

基于自适应注意力机制的知识图谱补全算法

王杰, 李晓楠, 李冠宇

引用本文

王杰, 李晓楠, 李冠宇. [基于自适应注意力机制的知识图谱补全算法](#)[J]. 计算机科学, 2022, 49(7): 204-211.

WANG Jie, LI Xiao-nan, LI Guan-yu. [Adaptive Attention-based Knowledge Graph Completion](#) [J]. Computer Science, 2022, 49(7): 204-211.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[四元数关系旋转的知识图谱补全模型](#)

Knowledge Graph Completion Model Using Quaternion as Relational Rotation

计算机科学, 2021, 48(5): 225-231. <https://doi.org/10.11896/jsjcx.200300093>

[知识图谱构建技术:分类、调查和未来方向](#)

Knowledge Graph Construction Techniques:Taxonomy,Survey and Future Directions

计算机科学, 2021, 48(2): 175-189. <https://doi.org/10.11896/jsjcx.200700010>

[基于犹豫模糊可信度的知识推理](#)

Approach for Knowledge Reasoning Based on Hesitate Fuzzy Credibility

计算机科学, 2019, 46(1): 131-137. <https://doi.org/10.11896/j.issn.1002-137X.2019.01.020>

[面向不确定性知识表示的描述逻辑 SROIQ\(D\)不确定性扩展](#)

Uncertain Extension of Description Logic SROIQ(D) for Uncertain Knowledge Representation

计算机科学, 2016, 43(Z11): 88-92. <https://doi.org/10.11896/j.issn.1002-137X.2016.11A.019>

[AGM 信念收缩算子的拓扑式刻画](#)

Topological Characterization of AGM Belief Contraction Operator

计算机科学, 2016, 43(9): 87-90. <https://doi.org/10.11896/j.issn.1002-137X.2016.09.016>

基于自适应注意力机制的知识图谱补全算法

王杰 李晓楠 李冠宇

大连海事大学信息科学技术学院 辽宁 大连 116026

(wjie@dmlu.edu.cn)

摘要 现有的知识图谱补全模型通常将多源信息整合为实体和关系学习单一的静态特征表示,但无法表征不同上下文出现的实体和关系的细差含义和动态属性,即实体和关系在涉及不同的三元组时可能有着不同的角色和含义,并因此表现出不同的属性。为此,提出了一种自适应注意力网络用于知识图谱补全,引入自适应注意力建模每个特征维度对特定任务的贡献程度,为目标实体和关系生成动态可变的嵌入表示。具体而言,所提模型通过定义邻居编码器和路径聚合器来处理实体邻域子图中的两种结构,自适应地调整邻居实体和关系路径的注意力得分,以捕获逻辑上与任务最相关的属性特征,为实体和关系赋予符合当前任务的细粒度语义。在链接预测任务中的实验结果表明,所提模型在 FB15K-237 数据集上的 *MeanRank* 指标比 PathCon 降低了 6.9%, *Hits@1* 比 PathCon 提高了 2.3%;在稀疏数据集 NELL-995 和 DDB14 上,其 *Hits@1* 分别达到了 87.9% 和 98%,证明了引入自适应注意力机制能够有效提取实体和关系的动态属性,为二者生成更全面的表示形式,从而提高知识图谱补全精度。

关键词: 知识图谱补全; 邻域子图; 自适应注意力; 知识表示

中图分类号 TP182

Adaptive Attention-based Knowledge Graph Completion

WANG Jie, LI Xiao-nan and LI Guan-yu

Information Science & Technology College, Dalian Maritime University, Dalian, Liaoning 116026, China

Abstract Existing knowledge graph completion models learn a single static feature representation for entities and relationships by integrating multi-source information. But they can't represent the subtle meaning and dynamic attributes of entities and relationships that appear in different contexts. That is, entities and relationships will show different attributes, because they have different roles and meanings when they are involved in different triples. To solve above problems, an adaptive attention network for knowledge graph completion is proposed, which uses adaptive attention to model the contribution of each task-specified feature dimension, and generates dynamic and variable embedding representations for target entities and relationships. Specifically, the proposed model defines the neighbor encoder and the path aggregator to process two structures in the entity neighborhood subgraph, adaptively learn the attention weights to capture the most logically related features of the task, and to give the entities and relationships with fine-grained semantics in line with the current task. Experimental results in link prediction task show that, the *MeanRank* of the proposed model on FB15K-237 dataset is 6.9% lower than PathCon, and *Hits@1* is 2.3% higher than PathCon. For the sparse datasets NELL-995 and DDB14, its *Hits@1* reaches 87.9% and 98% respectively. Therefore, it proves that the introduction of adaptive attention mechanism can effectively extract the dynamic attributes of entities and relationships to generate a more comprehensive embedding representation, and improves the accuracy of knowledge graph completion.

Keywords Knowledge graph completion, Neighborhood subgraph, Adaptive attention, Knowledge representation

1 引言

知识图谱(Knowledge Graph)以三元组形式存储大量的结构化事实,每个三元组(h, r, t)由头实体 h 、尾实体 t 和关系 r 组成,表示真实世界中的物体或抽象的概念间的联系。

大规模的通用的知识图谱如 Freebase^[1], WordNet^[2] 等包含数十亿个三元组,被广泛应用于关系提取^[3]、智能问答^[4] 和推荐系统^[5] 等领域,为人工智能系统提供了可处理的先验知识。然而,知识图谱大都由人工或通过半自动的方式构建,通常存在严重的数据不完整和数据稀疏等问题^[6],因此针对知识

到稿日期:2021-04-14 返修日期:2021-10-23

基金项目:国家自然科学基金(61976032,62002039)

This work was supported by the National Natural Science Foundation of China(61976032,62002039).

通信作者:李冠宇(rabitlee@163.com)

图谱补全的研究受到了广泛关注。

目前,针对知识图谱补全的一种主流的研究是基于图的嵌入表示方法^[7-9],该方法将知识图谱中实体和关系的邻域信息表示为一个低维嵌入的矢量或张量。然而,这类方法在聚合多跳邻居实体(节点)的消息传递过程中往往受限于简单的假设,即所有邻居实体具有同等的贡献,或者在所有任务中为邻域内的实体分配固定不变的权重(注意力得分)。基于以上简单假设的聚合策略只能为实体和关系学习到单一静态的特征表示,而不能精准表述当前上下文中的语义。

实际上,现实知识图谱中很少存在孤立的实体和关系,其属性和特征表示取决于周围的相邻实体和对应关系的信息。如图1所示,实体史蒂夫·乔布斯的邻居实体揭示了其不同的角色,左侧邻居体现其商业角色(实线),而右侧的邻居实体主要描述其家庭角色(虚线)。因此,对于补全三元组(史蒂夫·乔布斯,创始人,?)和(史蒂夫·乔布斯,子女,?)而言,模型需要侧重于实体史蒂夫·乔布斯的不同角色,根据特定任务筛选出高价值的属性信息,为实体动态地生成更精准的代表形式。同理,关系也具有多义性,当涉及不同的实体对时也会呈现出不同的含义。

模型在建模嵌入表示时应根据具体预测任务来识别实体和关系承担的角色,对不同三元组给予不同的重要性和关注度,因此我们引入自适应注意力网络来表征实体和关系的多重角色和动态属性,以提升补全的性能。本文的贡献有以下3点:1)提出了一个基于自适应注意力网络的新型知识图谱补全模型(Adaptive Attention Knowledge Graph Completion, AAKGC),将实体和关系的上下文含义和细粒度语义融入知识图谱补全中;2)定义自适应邻居编码器和路径聚合器两个组件,分别考虑邻居和路径的不同影响来增强实体及关系的自适应表示,捕获逻辑上与任务最相关的属性特征;3)在链接预测任务中,AAKGC在各个数据集上表现出了非常具有竞争力的性能,验证了本文模型的有效性和可行性。

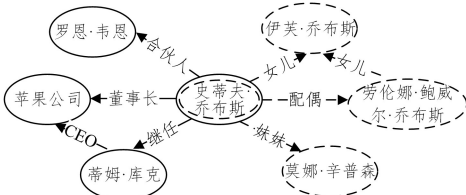


图1 史蒂夫·乔布斯的多种角色示例

Fig.1 Examples of multiple roles of entity Steve Jobs

本文第2节介绍了相关工作,分析了知识图谱补全任务关键技术的发展及存在的问题,阐述了本文研究的目的和意义;第3节详细介绍了本文提出的基于自适应注意力机制的补全算法的核心思想及两个重要组件;第4节利用基准数据集评估了模型在链接预测任务中的性能,并与其他最新算法进行了比较;最后总结全文并展望未来。

2 相关工作

2.1 知识图谱补全模型

针对 KGC 任务,本文通过梳理知识图谱表示学习算法

发现,早期的工作主要集中在静态知识图谱补全(Static Knowledge Graph Completion)场景,以 TransE^[10]为代表的翻译模型虽然获得了较好的效果,但在处理诸如 1-N, N-1 和 N-N 的复杂关系时还存在缺陷。为了解决该问题,学者们提出了 TransH^[11], TransR^[12] 及 STTransE^[13] 等扩展模型,扩展模型引入了特定于关系的平面或子空间,将三元组中的头尾实体投射到相应的关系空间中。此外, RotatE^[14] 进一步将计算扩展到复杂领域以处理多种关系模式,并获得了最优结果。然而,这些模型对超参数比较敏感,并且扩展性也较差。在现实世界中,新实体以及新关系会不间断地产生,翻译模型无法满足自动添加新实体以及新关系的需求。

在此背景下,已有研究人员逐渐把研究重心转移到动态知识图谱补全(Dynamic Knowledge Graph Completion)上,进而寻求能自动扩大知识图谱规模的新方法。语义匹配模型通常采用基于相似性的评分函数,其通过匹配实体的潜在语义和向量空间表示中包含的关系来度量给定三元组的可信性,如 DistMult^[15], ComplEx^[16], SimpleE^[17], QuatE^[18] 等。此外, ConvE^[19], ConvKB^[20] 和 CapsE^[21] 等模型借助卷积神经网络或胶囊神经网络来增强模型的表达能力,进一步将三元组作为一个整体来学习实体和关系的嵌入。此外,还有一些主流模型,如 DRUM^[22] 和 PTransE^[23] 等,在学习嵌入过程中融入实体属性、关系路径,并将其作为附加信息来提升嵌入和预测的精度。但是这些嵌入模型仅从单一三元组中学习实体或关系的静态全局表示,没有充分挖掘实体与其相邻节点间的消息传递和信息交互,忽略了实体周围局部邻域内蕴含的丰富语义,因此,对于实体描述信息较少、关系路径稀疏的知识图谱而言,模型的表达能力大大降低。

2.2 图神经网络算法

传统知识图谱补全模型由于仅关注三元组的拓扑结构信息,因此无法深层次地挖掘三元组各维度的属性特征^[24]。为解决该问题,本文尝试应用图神经网络(Graph Neural Network, GNN)来处理邻域结构信息^[25]以补全知识图谱。该类模型遵循神经消息传递^[26]的概念,整合多跳邻居节点和关系路径中的拓扑结构信息,进而提供更精细的特征表示。

TransGCN^[27]模型将翻译系列模型的思想引入到 GCN 中,综合 TransE 模型和图神经网络的优点,显著改善了链路预测的效果。针对知识图谱中的关系具有不同种类的特点,关系图卷积网络(R-GCN)^[28]通过 GCN 建模关系数据来解决知识图谱中不同关系对实体嵌入的影响。文献^[29]提出了一个全新的匹配网络 GMatching,其中邻居编码器假设所有邻居节点都对实体嵌入作出同等的贡献,以相同的权重集成单跳邻居节点的信息来增强实体嵌入,并定义 LSTM 匹配处理器通过 LSTM 块执行多步匹配。文献^[30]提出的 PathCon 考虑聚合实体周围的相邻关系以及实体对间的关系路径以生成关系上下文和关系路径嵌入表示,并将二者结合起来用于

知识图谱补全。此外,FSRL模型^[31]采用注意力机制为邻居实体分配不同的权重来改进算法,以解决小样本知识图谱补全的问题。但上述模型大都是简单地为相邻节点分配同等的权重,或者采用固定的注意力机制来汇总邻域结构信息,导致模型在涉及不同的任务关系时只能为实体和关系学习相同的静态表示,无法表达它们的动态特性和语义多样性,而本文试图通过自适应注意力网络提取深度自适应于实体和关系的动态表示提高模型的表达能力和补全精度。

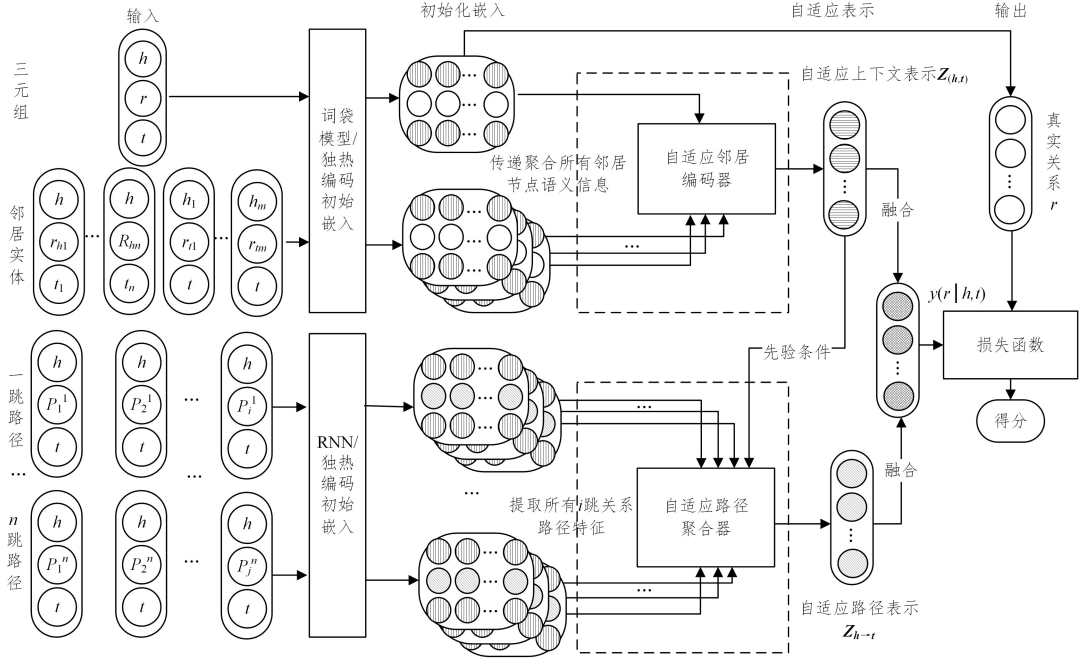


图2 AAKGC模型的总体框架

Fig. 2 Overall framework of AAKGC model

3.1 问题形式化

知识图谱 KG 由大量高质量的结构化三元组构成,表示为 $T = \{(e_i, r_k, e_j) \mid e_i, e_j \in E \wedge r_k \in R\}$, 其中, $E = \{e_1, e_2, \dots, e_M\}$ 为实体集, $R = \{r_1, r_2, \dots, r_N\}$ 为实体间的关系集, 粗体 h, r, t 分别为头实体 h 、关系 r 和尾实体 t 对应的嵌入表示, d 为嵌入向量的维度。对于训练集中的三元组 (h, r, t) , 为了将邻域子图内的相关信息转化为下游补全模型可以处理的输入形式, AAKGC 定义了自适应邻居编码器和路径聚合器, 将图中的拓扑结构和语义信息编码到低维嵌入空间中。在此过程中主要考虑邻域子图的结构, 即实体的邻居实体和实体对间的多跳关系路径。

定义 1(邻居实体) 对于任意实体 h , 局部邻域子图中存在其多个邻居实体, 可以通过三元组 (h, r_k, t_k) 表示实体 h 与实体 t_k 互为邻居实体并存在关系 r_k , 如三元组 (Steve Jobs, Chairman_of, Apple Inc.) 和 (Steve Jobs, Marry To, Laurene Powell Jobs) 表示实体 Apple Inc. 和 Laurene Powell Jobs 为实体 Steve Jobs 的邻居实体。

定义 2(关系路径) 将局部邻域子图内的一条 n 阶关系路径记作 $P^{s=n} = h \rightarrow r_0 \rightarrow \dots \rightarrow r_{n-1} \rightarrow t$, 其是连接实体 h 和 t 的一系列关系形成的序列, 路径的长度定义为序列中的关系数目。例如 Steve Jobs \rightarrow SucceededBy \rightarrow CEO_Of \rightarrow FamousFor \rightarrow

3 AAKGC 模型

本节将对 AAKGC 模型的架构、损失函数和训练过程进行详细介绍。图 2 给出了模型的总体框架, 该框架主要包含两大组件: 1) 自适应邻居编码器, 用于学习动态的自适应实体上下文表示; 2) 自适应路径聚合器, 用于汇总实体对间的关系路径, 学习自适应路径嵌入表示。最后 AAKGC 模型基于获取的自适应实体上下文和路径嵌入表示来预测缺失的关系, 以训练整个模型。

iphone 为长度为 3 的关系路径, 特别地, 可以将实体间的关系视为长度为 1 的关系路径。

给定包含丰富拓扑结构和语义的邻居实体和关系路径, AAKGC 的目标是基于上述两类子图结构学习动态地自适应于每个具体预测任务的实体和关系嵌入表示, 进而预测缺失的三元组信息。

3.2 自适应邻居编码器

知识图谱中实体的局部邻域子图中蕴含丰富的拓扑结构, 已有的补全工作证实了整合实体周围相邻节点信息能够为实体提供更精细的特征表示, 有利于知识图谱补全^[22-27]。然而, 这些方法在邻域消息传递过程中没有考虑注意力或者使用固定的注意力机制, 简单地对所有邻居实体赋予相同或固定的权重, 因此实体只能学习到单一的静态特征表示, 无法表达不同上下文中出现的实体的细差含义, 导致知识推理质量明显下降。为了解决上述问题, AAKGC 为实体设计了一种自适应邻居编码器 (Adaptive Neighbor Encoder), 该编码器的具体细节如图 3 所示。自适应邻居编码器可以根据预测任务和具体上下文动态调整邻居实体的注意力得分, 通过邻居实体的不同影响来捕获实体的多种角色, 以获取实体的自适应嵌入表示。

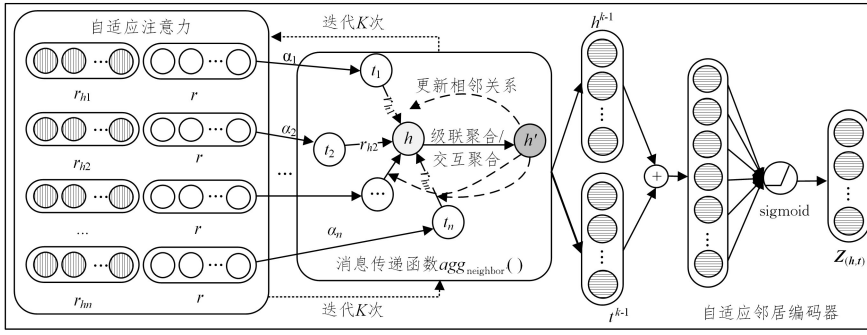


图3 自适应邻居编码器图解

Fig. 3 Diagram of adaptive neighbor encoder

在三元组 \$(h, r, t)\$ 中,实体的角色通常与关系 \$r\$ 高度相关,一个实体与其他实体间的关系可以反映出该实体的性质或类型等有价值的信息。例如,在图1中,如果预测关系 \$r\$ 为“任职”,则实体周围的关系“合伙人”与关系 \$r\$ 的相关性比“配偶”等关系更大,因为前者更能体现实体与职业相关的角色特征,所以自适应邻居编码器直接利用与邻居实体间的关系来建模目标实体角色。对于实体 \$h\$ 的相邻关系 \$r_k \in N_e\$, 首先定义一个度量函数 \$\varphi\$ 来计算任务关系 \$r\$ 和相邻关系 \$r_k\$ 之间的相关性得分。

$$\varphi(\mathbf{h}_{r_k}, \mathbf{h}_{r_0}) = \mathbf{h}_{r_k}^T \mathbf{W} \mathbf{h}_{r_0} + b \quad (1)$$

其中, \$\mathbf{W} \in R^{d \times d}\$ 和 \$b \in R^d\$ 分别为权重矩阵和偏置参数, \$\mathbf{h}_{r_k}\$ 表示相邻关系 \$r_k \in N_e\$ 的初始嵌入, \$\mathbf{h}_{r_0}\$ 表示关系 \$r\$ 的初始特征, 本文采用独热编码 (One-Hot) 或词袋模型 (Bag-of-words Model) 两种方式生成初始嵌入。

这意味着,根据式(1)计算得到的相关性得分 \$\varphi(\cdot)\$ 越高,相邻关系 \$r_k\$ 与任务关系 \$r\$ 就越相关。因此,自适应邻居编码器根据相关性得分为与预测关系最相关的相邻关系赋予更高的注意力得分^[32],使其在接下来的消息传递过程中扮演更重要的角色。

$$\alpha_{r_k} = \frac{\exp(\varphi(\mathbf{h}_{r_k}, \mathbf{h}_{r_0}))}{\sum_{r_k \in N_e} \exp(\varphi(\mathbf{h}_{r_k}, \mathbf{h}_{r_0}))} \quad (2)$$

其中, \$N_e\$ 表示实体 \$e\$ 相邻的关系集合; \$\alpha_{r_k}\$ 表示关系 \$r_k\$ 针对当前预测任务的自适应注意力得分,得分越高表示该关系对应的邻居实体将在实体嵌入中扮演更重要的角色。

为了提升实体嵌入的表达效果,自适应邻居编码器定义了一个消息传递函数 \$agg_{neighbor}(\cdot)\$ 来聚合实体周围 \$k\$ 跳邻居关系生成实体的自适应嵌入表示。整个过程由聚合和传递两个步骤组成,重复执行这两个步骤以便从多跳邻边收集消息,学习每个实体和关系的自适应表示。在第 \$i\$ 次迭代中,根据式(3)和式(4)生成实体 \$e\$ 的自适应嵌入表示 \$\mathbf{c}_e^i\$, 并更新关系 \$r\$ 的隐藏状态 \$\mathbf{h}^i\$。

$$\mathbf{c}_e^i = \sum_{r \in N(e)} \mathbf{h}_r^i \alpha_r^i \quad (3)$$

$$\mathbf{h}_r^{i+1} = agg_{neighbor}(\mathbf{c}_h^i, \mathbf{c}_t^i, \mathbf{h}_r^i), h, t \in N(r) \quad (4)$$

其中, \$N(r)\$ 表示关系 \$r\$ 的头尾实体集合,聚合操作抽象为 \$agg_{neighbor}(\cdot)\$, 本文采用级联聚合函数和交互聚合函数两种方式来实现具体聚合操作过程。级联聚合函数串联第 \$i\$ 次迭代中获得的关系嵌入 \$\mathbf{h}_r^i\$ 以及头尾实体 \$h, t\$ 的消息 \$\mathbf{c}_h^i, \mathbf{c}_t^i\$, 并将串联嵌入送入非线性激活函数以更新下一次迭代中的关系嵌入

\$\mathbf{h}_r^{i+1}\$, 如式(5)所示。

$$\mathbf{h}_r^{i+1} = \sigma([\mathbf{c}_h^i, \mathbf{c}_t^i, \mathbf{h}_r^i] \cdot \mathbf{W}^i + b^i), h, t \in N(r) \quad (5)$$

其中, \$\mathbf{W}^i\$ 和 \$b^i\$ 为可学习的参数矩阵和偏置。

级联聚合函数计算简单且保留了两个输入端点的顺序,但忽略了两个端点之间可能存在的丰富的特征交互。因此,在交互聚合函数中,我们首先计算头尾实体所有元素之间存在的特征交互,然后将交互矩阵 \$\mathbf{c}_h^i \mathbf{c}_t^{iT}\$ 展平为一个向量,并汇总关系嵌入 \$\mathbf{h}_r^i\$ 信息来更新 \$\mathbf{h}_r^i\$, 如式(6)和式(7)所示:

$$\mathbf{C}_h^i \mathbf{c}_t^{iT} = \begin{bmatrix} \mathbf{c}_h^{(1)} \mathbf{c}_t^{(1)T} & \dots & \mathbf{c}_h^{(1)} \mathbf{c}_t^{(d)T} \\ \vdots & \ddots & \vdots \\ \mathbf{c}_h^{(d)} \mathbf{c}_t^{(1)T} & \dots & \mathbf{c}_h^{(d)} \mathbf{c}_t^{(d)T} \end{bmatrix} \quad (6)$$

$$\mathbf{h}_r^{i+1} = \sigma(\text{flatten}(\mathbf{C}_h^i \mathbf{c}_t^{iT}) \mathbf{W}_1 + \mathbf{h}_r^i \mathbf{W}_2 + b^i), h, t \in N(r) \quad (7)$$

其中, \$\mathbf{W}_1, \mathbf{W}_2\$ 和 \$b\$ 为关系 \$r\$ 相关的权重矩阵和偏置参数, \$\sigma(\cdot)\$ 为非线性激活函数。

在上述聚合过程中,邻居实体仅仅收集并临时存储相邻关系中的语义信息,接着在传递过程中将其聚集的消息传递回的每一个相邻的关系用于更新关系嵌入。重复迭代 \$K\$ 次,实体和关系的自适应嵌入表示更新完毕,将最终的实体嵌入 \$\mathbf{c}_h^{K-1}, \mathbf{c}_t^{K-1}\$ 分别作为头尾实体 \$h\$ 和 \$t\$ 的自适应嵌入表示。基于增强的实体嵌入,将两者进一步组合可以得到实体对 \$(h, t)\$ 的自适应上下文表示 \$\mathbf{Z}_{(h,t)}\$:

$$\mathbf{Z}_{(h,t)} = agg_{neighbor}(\mathbf{c}_h^{K-1}, \mathbf{c}_t^{K-1}) \quad (8)$$

因此,自适应邻居编码器根据预测关系确定实体角色,在消息聚合函数中利用自适应注意力动态地为相邻关系分配动态注意力得分,以这种方式获得的实体和关系的自适应表示不仅聚合了局部邻域中的上下文信息,而且满足了特定任务对实体及关系的属性特征和多重角色的要求。

3.3 自适应路径聚合器

事实上,知识图谱中的实体除了通过直接关系相连,还可以通过多跳关系序列连接。关系路径作为邻域子图中连接两个实体的另一种方式,往往包含与直接关系相似的语义,例如实体 Steve Jobs 与 USA 之间的关系路径 Steve Jobs \$\rightarrow\$ birthplace \$\rightarrow\$ located \$\rightarrow\$ USA 意味着这两者之间的直接关系是 Nationality。这表明,实体之间不同的连接路径揭示了它们之间关系的本质,有助于进行预测。为此,本文设计了一种自适应路径聚合器,将多跳路径作为辅助信息建模实体对间潜在的关系,以尽可能完整地将实体间的交互关系保存到向量空间。

对于邻域子图中的一条从头实体 \$h\$ 到尾实体 \$t\$ 的 \$n\$ 阶

关系路径 P^{*n} ,本文参照文献[30]中的设置,删除多跳关系路径中的中间实体,仅保留关系类型的顺序和结构。通过这种方式,AAKGC能够适用于归纳设置,即对于训练集中不存在的新实体,自适应路径聚合器仍可以对其进行建模。因此,定义对应的关系路径 $P=(t_{r_0}, t_{r_1}, \dots, t_{r_{n-1}})$ 为原始多跳路径 P^{*n} 中所有关系类型所构成的序列,其中 t_{r_i} 是边 r_i 的关系类型。为了使AAKGC能够同时捕获实体间的多重关系含义,路径聚合器捕获头尾实体间可能存在的多条路径,生成最终的关系路径嵌入 $\mathbf{Z}_{h \rightarrow t}$ 。

$$\mathbf{Z}_{h \rightarrow t} = \text{agg}_{\text{path}}(\{\mathbf{S}_P\}, P \in P_{h \rightarrow t}) \quad (9)$$

其中, $\text{agg}_{\text{path}}()$ 表示自适应路径聚合函数, $P_{h \rightarrow t}$ 表示邻域子图中从 h 到 t 的所有关系路径的集合,每个关系路径 $P \in P_{h \rightarrow t}$ 的嵌入向量 \mathbf{S}_P 通过式(10)和式(11)获得,考虑到One-Hot编码方式简单高效,而RNN不依赖于相对路径的数量,并且可以基于关系序列捕获不同关系路径之间的相似性,本文采用上述两种编码方式建模关系路径 P 的表示形式 \mathbf{S}_P 。

$$\mathbf{S}_P = \text{RNN}(t_{r_0}, t_{r_1}, \dots) \quad (10)$$

$$\mathbf{S}_P = \text{One-Hot}(t_{r_0}, t_{r_1}, \dots) \quad (11)$$

然而,实体对 (h, r, t) 之间的多条路径可能指示不同的关系。对于具体的预测任务,并非所有的路径在逻辑上都能起作用,且每条路径的重要性也各不相同。考虑到邻居编码器已经捕获了适应于当前任务的实体角色,因此自适应邻居编码器可以借助实体嵌入表示来识别逻辑上与预测关系最相关的关系路径,并为每条路径分配可学习的注意力得分。具体实现方式如下:

$$\mathbf{Z}_{h \rightarrow t} = \text{agg}_{\text{path}}(\{\mathbf{S}_P\}, P \in P_{h \rightarrow t}) = \sum_{P \in P_{h \rightarrow t}} \alpha_P \mathbf{S}_P \quad (12)$$

$$\alpha_P = \frac{\exp(\mathbf{S}_P^T \mathbf{Z}_{(h,t)})}{\sum_{P \in P_{h \rightarrow t}} \exp(\mathbf{S}_P^T \mathbf{Z}_{(h,t)})} \quad (13)$$

其中, α_P 为每条路径 P 对于当前任务的注意力权重, $\mathbf{Z}_{h \rightarrow t}$ 为实体对 (h, t) 间的关系路径的表示。实体对自适应上下文表示 $\mathbf{Z}_{(h,t)}$ 用于帮助识别最重要的关系路径,自适应路径聚合器的具体细节如图4所示。通过这些操作,自适应路径聚合器进一步筛选出最相关的路径用于下游链接预测任务。

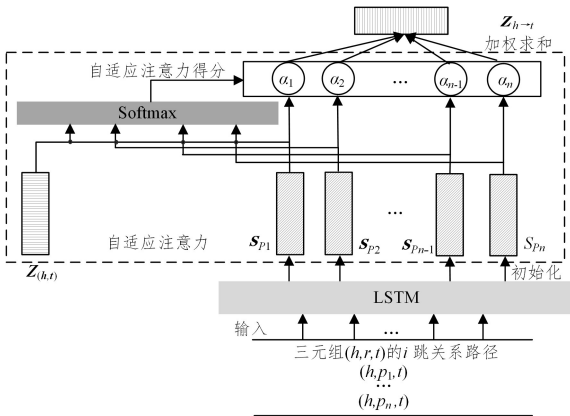


图4 自适应路径聚合器图解

Fig. 4 Diagram of adaptive path aggregator

3.4 AAKGC 模型训练

AAKGC基于自适应邻居编码器和路径聚合器两个组件,引入自适应注意力筛选出邻居实体和关系路径中与当前

任务最相关的信息。紧接着,AAKGC模型融合生成的实体对和关系路径中的语义信息来预测目标关系,并通过最小化损失函数 g 来训练整体模型。

$$\min L = \sum_{(h, r, t) \in Q_{\text{train}}} g(y(r|h, t), r) \quad (14)$$

其中, Q_{train} 为训练集, r 为训练三元组 $(h, r, t) \in Q_{\text{train}}$ 中的真实关系, $g(\cdot)$ 为交叉熵损失函数, L 表示预测关系 $y(r|h, t)$ 真实关系 r 间的损失值, $y(r|h, t)$ 的计算式如式(15)所示:

$$y(r|h, t) = \text{softmax}(\mathbf{Z}_{(h,t)} + \mathbf{Z}_{h \rightarrow t}) \quad (15)$$

为了更清晰地表达模型的处理流程,算法1详细展示了AAKGC的具体训练过程。

算法1 AAKGC 算法

输入:训练样本集 Q_{train} ,实体集 E 和关系集 R ,邻居实体跳数 K ,最大关系路径长度 N ,超参数集合 γ ,嵌入向量维度 d

输出: $e \in E$ 和 $r \in R$ 的嵌入向量,AAKGC的权重矩阵和偏置项参数

1. 初始化:

$r \leftarrow \text{torch.eye}(d)$ for each $r \in R$

$r \leftarrow \text{torch.nn.init.uniform}(r, (-\frac{\gamma}{d}, \frac{\gamma}{d}))$ for each $r \in R$

$\mathbf{S}_P \leftarrow \text{torch.zeros}(|P|, d)$ for each $P \in P_{h \rightarrow t}$

2. 执行步骤2—步骤15直至验证集的准确率性能收敛

3. $Q_{\text{batch}} \leftarrow \text{sample}(Q, b)$

4. $T_{\text{batch}} \leftarrow \emptyset$

5. for $(h, r, t) \in Q_{\text{batch}}$

6. $(h', r', t') \leftarrow \text{sample}(Q'_{(h,r,t)})$

7. $T_{\text{batch}} \leftarrow T_{\text{batch}} \cup \{(h, r, t), (h', r', t')\}$

8. end for

9. 定义自适应邻居编码器 $\text{agg}_{\text{neighbor}}()$ 和自适应路径聚合器 $\text{agg}_{\text{path}}()$

10. for $(h, r, t) \in T_{\text{batch}}$

11. 生成 $N_e, N(r)$ for each $e \in E, r \in R$,根据式(3)、式(4)更新 h_i^l, c_i^l for $i \in [1, K], r \in N_e, e \in N(r)$

12. end for

13. $\mathbf{Z}_{h \rightarrow t} = \text{agg}_{\text{path}}(\{\mathbf{S}_P\}, P \in P_{h \rightarrow t})$

14. $\mathbf{Z}_{(h,t)} = \text{agg}_{\text{neighbor}}(c_h^{K-1}, c_t^{K-1})$

15. 通过最小化损失函数 L 更新实体和关系嵌入向量及模型参数

4 实验

为了评估AAKGC模型的性能,本节在4个知识图谱数据集上进行了链接预测实验,并与其他具有代表性的补全方法的性能进行了对比。实验结果表明,与现有模型相比,AAKGC获得了更具竞争力的实验性能,特别是在NELL-995和DDB14等稀疏数据集上也能达到较好的性能效果。此外,本文还在链接预测任务中对模型进行了消融研究,分别测试了为邻居编码器和路径聚合器添加自适应注意力机制对于补全效果的重要性。

4.1 数据集

本文使用了4个基准数据集,即FB15k, FB15k-237, NELL-995和DDB14。其中,FB15k, FB15k-237和NELL-995为3个常用的链接预测数据集,DDB14为医学领域数据集。FB15K^[8]是知识图谱Freebase的子集,包含了大规模的通用的人类知识,但也存在大量逆反关系;FB15k-237^[33]通过删除所有逆关系来改进FB15K,防止通过反转训练集中的

三元组直接推断出测试三元组的结果;NELL-995^[34]是从NELL系统的第995次迭代中提取的通用知识数据集;DDB14为文献[30]提出的一个新的医学领域数据集,其中包含从疾病数据库中收集的术语和概念,如疾病、症状、药物及其关系,并从原始的三元组中分别随机抽取4000个三元组作为验证集和测试集,各个数据集的详细信息如表1所列。根据表1中所列的平均节点度可以将上述4个数据集分为FB15k和FB15k-237这两个稠密数据集以及NELL-995和DDB14这两个较稀疏的数据集,它们可以有效评估AAKGC对不同疏密程度的知识图谱的补全效果。

表1 数据集的统计信息
Table 1 Statistics of datasets

| Dataset | E | R | # Triples | | | Avg. degree |
|-----------|-------|------|-----------|-------|-------|-------------|
| | | | Train | Valid | Test | |
| FB15K | 14951 | 1345 | 483142 | 5000 | 59071 | 64.6 |
| FB15K-237 | 14541 | 237 | 272115 | 17535 | 20466 | 37.4 |
| NELL-995 | 63917 | 198 | 137645 | 5000 | 5000 | 4.3 |
| DDB14 | 9203 | 14 | 36561 | 4000 | 4000 | 7.9 |

4.2 实验设置

4.2.1 负例三元组生成

对于上述4个数据集,令 $T = \{x_1, x_2, \dots, x_{|T|}\}$ 为测试集,其中 $x_i = (h_i, r_i, t_i)$ 表示第 i 个三元组;令 $y_i = \{0, 1\}$ 为标签,记录三元组的真实性,即正例三元组集 $Q = \{(h_i, r_i, t_i) | y_i = 1\}$,负例三元组集 $Q' = \{(h_i, r_i, t_i) | y_i = 0\}$ 。由于数据集中仅包含真实三元组,依照文献[30]中的设置,通过随机采样,用关系 r' 替换三元组 (h, r, t) 中的真实关系 r ,创建一组负例三元组 (h, r', t) 。在这个过程中需要对候选集进行筛选过滤,避免存在一对多、多对一和多对多等复杂关系,导致某些负例三元组恰好是数据集中的真实三元组。按照先前工作^[35]的标准,我们删除出现在训练集、测试集或验证集中所有的负例三元组,以确保实验结果的有效性。

4.2.2 基线设置

AAKGC选用多个现有模型进行性能对比实验,包括TransE^[7], ComplEx^[13], DisMult^[12], RotatE^[11], SimpleE^[14]和QuatE^[15],此类方法利用知识图谱中的结构信息建模实体和关系嵌入表示。此外,AAKGC还与DRUM^[19]和PathCon^[26]模型进行了对比,其中DRUM提取邻域中的关系路径中包含的细粒度语义,而PathCon融合实体上下文和关系路径中的上下文信息以提高预测精度。

4.2.3 参数设定

本文使用Adam作为优化器,采用多次实验来寻找最佳参数设置,其中实体和关系嵌入维度 $d \in \{8, 16, 32, 64\}$,学习率 $\lambda \in \{0.001, 0.005, 0.01, 0.05, 0.1\}$,邻居实体跳数 $k \in \{1, 2, 3, 4\}$,最大路径长度 $n \in \{1, 2, 3, 4\}$,并使用 L_2 正则化用于防止过拟合,损失项的权重 $l \in \{10^8, 10^7, 10^6, 10^5\}$ 。根据验证集中关系预测的准确率结果,得到的最优超参数设置如下:实体和关系嵌入维度 $d=64$,批处理大小 $batch=128$, $epoch=20$,学习率 $\lambda=0.005$, L_2 损失项的权重 $l=10^7$ 。

4.3 链接预测

链接预测任务旨在根据学习到的知识预测三元组中缺失的实体或关系,是知识图谱补全的标准评估任务。在测试

过程中,原始三元组的关系被移除,使用知识图谱中的其他关系进行随机替换,各个基线模型根据损失函数对替换后的三元组整体进行打分。此任务不需要为三元组预测一个最佳答案,而是根据得分进行降序排列,记录正确三元组的排名。

4.3.1 评估指标

为了准确地评估所提模型,本文采用以下3个指标来衡量各个模型的性能^[36]:1)平均倒数排名(Mean Reciprocal Rank, MRR),取正确关系在所有候选关系集中的排序的倒数并求均值;2)MeanRank(MR),衡量正确关系在所有关系中的平均排名;3)Hits@K,正确关系位于候选关系集前 k 位所占的比例。MR值越低表示性能较好,而MRR值和Hits@K值越高则对应的预测结果越准确。

4.3.2 实验结果及分析

链接预测任务的实验结果如表2和表3所列,其中本文参照的各个模型的实验结果来自文献[30]。本文在相同的设置下将AAKGC训练1000个epoch,并将验证集准确率最高时的各个评估指标作为实验结果,重复3次后取均值作为最终结果,以粗体突出显示最佳性能。

表2 模型在FB15K和FB15K-237稠密数据集中的关系预测结果
Table 2 Relationship prediction results of models in FB15K and FB15K-237 dense datasets

| Model | FB15K | | | FB15K-237 | | |
|---------|--------------|--------------|--------------|--------------|--------------|--------------|
| | MRR | MR | Hits@1 | MRR | MR | Hits@1 |
| TransE | 0.962 | 1.684 | 0.940 | 0.966 | 1.352 | 0.946 |
| Complex | 0.901 | 1.553 | 0.844 | 0.966 | 1.352 | 0.946 |
| DisMult | 0.661 | 2.555 | 0.439 | 0.875 | 1.927 | 0.806 |
| RotatE | 0.979 | 1.206 | 0.967 | 0.970 | 1.315 | 0.951 |
| SimpleE | 0.983 | 1.308 | 0.972 | 0.970 | 1.315 | 0.951 |
| QuatE | 0.984 | 1.207 | 0.972 | 0.974 | 1.283 | 0.958 |
| DRUM | 0.945 | 1.527 | 0.945 | 0.959 | 1.541 | 0.905 |
| PathCon | 0.984 | 1.526 | 0.974 | 0.979 | 1.187 | 0.964 |
| AAKGC | 0.992 | 1.443 | 0.988 | 0.992 | 1.104 | 0.987 |

表3 模型在NELL-995和DDB14稀疏数据集中的关系预测结果
Table 3 Relationship prediction results of models in NELL-995 and DDB14 sparse datasets

| Model | NELL-995 | | | DDB14 | | |
|---------|--------------|--------------|--------------|--------------|--------------|--------------|
| | MRR | MR | Hits@1 | MRR | MR | Hits@1 |
| TransE | 0.841 | 5.253 | 0.781 | 0.966 | 1.161 | 0.948 |
| Complex | 0.703 | 23.040 | 0.625 | 0.953 | 1.287 | 0.931 |
| DisMult | 0.634 | 23.530 | 0.524 | 0.927 | 1.419 | 0.886 |
| RotatE | 0.729 | 23.894 | 0.691 | 0.953 | 1.281 | 0.934 |
| SimpleE | 0.716 | 26.120 | 0.671 | 0.924 | 1.540 | 0.892 |
| QuatE | 0.752 | 21.340 | 0.706 | 0.946 | 1.347 | 0.922 |
| DRUM | 0.715 | 18.203 | 0.640 | 0.958 | 1.140 | 0.930 |
| PathCon | 0.896 | 2.258 | 0.844 | 0.980 | 1.058 | 0.966 |
| AAKGC | 0.914 | 3.516 | 0.879 | 0.987 | 1.056 | 0.980 |

由表2—表4可以看出:1)AAKGC的绝大多数指标在4个基准数据集上获得了较为突出的实验结果,其中,MRR和Hits@1指标在所有的数据集上均明显优于其他模型,证明了本文模型的可行性和鲁棒性;2)与仅学习单一三元组结构信息的TransE,DisMult和ComplEx等嵌入模型相比,本文模型的预测性能提升显著,说明AAKGC成功提取出了多跳局部邻域子图内的复杂特征和结构,为实体和关系表示提供了丰富的特征语义;3)AAKGC和本文的基模型PathCon都考虑了融合邻域信息来提升链接预测的性能,而AAKGC

模型性能在多项指标上的整体表现更好,在 FB15K-237 数据集上,其 MR 值比 PathCon 降低了 6.92%, Hits@1 值比 PathCon 提高了 2.39%,说明在聚合实体邻域信息过程中,引入自适应注意力能有效捕获实体和关系在不同上下文中的细差含义和动态属性,而这对于提高预测精准度同等重要;4)在 NELL-995 数据集上,本文模型的 MRR 和 Hits@1 值达到了 91.4%和 87.9%,表明即使对于稀疏的知识图谱,本文模型也能保持良好的泛化性能。为了进一步分析模型的性能,本文对模型进行了消融研究,以测试自适应注意力机制对于补全效果的重要性。根据是否为邻居编码器和路径聚合器加入自适应注意力机制可将模型表示为 AAKGC($agg_{neighbor}$)和 AAKGC(agg_{path})。图 5 给出了 3 个模型在 NELL-995 数据集上的 MRR 性能对比结果,在其他数据集上也观察到了类似的变化。

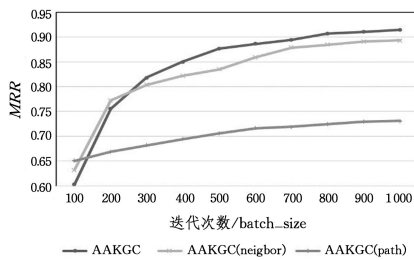


图 5 3 个模型在 NELL-995 数据集上的 MRR 性能对比结果

Fig. 5 MRR performance comparison results of 3 models on NELL-995 dataset

由图 5 可以看出,随着迭代次数的增加,3 个对比模型在 NELL-995 数据集上的 MRR 值逐渐上升,最终 AAKGC 稳定在 0.91 左右,而 AAKGC($agg_{neighbor}$)和 AAKGC(agg_{path})仅仅达到 0.89 和 0.73,这表明:1)通过为邻居编码器引入自适应注意力,AAKGC 可以根据预测任务和上下文动态调整邻居实体的注意力得分,使得所提模型充分考虑到邻居实体的不同影响,从而捕获实体的多重角色和动态属性,极大地提升了知识推理的质量;2)为邻居编码器引入自适应注意力,AAKGC 可以有效识别在逻辑上与预测关系最相关的关系路径,建模实体对间潜在的交互作用,揭示它们之间关系的本质,有助于关系预测。由此证明了 AAKGC 为邻居编码器和路径聚合器引入自适应注意力机制的有效性和实用性。

此外,我们在实验过程中还发现,对于不同的数据集,选取的邻居节点的跳数、最大路径长度以及邻居聚合器会对模型预测的性能产生影响。因此,本文针对每个数据集分别调整了超参数的取值,各个数据集在实验中达到最优性能所选用的参数值如表 4 所列。

表 4 不同数据集的最佳超参数设置

Table 4 Optimal parameter settings for different datasets

| Parameter | FB15K | FB15K-237 | NELL-995 | DDB14 |
|---------------------|--------|-----------|----------|-------|
| Neighbor hops | 2 | 2 | 2 | 3 |
| Maxpath length | 2 | 3 | 3 | 4 |
| Neighbor Aggregator | Concat | Concat | Concat | Cross |

结束语 本文针对知识图谱补全提出了一种基于自适应

注意力机制的知识图谱补全算法,利用实体局部邻域中的数据特征和子图结构建模实体和关系嵌入表示。所提模型在消息聚合的过程中引入自适应注意力机制,使得自适应邻居编码器和路径聚合器重点关注与当前预测关系在逻辑上最为相关的角色和属性,捕获实体和关系面向当前任务的嵌入表示,进而提高知识图谱补全的精度。实验结果表明,本文算法的整体性能明显优于其他基线模型,证实了自适应邻居编码器和路径聚合器能有效捕获实体和关系的细差含义和动态属性。此外,所提模型在 NELL-995 数据集上达到了极高的准确率,表明所提算法在处理稀疏知识图谱时也具有一定的归纳能力。本文模型的源码及相应数据集可通过 GitHub¹⁾ 获得。今后我们将考虑为自适应邻居编码器和路径聚合器融入其他多源信息,如文本信息或关系顺序等,以提升链接预测的准确性。

参考文献

- [1] BOLLACKER K, EVANS C, PARITOSH P, et al. Freebase: A Collaboratively Created Graph Database for Structuring Human Knowledge[C]// Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, 2008: 1247-1250.
- [2] MILLER G A. WordNet: A Lexical Database for English[J]. Communications of the ACM, 1995, 38(11): 39-41.
- [3] XIANG R, WU Z, HE W, et al. Cotype: Joint Extraction of Typed Entities and Relations with Knowledge Bases[C]// Proceedings of the 26th International Conference on World Wide Web, Perth: ACM, 2017: 1015-1024.
- [4] HUANG X, ZHANG J, LI D, et al. Knowledge Graph Embedding based Question Answering [C] // Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, Melbourne: ACM, 2019: 105-113.
- [5] WANG H, ZHAO M, XIE X, et al. Knowledge Graph Convolutional Networks for Recommender Systems[C]// Proceedings of the World Wide Web Conference, San Francisco: ACM, 2019: 3307-3313.
- [6] DING J H, JIA W J. Summary of knowledge graph completion algorithms[J]. Information and Communication Technology, 2018, 12(1): 56-62.
- [7] ARORA S. A survey on graph neural networks for knowledge graph completion[J]. arXiv: 2007. 12374, 2020.
- [8] OH B, SEO S, LEE K H. Knowledge graph completion by context-aware convolutional learning with multi-hop neighborhoods [C]// Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018: 257-266.
- [9] WANG C, SHA Y. Neighborhood Aggregation Embedding Model for Link Prediction in Knowledge Graphs[C]// International Conference on Web Engineering. Cham: Springer, 2020: 188-203.
- [10] BORDES A, USUNIER N, et al. Translating Embeddings for Modeling Multi-relational Data[C]// Proceedings of the NIPS. Cambridge: MIT Press, 2013: 2787-2795.
- [11] WANG Z, ZHANG J, FENG J, et al. Knowledge Graph Embedding by Translating on Hyperplanes[C]// Proceedings of the

¹⁾ <https://github.com/WJ-RECORD/AAKGC>

- AAAI. Menlo Park, CA: AAAI, 2014; 1112-1119.
- [12] LIN Y, LIU Z, SUN M, et al. Learning Entity and Relation Embeddings For Knowledge Graph Completion[C]//Proceedings of the AAAI. Menlo Park, CA: AAAI, 2014; 2181-2187.
- [13] NGUYEN D Q, SIRTS K, QU L, et al. STransE: A Novel Embedding Model of Entities and Relationships in Knowledge Bases [C]//Proceedings of the North American Chapter of the Association for Computational Linguistics. San Diego California: The Association for Computational Linguistics, 2016; 460-466.
- [14] SUN Z, DENG Z, NIE J, et al. Rotate: Knowledge Graph Embedding by Relational Rotation in Complex Space[J]. arXiv: 1902.10197, 2019.
- [15] YANG B, YIH W, HE X, et al. Embedding Entities and Relations for Learning and Inference in Knowledge Bases[J]. arXiv: 1412.6575, 2014.
- [16] TROUILLON T, WELBL J, RIEDEL S, et al. Complex Embeddings for Simple Link Prediction[C]//Proceedings of the 33rd International Conference on Machine Learning. New York City: JMLR.org, 2016; 2071-2080.
- [17] KAZEMI S, POOLE D. Simple Embedding for Link Prediction in Knowledge Graphs[C]//Advances in Neural Information Processing Systems. Canada, 2018; 4284-4295.
- [18] ZHANG S, YI T, YAO L, et al. Quaternion Knowledge Graph Embeddings[C]//Advances in Neural Information Processing Systems. Vancouver, 2019; 2731-2741.
- [19] DETTMERS T, MINERVINI P, STENETORP P, et al. Convolutional 2D Knowledge Graph Embeddings[C]//Proceedings of the AAAI. New Orleans: AAAIPress, 2018; 1811-1818.
- [20] DAI Q N, TU D N, NGUYEN D Q, et al. A Novel Embedding Model for Knowledge Base Completion based on Convolutional Neural Network[C]//Proceedings of the NAACL. New Orleans: Association for Computational Linguistics, 2018; 327-333.
- [21] VU T, NGUYEN T D, NGUYEN D Q, et al. A Capsule Network-Based Embedding Model for Knowledge Graph Completion and Search Personalization [C] // Proceedings of the NAACL-HLT. Minneapolis: Association for Computational Linguistics, 2019; 2180-2189.
- [22] SADEGHIAN A, ARMANDPOUR M, DING P, et al. DRUM: End-To-End Differentiable Rule Mining On Knowledge Graphs [C] // Advances in Neural Information Processing Systems. Vancouver, 2019; 15321-15331.
- [23] LIN Y, LIU Z, LUAN H, et al. Modeling Relation Paths for Representation Learning of Knowledge Bases[C]//Proceedings of the EMNLP. Lisbon, The Association for Computational Linguistics, 2015; 705-714.
- [24] NIU G, LI YANG, TANG C, et al. Relational Learning with Gated and Attentive Neighbor Aggregator for Few-Shot Knowledge Graph Completion[J]. arXiv: 2104.13095, 2021.
- [25] WANG K, LIU Y, XU X, et al. Enhancing knowledge graph embedding by composite neighbors for link prediction[J]. Computing, 2020, 102(12); 2587-2606.
- [26] LIU X, TAN H, CHEN Q, et al. RAGAT: Relation Aware Graph Attention Network for Knowledge Graph Completion [J]. IEEE Access, 2021, 9; 20840-20849.
- [27] CAI L, YAN B, MAI G, et al. TransGCN: Coupling Transformation Assumptions with Graph Convolutional Networks for Link Prediction[C] // Proceedings of the 10th International Conference on Knowledge Capture. Marina Del Rey: ACM, 2019; 131-138.
- [28] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling Relational Data with Graph Convolutional Networks[C]//15th International Conference on Extended Semantic Web Conference. ESWC, 2018; 593-607.
- [29] XIONG W, YU M, CHANG S, et al. One-Shot Relational Learning for Knowledge Graphs[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels: Association for Computational Linguistics, 2018; 1980-1990.
- [30] WANG H, REN H, LESKOVEC J. Entity Context and Relational Paths for Knowledge Graph Completion[J]. arXiv: 2002.06757, 2020.
- [31] ZHANG C, YAO H, HUANG C, et al. Few-Shot Knowledge Graph Completion [C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI Press, 2020; 3041-3048.
- [32] ZHAO X J, JIA Y, LI A P, et al. Research on Link Prediction Model Based on Hierarchical Attention Mechanism[J]. Journal of Communications, 2021, 42(3); 36-44.
- [33] XIE Z, ZHOU G, LIU J, et al. ReInceptionE: Relation-aware Inception Network with Joint Local-Global Structural Information for Knowledge Graph Embedding[C]//Proceedings of the 58th ACL. Online: Association for Computational Linguistics, 2020; 5929-5939.
- [34] XIONG W, HOANG T, WANG W Y. DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning[C]//Proceedings of the EMNLP. Copenhagen: Association for Computational Linguistics, 2017; 564-573.
- [35] ZHANG Z, ZHUANG F, ZHU H, et al. Relational Graph Neural Network with Hierarchical Attention for Knowledge Graph Completion[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(5); 9612-9619.
- [36] SUN Z, VASHISHTH S, SANYAL S, et al. A Re-evaluation of Knowledge Graph Completion Methods[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online: Association for Computational Linguistics, 2020; 5516-5522.



WANG Jie, born in 1997, postgraduate, is a member of China Computer Federation. Her main research interests include intelligent information processing and knowledge graph completion.



LI Guan-yu, born in 1963, Ph.D, professor, is a member of China Computer Federation. His main research interests include intelligent information processing and knowledge graph.