研究不仅为时序知识图谱的补全提供了一种新的视角，也为智能信息处理、文本挖掘和信息抽取技术的发展提供了新思路，进而促进知识图谱在推荐系统、智能问答、事件预测等领域的广泛应用。

1 相关工作

1.1 静态知识图谱补全方法

静态知识图谱补全模型是指针对已构建完成且不随时间变化的知识图谱进行的缺失信息预测或填充的过程。现有的静态知识图谱补全模型主要以 TransE 法为核心，针对空间映射等场景做的改进。而通过简单的向量加法或减法来建模实体间的关系，这在处理复杂关系模式（如对称性、反对称性等）时可能不够灵活。RotatE^[12] 模型由 Sun 等提出，其创新之处在于它将实体和关系映射到复数向量空间，并利用复数的旋转属性来建模关系，从而能够更自然和精确地表达这些复杂关系模式。

RotatE 模型通过其独特的复数旋转机制，不仅在理论上提供了新的视角来理解知识图谱中的关系，也在实践应用中展现了显著的性能提升，成为知识图谱补全领域的重要方法之一。

1.2 动态知识图谱补全方法

动态知识图谱补全方法可分为两大类：基于符号逻辑的方法和基于知识表示学习的方法。

1.2.1 基于符号逻辑的方法

CTDNE^[13] 开创性地引入了时态路径的概念，强调在构建路径时需遵循时间序列的约束条件，确保所采样路径蕴含时序依赖性，从而精准反映知识图谱随时间的演变逻辑。Liu^[14] 等人通过 TLogic 模型专注于多跳关系推理中的时间维度，创新性地提出了循环时态逻辑规则，增强了推理的时序合理性。T-EDGE^[15] 针对带权图谱的特性，设计了一套新颖的时序约束采样策略，因而被赋予更高的采样概率，促进了时序信息的有效捕获。FIGT-NE^[16] 模型则通过时间增强的随机游走策略，动态调整实体的采样概率，依据其已被访问的频次，有效缓解了采样偏差，更全面地揭示了时态知识图谱的发展规律。ATGEB^[17] 则认识到实体活跃度的非均匀分布特性，首先依据实体活跃时间进行排序，并采用 DBSCAN 聚类算法提升了对时序知识图谱动态性的理解与建模精度。

以上方法从本体或规则出发，可解释性较强，但计算效率低下，面对日益增长的知识数量难以满足要求，且推理过程中计算太过复杂。

1.2.2 基于知识表示学习的方法

基于知识表示学习的 TKGE 模型大多是由 TransE 和 DistMult 扩展而来的，其中，Duration-HyTE^[18] 在 HyTE 的基础上进一步区分了持续型关系和瞬发型关系，并对知识的有效时间进行建模，引入了知识有效可信度的概念。ATiSE^[19] 模型则采用加性时间序列分解方法，将时序知识图谱的嵌入表示映射到多维高斯分布的空间中，以提高模型的鲁棒性。

除了基于 TransE 的扩展模型外，研究者还在 DistMult 的基础上进行了时序扩展。文献 [20] 中提出的 DE-SimpleE 模型通过一种历时实体嵌入函数来学习时间信息，这种方法能够更精细地捕捉实体在不同时间点的特征。另外，TNT-omplex^[21] 模型使用正则化组件对静态模型 Complex 进行了扩展，将时态知识图谱表示为一个 4 阶张量，这种方法将时态信息直接融入张量表示中，提高了对时序知识的建模能力。

除此之外，基于先前的 RotatE 模型，ChronoR 应运而生，通过采用高维度旋转操作来编码实体、关系及其随时间的变化，有效捕捉了时序数据中时间维度与多重关系间的复杂交互。DBKGE^[22] 模型则另辟蹊径，利用动态贝叶斯框架赋予实体以高斯分布表示，在数据流的持续输入中动态更新，展现出对未来状态的强大预测能力。TBDRI^[13] 聚焦于逆关系的重要性，引入基于关系交互的块分解策略来独立优化逆关系的学习，深化了对关系复杂性的把握。HTTR^[23] 通过提出户主转换的概念，以正交矩阵的形式表达了实体间关系的时间演变过程，为关系

动态性的建模提供了一种新思路。TASTER^[24]模型独创性地运用稀疏变换矩阵来揭示实体随时间的演化路径,同时兼顾了局部与全局信息的建模,有效提升了模型的可扩展性。TGeomE^[25]则将时序知识图谱映射到超复数(四元数)空间,巧妙地将时间戳嵌入关系之中,定义了随时间变化的关系嵌入表达,并利用四元数的几何乘积和 clifford 共轭运算来计算关系得分,展现了独特的数学美感与高效性。BiQCap^[26]模型通过在双四元数空间中结合欧氏旋转和双曲旋转来表达关系,同时将实体在不同时间点的表示视为平移,为实体的动态特性提供了全新的几何解释,进一步丰富了时序知识图谱补全的技术版图。

与 TransE 和 DistMult 类似,尽管各种各样的知识图谱补全模型已被开发来学习实体之间的多关系交互,但它们都存在推断临时关系的问题,这些临时关系仅在某个时间点有效或持续某个时间段。

2 基于混合注意力机制的时间旋转知识图谱补全模型

2.1 模型整体架构

本文提出了如图 1 所示的时序图谱补全模型架构。首先利用时间旋转模块,将实体的时间演化过程视为复数空间中从开始时间到当前时间的旋转,将时序图谱中的四元组映射至复数空间,可以捕捉实体和关系随时间演化的复杂模式。其次在模型中设计了 CBAM 混合注意力机制模块,能够从多维度识别特征图中实体和空间的结构与时间信息,捕捉如时间顺序和时间间隔等细粒度的时间信息。对实体嵌入进行加权求和,得到上下文敏感表示。最后,设计用于衡量四元组可信度的得分函数,并使用负采样损失函数对模型进行优化训练。

2.1.1 时间旋转模块

为克服静态嵌入模型在处理时间动态性方面的局限,根据文献[27],本文融入时间维度于嵌入模型训练过程,并将实体随时间的演变在复数空间中形象化为从初始时刻至当前时刻的向量旋转。对于包含时间区间的事实,采用双复数嵌入策略,分别对应关系的起点和终点,使得模型能够处理时序数据中的非均匀时间间隔、速度差异或非线性时间变化。时间旋转模型能够捕捉实体和关系随时间演化的复杂模式,增强对复杂事件动态的建模能力,提高了时序知识图谱补全的准确性。具体而言,原本静态的头实体 h 和尾实体 t 的向量表示,在纳入时间考量后,需将其映射成复数空间中向量 h_i 和 t_i 。映射函数如下:

$$h_i = h \circ \tau, \quad t_i = t \circ \tau \quad (1)$$

其中, $\circ$ 表示厄米点积,将每个时间步 $\tau_i \in \mathbb{C}^d$ 的模约束为 1。利用欧拉公式 $e^{i\theta} = \cos\theta + i\sin\theta$,可将此映射过程视作实体嵌入向量在复平面上绕原点逆时针旋转一个角度 $\theta_{\tau,i}$,此旋转操作本质上仅调节了复数向量的相位,而对实体在实数嵌入空间中的位置信息给予了动态的时间依赖性表达。通过这种方式,模型能适应不同实体和关系随时间推移的不同速率变化,有效对齐时间序列,便于进一步分析。

2.1.2 混合注意模块

在时间旋转的时序知识图谱补全模型中,引入混合注意力机制可以帮助模型更好地捕捉和利用时序信息,以及不同实体和关系之间的相关性。本文选用卷积块注意力模块 (Convolutional Block Attention Module, CBAM) 模型^[28]。CBAM 主要包括两个部分:通道注意力模块 (Channel Attention Module) 和空间注意力模块 (Spatial Attention Module)。

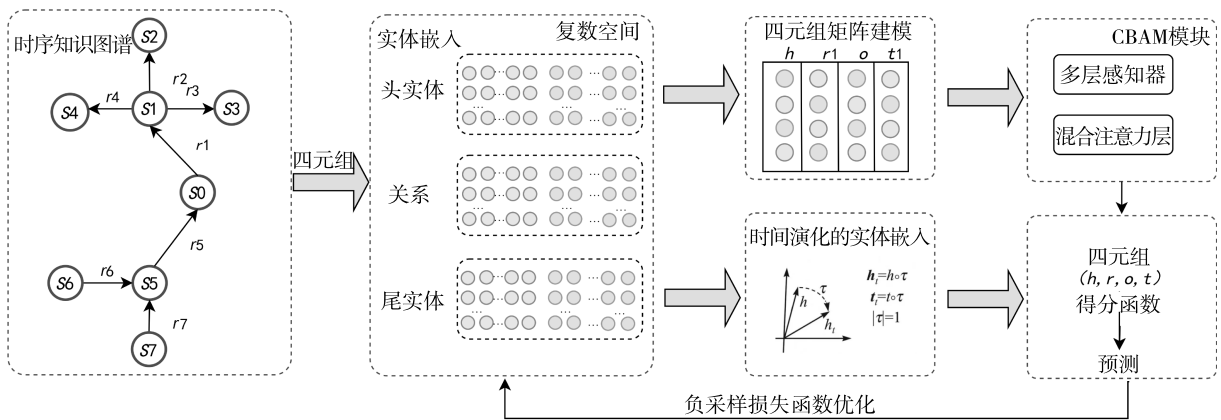


图 1 本文模型框架

CBAM 通过结合通道注意力和空间注意力, 实现了对输入特征的双重精炼, 能够根据任务需求和内容上下文动态地调整特征图中每个通道和空间位置的重要性。这种自适应重标定机制有助于模型更好地关注于关键特征, 提高识别性能。CBAM 中的模块流程图如图 2 所示。

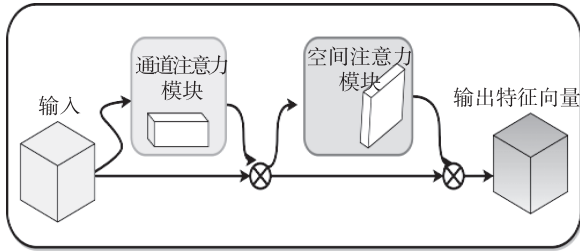


图2 CBAM 流程图

研究应用 CBAM 模块, 经过时间旋转模块后的嵌入向量建模为一个四元组向量矩阵。矩阵经过通道注意力模块后, 得到空间注意力图 $M_c(F)$, 公式如下:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (2)$$

其中, F 表示输入向量矩阵, AvgPool 和 MaxPool 分别表示全局平均池化和最大池化操作, MLP 表示多层感知机, σ 表示 Sigmoid 激活函数。

得到的 $M_c(F)$ 与输入进行相乘操作, 得到 F' :

$$F' = M_c(F) \otimes F \quad (3)$$

F' 继续进入到空间注意力模块, 得到空间注意力图 $M_s(F')$, 公式如下:

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \quad (4)$$

其中, $f^{7 \times 7}$ 表示 7×7 的卷积操作, $[\text{AvgPool}(F); \text{MaxPool}(F)]$ 表示将平均池化和最大池化结果沿通道轴拼接起来。

得到的 $M_s(F')$ 与输入图像 F' 进行相乘操作, 得到 F'' :

$$F'' = M_s(F') \otimes F' \quad (5)$$

2.2 训练过程

根据 TransE 模型的基础得分函数, 本研究设计了一种对于间隔时间内在复数空间下的得分函数, 将经过混合注意力模块的头实体嵌入向量 h_i 和关系嵌入向量之和当作尾实体共轭嵌入向量 $\bar{t}_i$ 的翻译。对于在某个时间间隔内的事实, 分为开始时间和结束时间, 用对偶集合 R_k 和 R_j 作为关系的开始和结束。具体表达式如下:

$$f(h, r, t, [\tau_k, \tau_j]) = \frac{1}{2} ((h_{i,k} + r_k - \bar{t}_{i,k}) + (h_{i,j} + r_j - \bar{t}_{i,j})) \quad (6)$$

其中, $r_k, r_j \in \mathbf{R}$, $h_{i,k}$ 和 $h_{i,j}$ 表示嵌入向量开始时间和结束时间的复数空间下的旋转。

对于缺少开始时间或结束时间的事实, 其得分为已知时间点的四元组得分。对于缺少开始时间的四元组得分函数表达式如下:

$$f(h, r, t, [\tau_k, -]) = (h_{i,k} + r_k - \bar{t}_{i,k}) \quad (7)$$

本文采用的基于负采样的损失函数^[28], 在许多基于翻译的嵌入模型中已经被证实十分有效, 其表达式如下:

$$L = -\log[\rho(\gamma - f(\xi_i)) + s_{\tau,i}] - \sum_{j=1}^n \frac{1}{n} \log[\rho(f(\xi'_j) - \gamma - s'_{\tau,j})] \quad (8)$$

其中, γ 是一个边际参数, 用于控制正负样本之间的间隔; f 表示模型中嵌入向量间的特定翻译或相似度计算函数; $\xi_i \in Q^+$ 代表正样本对; $\xi'_j \in Q^-$ 则指第 j 个负样本对; n 表示每个正样本负采样的数量, 确保了训练中的负采样平衡。整体而言, 该损失函数旨在最大化正样本对的得分同时最小化负样本对的得分, 以此优化模型的嵌入表示能力。

3. 实验过程及结果分析

3.1 数据集

本研究选择在 ICEWS14、ICEWS18、YAGO11k、WIKI12k 四个公开的数据集上进行链接预测实验。具体的统计信息如表 1 所示。ICEWS14 包含了 2014 年 1 月 1 日到 2014 年 12 月 31 日所有发生的事件。ICEWS18 则包含了 2018 年 1 月 1 日到 2018 年 10 月 31 日所有发生的事件。YAGO11k 是由德国马普研究所研制的链接数据库, 从其中包含起始、结束时间的样本选择 10 种最依赖时间的关系类型。WIKI12k 是基于维基百科的开放知识库, 将知识以结构化的数据进行存储, 从其中包含起始、结束时间的样本中选择前 24 种最依赖时间的关系类型。将数据集按照时间步进行排序, 划分成训练集: 验证集: 测试集 = 8 : 1 : 1。

表 1 数据集信息统计

数据集	实体	关系	训练集	验证集	测试集	时间步	间隔
ICEWS14	7 128	230	63 685	13 823	49 545	365	每天
ICEWS18	23 033	256	373 018	45 995	13 222	304	每天
YAGO11k	10 623	10	161 540	19 523	20 026	189	每年
WIKI12k	12 554	24	539 286	67 538	63 110	232	每年

3.2 评价指标

时序知识图谱链接预测的评价指标旨在衡量模型在预测知识图谱中缺失实体或关系时的性能。通常选取的评价指标为: 平均倒数排序 (Mean Reciprocal Ranking,

MRR) 和命中率 Hits@ k 。通常 k 的取值为 1、3、10。平均倒数排序表示在测试集中正确事实的平均排名倒数, MRR 的数值越大, 说明 KGC 模型的链接预测效果越好。Hits@ k 表示在测试集中得分排名前 n 的正确事实所占的比率, 例如, Hit@10 表示正确答案出现在模型预测的前 10 个实体中的概率; Hit@1 (也称为 Recall@1) 特别关注模型能否直接预测出最有可能的正确答案。Hits@ k 的数值越大, 说明 KGC 模型的链接预测效果越好。MRR 和 Hits@ k 计算方法分别如式 (9)、式 (10) 所示:

$$\text{MRR} = \frac{1}{|D_{\text{test}}|} \sum_{(s,r,o,t) \in D_{\text{test}}} \left(\frac{1}{\text{rank}(o | s, r, t)} \right) \quad (9)$$

$$\text{Hit@}k = \frac{1}{|D_{\text{test}}|} \sum_{(s,r,o,t) \in D_{\text{test}}} (\text{rank}(o | s, r, t) \leq k ? 1 : 0) \quad (10)$$

其中, D_{test} 表示测试集四元组个数, rank 表示打分后的排序名次。

3.3 实验设置

模型的嵌入维度均设置为 500 维, 负样本数量 $n = 10$ 。经过参数敏感度分析, 模型的超参数分别设置为: 在 ICEWS14 上, 学习率为 0.1, $\gamma = 110$; 在 ICEWS18 上, 学习率为 0.1, $\gamma = 120$; 在 YAGO11k 上, 学习率为 0.1, $\gamma = 50$; 在 WIKI12k 上, 学习率为 0.3, $\gamma = 20$ 。

3.4 对比实验

研究选择的对比模型涵盖了静态与时序知识图谱补全两大范畴。在静态知识图谱补全领域, 选用的模型包括 TransE、DistMult、ComplEx-N3、RotatE、QuatE。而在时序知识图谱补全方面, 则选取了 TTransE、HyTE、TA-TransE、TA-DistMult、DE-SimpleE、ATiSE 和 TeRo。通过与这些方法的综合对比, 旨在全方位、深入地检验 YiTX

在处理时序数据、捕捉动态关系演变方面的效能与优势。

所有选定的基线模型均提供了开源代码, 确保了实验的可复现性与公平性。为了确保对比的一致性和公正性, YiTX 模型与各基线模型采用了相同的配置: 优化算法、嵌入向量的维度、初始学习速率及批处理尺寸维持一致。同时, 各模型依据其原始论文推荐的设置进行了初始化, 并针对具体实验进行了微调, 旨在达到各个模型的最佳性能状态。实验结果如表 2 和表 3 所示, 最优性能的数值以粗体突出, 便于快速识别和对比各模型的表现差异。

由表 2 和表 3 可以看出: (1) 在 ICEWS14 和 WIKI12 数据集中, 本文所提出模型的实验结果均等于或超过对比模型中最好的实验结果; 在四个数据集中, MRR 均等于或大于对比模型; (2) 与单一学习三元组信息的静态补全方法对比, 包括 YiTX 在内的动态补全方法各项指标都远超前于静态方法, 证明处理知识图谱补全任务中, 利用时间信息的必要性; (3) 与同样将时间旋转方法纳入建模的 TeRo 相比, YiTX 在多项指标上的整体表现更好, 在 ICEWS14 数据集上, 模型比 TeRo 在 Hits@10 上提高了 1.8 个百分点, 在 YAGO11k 数据集上, 模型比 TeRo 在 Hits@10 上提高了 0.7 个百分点, 证明混合注意力机制对模型的改进是有效的, 有利于充分拟合时序知识图谱中的复杂关系, 提高补全的准确性; (4) 除在 ICEWS18 和 YAGO11k 中 Hit@1 和 Hit@3 指标略低于 TeRo, 其他指标均高于其他基线模型。根据表 1 的数据集统计信息, 该两个数据集实体关系数量多, 时间数量多, 时间跨度大, 降低了时间信息建模的难度, 对于时间独立补全方法 TeRo 效果更好。本文方法虽然也对时间信息单独建模, 但提取多时间特征仍存在提升空间。

表 2 不同方法在数据集 ICEWS14 和 ICEWS18 上的预测结果

Datasets	ICEWS14				ICEWS18			
Metrics	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@3	Hits@10
TransE	0.280	0.094	—	0.637	0.294	0.090	—	0.663
DistMult	0.439	0.323	—	0.672	0.456	0.337	—	0.691
ComplEx-N3	0.467	0.467	0.527	0.716	0.481	0.362	0.535	0.729
RotatE	0.418	0.291	0.478	0.690	0.304	0.164	0.355	0.595
QuatE	0.471	0.353	0.530	0.712	0.482	0.370	0.529	0.727
TTransE	0.255	0.074	—	0.601	0.271	0.084	—	0.616
HyTE	0.297	0.108	0.416	0.655	0.316	0.116	0.445	0.681
TA-TransE	0.275	0.095	—	0.625	0.299	0.096	—	0.668
TA-DistMult	0.477	0.363	—	0.686	0.474	0.346	—	0.728
DE-SimpleE	0.526	0.418	0.592	0.725	0.513	0.392	0.578	0.748
ATiSE	0.550	0.436	0.629	0.750	0.519	0.378	0.606	0.794
TeRo	0.562	0.468	0.621	0.732	0.586	0.469	0.668	0.795
YiTX	0.563	0.468	0.625	0.750	0.588	0.465	0.669	0.795

表 3 不同方法在数据集 YAGO11k 和 YIKI12k 上的预测结果

Datasets	YAGO11k				YIKI12k			
	MRR	Hits@ 1	Hits@ 3	Hits@ 10	MRR	Hits@ 1	Hits@ 3	Hits@ 10
TransE	0.100	0.015	0.138	0.244	0.178	0.100	0.192	0.339
DistMult	0.158	0.107	0.161	0.268	0.222	0.119	0.238	0.460
ComplEx-N3	0.167	0.106	0.154	0.282	0.233	0.123	0.253	0.436
RotatE	0.167	0.103	0.167	0.305	0.221	0.116	0.236	0.461
QuatE	0.164	0.107	0.148	0.270	0.230	0.125	0.243	0.416
TTransE	0.108	0.020	0.150	0.251	0.172	0.096	0.184	0.329
HyTE	0.105	0.015	0.143	0.272	0.180	0.098	0.197	0.333
TA-TransE	0.127	0.027	0.160	0.326	0.178	0.030	0.267	0.429
TA-DistMult	0.161	0.103	0.171	0.292	0.218	0.122	0.232	0.447
ATiSE	0.170	0.110	0.171	0.288	0.280	0.175	0.317	0.481
TeRo	0.187	0.121	0.197	0.319	0.299	0.198	0.329	0.507
YiTX	0.188	0.121	0.196	0.326	0.302	0.198	0.331	0.508

3.5 消融实验

本节设计了两组实验，实验设置不变的条件下，分别去除混合注意力机制和将混合注意力机制改变成传统注意力机制，分别记为 Model-A 和 Model-B 模型，以此验证模型中混合注意力模块的有效性。在 ICEWS18 和 YA-

GO11k 两组数据集上进行研究。消融实验结果如表 4 所示，可以观察到对照模型在两个数据集上的性能都不如完整模型，移除和替换单个组件后模型的性能有所下降，证明了本文所提出混合注意力模块各个组件的有效性。

表 4 消融实验结果

Datasets	ICEWS18				YAGO11k			
	MRR	Hits@ 1	Hits@ 3	Hits@ 10	MRR	Hits@ 1	Hits@ 3	Hits@ 10
Model-A	0.586	0.465	0.668	0.795	0.187	0.121	0.193	0.319
Model-B	0.583	0.464	0.667	0.794	0.186	0.12	0.192	0.319
YiTX	0.588	0.465	0.669	0.795	0.188	0.121	0.196	0.326

4 结论

本文提出的基于混合注意力机制的时间旋转时序知识图谱补全模型，将实体的时间演化过程视为复数空间下从开始时间到当前时间的旋转，并加入混合注意力机制，通过整合时间序列分析、特征抽取和注意力引导的信息筛选，在广泛使用的基准时序知识图谱数据集上进行了实验，结果验证了所提出模型的有效性，该研究实现了对复杂时序知识图谱的有效理解和预测。

与目前动态补全工作相比，基于混合注意力机制的时间旋转补全方法在多项衡量指标上获得提升，取得较好的动态知识图谱补全效果。但在更复杂的动态知识库中，提出的方法仍存有提升空间，未来将改进方法，以适用于补全多时间的动态知识图谱。除了解决动态知识库补全任务，时间旋转也提供了将四元组建模的思路。未来将尝试自注意力机制等注意力模型，以适用于动态知识图谱的推理及问答等应用。

参考文献

- [1] HOGAN A, BLOMQUIST E, COCHEZ M, et al. Knowledge graphs [J]. ACM Computing Surveys (CSUR), 2021, 54 (4): 1-37.
- [2] TAI C Y, HUANG L Y, HUANG C K, et al. User-centric path reasoning towards explainable recommendation [C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York; ACM, 2021: 879-889.
- [3] 孙亚伟, 程龚, 厉肖, 等. 基于图匹配网络的可解释知识图谱复杂问答方法 [J]. 计算机研究与发展, 2021, 58 (12): 2673-2683.
- [4] JI H Z, KE P, HUANG S H, et al. Language generation with multi-hop reasoning on commonsense knowledge graph [C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Stroudsburg: Association for Computational Linguistics, 2020: 725-736.

- [5] 王昊奋, 丁军, 胡芳槐, 等. 大规模企业级知识图谱实践综述 [J]. 计算机工程, 2020, 46 (7): 1–13.
- [6] JIN W, QU M, JIN X, et al. Recurrent event network: autoregressive structure inference over temporal knowledge graphs [C]//Proc. of EMNLP, 2020: 6669–6683.
- [7] BORDES A, USUNIER N, GARCIA-DURAN A, et al. Translating embeddings for modeling multi-relational data [C] //Proc. of Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2013: 2787–2795.
- [8] LI Z X, GUAN S P, JIN X L. 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.
- [9] DASGUPTA S S, RAY S N, TALUKDAR P. Hyte: hyperplane-based temporally aware knowledge graph embedding [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018: 2001–2011.
- [10] LIAO S, LIANG S, MENG Z, et al. Learning dynamic embeddings for temporal knowledge graphs [C]//Proc. of WSDM, 2021.
- [11] SADEGHIAN A, ARMANDPOUR M, COLAS A, et al. ChronoR: rotation based temporal knowledge graph embedding [C]//Proc. of AAAI, 2021: 6471–6479.
- [12] SUN Z, DENG Z H, NIE J Y, et al. RotatE: knowledge graph embedding by relational rotation in complex space [C]//Proceedings of the 6th International Conference on Learning Representations, 2018.
- [13] NGUYEN G H, LEE J B, ROSSI R A, et al. Continuoustime dynamic network embeddings [C]//Proceedings of the Web Conference 2018, 2018: 969–976.
- [14] LIU Y, MA Y, HILDEBRANDT M, et al. TLogic: temporal logical rules for explainable link forecasting on temporal knowledge graphs [C]//Proceedings of the 26th AAAI Conference on Artificial Intelligence, 2021: 4120–4127.
- [15] LIN D, WU J, YUAN Q, et al. T-EDGE: temporal weighted multidigraph embedding for Ethereum transaction network analysis [J]. Frontiers in Physics, 2020, 8: 204.
- [16] LIU Z, ZHOU D, ZHU Y, et al. Towards fine-grained temporal network representation via time-reinforced random walk [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 4973–4980.
- [17] WU A, YUAN Y, MA Y, et al. Node embedding research over temporal graph [J]. International Journal of Software and Informatics, 2021, 11 (1): 5–28.
- [18] 崔员宁, 李静, 沈力, 等. Duration-HyTE: 基于持续时间建模的时间感知知识表示学习方法 [J]. 计算机研究与发展, 2020, 57 (6): 1239–1251.
- [19] XU C J, NAYYERI M, ALKHOORY F, et al. Temporal knowledge graph embedding model based on additive time series decomposition [EB/OL]. [2024–06–01]. <http://arxiv.org/abs/1911.07893>.
- [20] GOEL R, KAZEMI S M, BRUBAKER M, et al. Dia-chronic embedding for temporal knowledge graph completion [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34 (4): 3988–3995.
- [21] LACROIX T, OBOZINSKI G, USUNIER N. Tensor decompositions for temporal knowledge base completion [EB/OL]. [2024–06–01]. <http://arxiv.org/abs/2004.04926>.
- [22] XU C, NAYERI M, ALKHOORY F, et al. Temporal knowledge graph completion based on time series gaussian embedding [C]//Proc. of ISWC, 2020: 654–671.
- [23] ZHAO X, LI A, JIANG R, et al. Householder transformation-based temporal knowledge graph reasoning [J]. Electronics, 2023, 12 (9): 2001.
- [24] WANG X, LYU S, WANG X, et al. Temporal knowledge graph embedding via sparse transfer matrix [J]. Information Sciences, 2023, 623: 56–69.
- [25] XU C, NAYERI M, CHEN Y Y, et al. Geometric algebra based embeddings for static and temporal knowledge graph completion [J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35 (5): 4838–4851.
- [26] ZHANG, LIANG X, LI Z, et al. BiQCap: a biquaternion and capsule network-based embedding model for temporal knowledge graph completion [C]//Database Systems for Advanced Applications, 2023: 673–688.
- [27] XU C J, NAYYERI M, ALKHOORY F, et al. TeRo: a time-aware knowledge graph embedding via temporal rotation [C]//Proceedings of the 28th International Conference on Computational Linguistics. Stroudsburg: International Committee on Computational Linguistics, 2020: 1583–1593.
- [28] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision, 2018: 3–19.

(收稿日期: 2024–09–13)

作者简介:

王璐璐 (2000–), 通信作者, 女, 硕士研究生, 主要研究方向: 时序知识图谱推理. E-mail: 369090559@qq.com.

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.com