

文章编号: 2096-1618(2022)01-0055-07

结合时间映射和卷积神经网络的知识补全

陈新元^{1,2}, 谢晟祎³, 陈庆强⁴, 刘羽⁵

(1. 闽江学院计算机与控制工程学院, 福建 福州 350121; 2. 福州墨尔本理工职业学院信息工程系, 福建 福州 350121; 3. 福建农业职业技术学院实验实训中心, 福建 福州 350300; 4. 福建工程学院信息科学与工程学院, 福州 350118; 5. 福州墨尔本理工职业学院现代教育技术中心, 福建 福州 350121)

摘要: 现有知识库存在大量缺失事实且事实常携带时间信息。针对主流嵌入表示方法在知识补全时常忽略时间维度的问题, 设计一种时间敏感的三元组嵌入表示方法 TSKGC (time sensitive knowledge graph completion)。通过为时间戳分配超平面, 将时序信息合并到实体关系空间中, 并进一步将映射后三元组的 3 列 k 维矩阵表示用作卷积神经网络的输入, 在不同超平面对应的多通道中并行处理, 提取三元组特征用于知识补全。在 YAGO11k 和 Wikidata12k 数据集上的实验证明, TSKGC 具备一定的时间预测能力, 并能有效利用时间信息提高链路预测的性能表现, 特别在 1-M、M-1 和 M-M 复杂关系类型上相比主流模型具备一定优势。

关键词: 知识表示; 知识图谱补全; 时间嵌入; CNN; 链路预测

中图分类号: TP18

文献标志码: A

doi: 10.16836/j.cnki.jcuit.2022.01.010

0 引言

知识图 (knowledge graph, KG) 中的节点和边分别对应实体和关系, 以三元组 (头实体, 关系, 尾实体) 的形式编码事实, 即 (h, r, t) , h 和 t 分别为头、尾实体, r 为实体间关系, 如 $(\text{Beijing}, \text{isCapitalOf}, \text{China})$ 。主流知识库如 NELL^[1], YAGO^[2], Freebase^[3] 等, 常用于信息检索^[4]、问题回答^[5-7]、语义搜索引擎^[8]、专家系统和社交网络分析等领域^[9]。现有知识库数据不完整^[10], 知识图谱补全^[11]是近年的热点方向之一^[12-14], 主流思路是学习 KG 中节点和关系的多维向量表示, 同时保留原有结构信息和知识约束。知识有效性常受时间限制, 或仅在特定时段内正确, 例如 $(\text{Donald Trump}, \text{isPresidentOf}, \text{United States})$ 目前仅在 2016–2020 年有效; $(\text{Barack Obama}, \text{presidentOf}, \text{USA})$ 仅在 2008–2017 年有效。部分知识库已开始为事实打上时间标签, 如 YAGO 和 Wikidata^[15] 等。然而主流方法在学习实体嵌入和关系翻译时, 常将知识图视作静态, 忽略了时间信息, 因此相关研究较少; Jiang 等^[16]提出利用时序信息辅助知识补全, 但并非直接将时间整合到嵌入表示中, 而是先学习关系间的相对顺序, 再在嵌入阶段将关系顺序引入为约束。

文中提出一种基于时间映射和卷积神经网络的嵌入表达技术 TSKGC, 将时序信息直接纳入事实嵌入学

习, 旨在提高链路预测的性能表现。TSKGC 先将输入的知识图根据不同的时间戳分为多个子图, 将实体和关系投影至对应时间戳的超平面上, 从而学习携带时间信息的向量表示。主要工作如下: 探讨较少涉及的时间敏感性问题, 将时间信息集成至三元组的嵌入表示中; 将映射三元组的矩阵表示用作卷积神经网络 (CNN) 的输入, 通过多个通道 (不同时间节点对应的超平面) 的卷积遍历提取事实特征。

1 相关工作

1.1 关系嵌入学习

嵌入学习是知识表达的热门领域之一, 常用于知识图谱补全或链路预测, 大致分为两类。一是基于距离的平移嵌入方法, 通过降低关系平移后的实体向量距离优化嵌入表示, 如 TransE^[17], TransH^[18], TransR^[19], TransD^[20], STransE^[21] 和 TransSparse^[22] 等。二是基于矩阵化参数的方法, 如基于矩阵嵌入的 SE^[23] 认为若三元组有效, 头实体在子空间中的映射应接近尾部向量; RESCAL^[24] 使用双线性模型表示特征和关系间的相互作用; DistMult^[25] 和 ComplEx^[26] 使用三重积计算三元组得分。此外, NTN^[27] 使用张量网络提升表达能力; Nguyen 等^[28] 和 Toutanova 等^[29] 引入了关系路径; 也有研究使用文本信息增强等其他方法^[30-34]。

TransE 在层次结构的概念树和词嵌入空间中矢量表示的基础上设计模型, 将标记边对应的关系映射

收稿日期: 2020-11-17

基金项目: 福建省教育科学“十三五”规划 2020 年度课题资助项目 (FJJKG20-402)

为平移向量,即 $v_h + v_r \approx v_t$, 其中 v_h, v_r, v_t 分别为实体 h , t 和关系 r 的嵌入表示。 $d = \|v_h + v_r - v_t\|_p$, 计算当 p 取 $L1$ 或 $L2$ 范数时三元组得分, 该算法相对简单, 但在复杂关系三元组上学习能力欠佳。

TransH 设计了关系特定的超平面 w_r , 将实体进行差异化投影。本文在此基础上, 尝试为时间信息建立超平面以学习知识表达。TransR 使用投影矩阵 W_r 进一步提高表达能力, 但计算开销增加。

近年来, CNN 广泛应用于 NLP^[35] 领域, 代表模型如 ConvE^[13], 使用 CNN 提取头实体和关系特征, 但二维卷积可能导致全局特征丢失。ConvKB^[36] 在 ConvE 基础上调整, 过滤器尺寸 size 设置为 1×3 。

1.2 时间信息嵌入

实体之间的关系基数可分为 1-1, M-1, 1-M 和 M-M 4 种类别, 简单方法在 1-1 关系上可获得较好的建模结果, 而对 M-1, 1-M 和 M-M 等复杂关系, 目前尚无统一的建模标准。部分复杂关系与不同时间节点对应的不同尾实体有关, 但主流关系嵌入模型的研究对象多为静态知识图, 忽略了对时间维度的分析, 因而无法直接对这类关系进行区分。

早期的 T-YAGO^[37] 从 Wikipedia Infoboxes 的半结构化数据中提取时间信息, 使用正则表达式分类; UzZaman 等^[38] 设计了时间提取和表达方案并根据顺序与事实对应; t-TransE^[16] 在 TransE 的基础上通过关系排序学习时间感知的嵌入表达; 双线性模型 Know-Evolve^[39] 使用 RNN 捕获非线性时间变化的动态特征; Leblay 等^[40] 使用知识图中静态的辅助信息进行时间嵌入学习; Dasgupta 等^[41] 将时间对应的超平面与实体关系融合; 另外有许多研究致力于时间相关的推理^[42-43]。文中重点是设计关系嵌入学习模型, 以及进一步提高特征提取能力用于知识推理或补全, 而非提取时序信息。

2 TSKGC

将时间跨度信息以 list 形式合并到知识库的三元组中, 即 $(h, r, t, [\tau_s, \tau_e])$, τ_s 和 τ_e 分别表示有效期的起始和终止; 设定时间点将知识图划分为多个静态子图, 每个子图内包括在对应时间点上有效的所有三元组, 单个三元组可在多个子图上存在。设图 $G = G_{T_1} \cup G_{T_2} \cup G_{T_i} \dots \cup G_{T_N}$, 其中 $i \in 1, 2, \dots, N$ 表示离散的时间节点的编号, 将其对应到 N 个超平面, 即对于 $G_{T_1}, G_{T_2}, \dots, G_{T_N}$ 分别有 $\omega_{T_1}, \omega_{T_2}, \dots, \omega_{T_N}$ 与之对应。若 T_i, T_{i+1}, \dots, T_m 落在三元组的时间跨度 $[\tau_s, \tau_e]$, 即 $\tau_s \leq T_i \leq T_{i+1} \leq \dots \leq T_m \leq \tau_e$,

则该三元组在 $G_{T_1}, G_{T_{i+1}}, \dots, G_{T_m}$ 上有效, 将其投影至这 m 个超平面, 如图 1 所示, 从而保留三元组的时间信息, 并通过缩小平移距离优化嵌入表示。

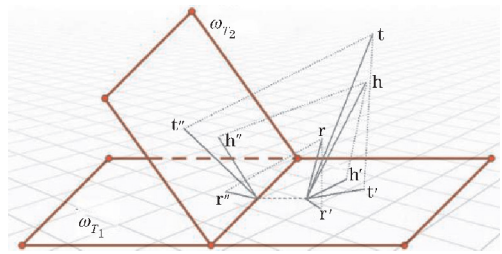


图1 单一三元组映射至多个超平面
(多个时间节点内有效)

使用 CNN 提取三元组嵌入表示的非线性时间演化特征。将对应不同时间节点的 N 个超平面作为卷积层的多个通道, 每通道多个卷积核遍历矩阵, 并行提取特征。过滤器的尺寸 size 和步长 stride 对特征提取和计算开销影响较大, 使用 $\omega \in \mathbb{R}^{1 \times 3}$ 提取同一维度嵌入向量的特征, 步长设置为 3, 提取全局特征的同时避免无意义的局部特征。特征映射 v_i 为

$$v_i = g(\omega \cdot A_{i,:} (A_{i,:}^m) + b) \quad (1)$$

g 为非线性激活函数, b 为偏置系数, 本文使用 ReLU。令 Ω 和 τ 分别表示 ω 的卷积核和核数, 即 $\tau = |\Omega|$ 。卷积层处理后, 特征映射的规模可表示为 $k \times m (\times \tau)$ 。

评分函数 $f(h, r, t)$ 定义为

$$f(h, r, t) = \text{concat}(g([v_h, v_r, v_t] * \Omega)) \cdot w \quad (2)$$

其中, $*$ 表示卷积运算, w 为权重参数。非线性函数的参数通过关系融合和平移转换得到。流程框架如图 2 所示。

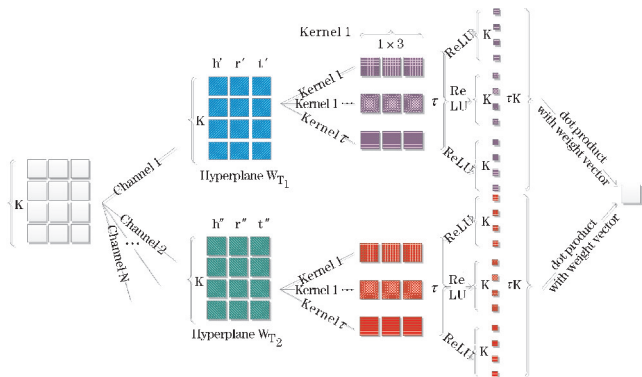


图2 TSKGC 的卷积流程

使用 Adam^[44] 训练 TSKGC^[26], 损失函数 \mathcal{L} 定义为

$$\mathcal{L} = \sum_{(h,r,t) \in \mathcal{G} \cup \mathcal{G}'} \lg(1 + \exp(l_{(h,r,t)} \cdot f(h,r,t))) + \frac{\lambda}{2} \|w\|_2^2 \quad (3)$$

其中, $l_{(h,r,t)} = \begin{cases} 1, & (h,r,t) \in \mathcal{G} \\ -1, & (h,r,t) \in \mathcal{G}' \end{cases}$, \mathcal{G}' 表示无效三元组。

模型算法如 Algorithm 1 所示,开销控制为 $O(n_e k + n_r k)$, 与 TransE 等主流模型在相同量级上。考虑到数据集 中的重数关系对实体/关系预测的影响,即一组头实体 + 关系,或一组头实体 + 尾实体,可能对应单一或多种 尾实体或关系。而单一情况下预测结果应相对较好, 因此根据重数关系,为损失函数添加权重系数,使单一 取值预测错误的惩罚较高。

Algorithm 1: Parameter Optimization for TSKGC

Input: KB \mathcal{G} , entity set \mathcal{E} , relation set \mathcal{R} , embedding dimension k , batch size b , L2 regularizer λ , pre-trained embeddings produced by TransE^[45]

```

1  Initialize variables for each  $v_h, v_r, v_t$ 
   //Using a truncated normal distribution with init_fil-
   ter_normal=True while using  $[0.1, 0.1, -0.1]$  with init_
   _filter_normal = False

2  For each  $\omega$  initialize with init_filter_normal

3   $w \leftarrow \text{uniform}(-\frac{\sqrt{6}}{\sqrt{k \times |\Omega| + 1}}, \frac{\sqrt{6}}{\sqrt{k \times |\Omega| + 1}})$ 

4  For  $i = 1, 2, \dots, n$ ,  $n = \text{epoch number}$ 

5  For  $j = 1, 2, \dots, \frac{|\mathcal{G}|}{b}$ 

6  Batch  $\leftarrow \text{Sample}(\mathcal{G}, b)$ 

7  IBatch  $= \emptyset$  //I short for invalid

8  For each triplet in Batch

9   $(h', r, t') \leftarrow \text{ISample}()$  //sample a corrupted
triplet

10 IBatch  $\leftarrow \text{IBatch} \cup (h', r, t')$ 

11 Batch  $\leftarrow \text{Batch} \cup \text{IBatch}$ 

12 For each triplet  $\in \text{Batch}$ 

13  $f(h, r, t) = \text{concat}(g([v_h, v_r, v_t] * \Omega)) \cdot w$ 

14 compute  $l_{(h, r, t)}$ 

15  $\nabla \mathcal{L}_{\text{Batch}} = \sum_{(h, r, t) \in \text{Batch}} \lg(1 + \exp(l_{(h, r, t)} \cdot$ 
 $f(h, r, t))) + \frac{\lambda}{2} \|w\|_2^2$ 

16 Update weight vector  $w$  and filters  $\Omega$  w. r. t.
```

$\nabla \mathcal{L}_{\text{Batch}}$

无效三元组分为两类:

与时间信息无关:参考 Bordes 等^[6],通过替换实 体创建无效三元组:给定正确三元组 (h, r, t) , η_h 和 η_t 是每个尾部对应头实体和每个头部对应尾实体的 平均数量, $\frac{\eta_t}{\eta_h + \eta_t}$ 和 $\frac{\eta_h}{\eta_h + \eta_t}$ 分别表示生成 (h', r, t) 和 (h, r, t') 的概率。

与时间信息相关:随机选择 KG 中存在的三元组, 但修改其时间节点归属。

3 实验结果与分析

在 YAGO11k 和 Wikidata12k 数据集上进行链路 预测和时间预测^[45],将 TSKGC 算法与 TranE、t-TranE、 HyTE 和 ConvKB 共 4 种主流算法比较。YAGO11k 从 YAGO3^[46]数据集中提取包含 occurSince 和 occurUntil 标注的事实,选择频数最高的 10 种关系并去除部分稀 疏实体。Wikidata12k 的处理流程类似,选择最常见的 24 种关系。数据集统计信息如表 1 所示。

表 1 数据集统计信息

Datasets	Entity	Relation	Train Set	Valid. Set	Test Set
YAGO11k	10623	10	16298	2038	2038
Wikidata12k	12554	24	32536	4065	4065

3.1 链路预测

链路预测在给定关系和头实体或尾实体的前提下 推测丢失实体。文中使用 MR(平均排名)和 Hits@10 (排名在前 10 位的有效实体的百分比)作为评估指 标。实验使用第 1 类方法创建无效样本,取验证集上 Hits@10 得分最高的模型用于测试。

TransE 的超参数初始化设置如下: $k \in [50, 100]$, 学习率设定为 $[1e^{-4}, 5e^{-4}]$, L1 范数, margin $\gamma \in [1, 3, 5, 7]$ 。在 YAGO11k 上,当 $k = 50$,学习率 $= 5e^{-4}$, L1 范数, $\gamma = 1$ 时得分最高;在 Wikidata12k 上,当 $k = 100$, $\gamma = 3$ 且其他参数相同时,得分最高。

t-TransE 的初始设置为 $k \in [20, 50, 100]$,学习率 $\in [1e^{-1}, 1e^{-2}, 1e^{-3}]$, $\gamma_1, \gamma_2 \in [1, 2, 4, 10]$, $\lambda \in [1e^{-1}, 1e^{-2}, 1e^{-3}, 1e^{-4}]$,当 $k = 100$,学习率 $= 1e^{-3}$, $\gamma_1 = \gamma_2 = 4$, $\lambda = 1e^{-2}$ 时表现最优。

HyTE 的参数初始化设置为:batch size $= 50k$, $k \in [50, 100, 200]$,学习率 $\in [1e^{-2}, 1e^{-3}, 1e^{-4}]$, L1 范 数, $\gamma \in [1, 2, 5, 10]$,当 $k = 100$,学习率 $= 1e^{-3}$, $\gamma = 5$ 时表现最佳。

TSKGC 学习率初始设置为 $\in [1e^{-5}, 1e^{-4}, 5e^{-4}]$, batch size 设为 $[128, 256]$, kernel 数 $\tau \in [100, 200, 500]$,轮数 epoch $\in [200, 500, 1000]$, L2 正则化 $\lambda = 0.001$,使用 Algorithm 1 中描述的正态分布初始化过滤 器,使用 Adam 优化器,将 ReLU 作为非线性激活函数。 在 YAGO11k 上,当 $k = 50$,学习率 $= 5e^{-4}$, $\tau = 100$,使用 $[0.1, 0.1, -0.1]$ 的过滤器分布时表现最佳;Wikida- ta12k 上,当 $k = 100$, $\tau = 200$,truncated 分布,其他参数 相同时,得分最优;两个数据集上 batch 大小都为 256, epoch $= 100$ 。

链路预测结果如表 2 所示(最优表现设置为黑体,其次设置为下划线),HyTE、ConvKB 和 TSKGC 的 MR 得分和 Hit@ 10 成绩明显优于 TransE 和 t-TransE,

而 TSKGC 成绩总体上相较 HyTE 和 ConvKB 又有所提高,说明将时间信息集成到实体关系映射中有助于知识补全,使用多通道卷积网络也利于提取全局特征。

表 2 YAGO11k 和 Wikidata12k 上的链路预测结果

Method	YAGO11k				Wikidata12k			
	MR		Hits@ 10/%		MR		Hits@ 10/%	
	Tail	Head	Tail	Head	Tail	Head	Tail	Head
TransE	613	2652	4.9	2.5	738	905	17.4	15.8
t-TransE	338	1840	5.7	2.8	592	784	25.3	17.9
HyTE	225	<u>1549</u>	<u>9.2</u>	4.1	305	437	<u>43.2</u>	36.5
ConvKB	214	1775	8.6	<u>6.5</u>	213	<u>368</u>	38.6	34.8
TSKGC	196	1388	10.8	7.9	<u>247</u>	348	47.5	<u>35.4</u>

YAGO11k 相较 Wikidata12k,复杂关系的比例更高,而相应 Hits@ 10 得分中,TSKGC 的优势相比 HyTE 和 ConvKB 也较明显,因此将 YAGO11k 中的三元组按重数关系分为 4 类,进一步比较 HyTE、ConvKB 和 TSKGC 这 3 种模型在不同关系上的性能,每类别的预测头/尾实体的 Hits@ 10 得分如图 3、图 4 所示。在 1-1 关系中,无论是预测头部或尾部,HyTE、ConvKB 和 TSKGC 的性能相仿;预测 1-M 类型的头部和 M-1 类型的尾部,三者差距也不大;在 M-1 和 M-M 类型的头部预测,以及 1-M 和 M-M 类型的尾部预测这 4 种情况下,TSKGC 的得分均最高,说明其能更好地识别上述复杂关系。

同数据集上以相同参数训练、验证并测试;由于其他主流模型无法用于时间预测,故只比较 TSKGC 与 HyTE 的表现。实验使用调整的 MR 指标^[41],考察模型的时间预测性能;将三元组的构成元素分别映射至所有超平面上,按概率降序排列,正确时间节点对应超平面的排序越前越好。以年为最小时间跨度,允许稀疏事实的年份合并,阈值设置为 500, YAGO11k 和 Wikidata12k 分别取 51 和 63 个时间节点。实验结果如表 3 所示,可见 TSKGC 的时间预测能力在 YAGO11k 中与 HyTE 基本持平,在 Wikidata12k 中略有提升。若时间信息完备,该类算法还可扩展为时间范围预测。

表 3 时间预测的 MR 得分

Model	YAGO11k	Wikidata12k
HyTE	4.36	11.12
TSKGC	4.45	9.38

4 结束语

基于超平面的映射,将事实三元组蕴含的时间信息集成到实体关系嵌入空间中,并进一步结合卷积神经网络,在减小参数规模的同时提取特征并评分,以解决主流算法多在静态图上进行知识表示或推理,没有利用时间信息的局限。实验证明 TSKC 模型的集成时间信息有助于提高链路预测的性能表现。但实验采用的是带时间信息的数据集,这类数据集难以获取,或手工标注成本较高,因此算法的适用范围需进一步泛化。如根据事实是否携带时间信息,选择不同的嵌入模型并在同一框架内计算其向量表示,以便共享参数,用于知识补全等任务。此外,近年来基于关系路径,以及注意力机制的知识表达和推理模型也取得了较好的成果,计划在未来工作中结合上述思路进一步提升模型性能。

图 3 YAGO11k 上针对 4 种类型关系的尾实体预测性能比较 (Hits@ 10)

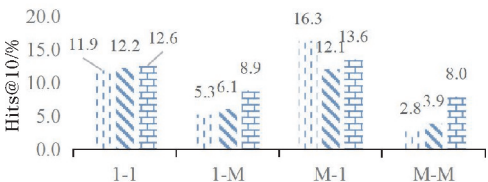


图 3 YAGO11k 上针对 4 种类型关系的尾实体预测性能比较 (Hits@ 10)

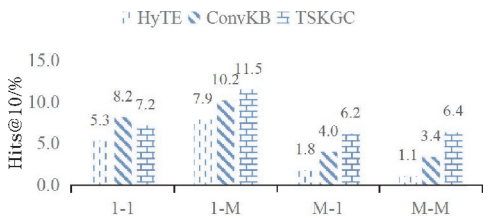


图 4 YAGO11k 上针对 4 种类型关系的头实体预测性能比较 (Hits@ 10)

3.2 时间预测

通过超平面映射将时间维度信息直接集成至事实向量,本文模型可用于给定三元组 (h, r, t, ?) 的时间节点预测。实验采用第 2 类方法创建无效样本,在相

参考文献:

- [1] Mitchell T, Cohen W, Hruschka E, et al. Never-ending Learning [J]. Communications of the ACM, 2018, 61(5): 103–115.
- [2] Rebele T, Suchanek F, Hoffart J, et al. YAGO: A multilingual knowledge base from wikipedia, wordnet, and geonames [C]. International Semantic Web Conference. Cham: Springer, 2016: 177–185.
- [3] 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. Vancouver: ACM, 2008: 1247–1250.
- [4] Xiong C, Callan J. Query Expansion with Freebase [C]. Proceedings of the 2015 international conference on the theory of information retrieval. New York: ACM, 2015: 111–120.
- [5] Dong L, Wei F, Zhou M, et al. Question Answering over Freebase with Multi-column Convolutional Neural Networks [C]. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Beijing: ACL, 2015: 260–269.
- [6] Bordes A, Usunier N, Chopra S, et al. Large-scale Simple Question Answering with Memory Networks [EB/OL]. arXiv preprint arXiv:1506.02075, 2015.
- [7] Hao Y, Zhang Y, Liu K, et al. An End-to-end Model for Question Answering over Knowledge Base with Cross-attention Combining Global Knowledge [C]. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Pennsylvania: ACL, 2017: 221–231.
- [8] Xiong C, Power R, Callan J. Explicit Semantic Ranking for Academic Search via Knowledge Graph Embedding [C]. Proceedings of the 26th international conference on world wide web. Perth: ACM, 2017: 1271–1279.
- [9] LIU Q, LI Y, DUAN H, et al. Overview of Knowledge Map Construction Technology [J]. Computer research and development, 2016, 53(3): 582–600.
- [10] West R, Gabrilovich E, Murphy K, et al. Knowledge Base Completion via Search-based Question Answering [C]. Proceedings of the 23rd international conference on World wide web. Seoul: ACM, 2014: 515–526.
- [11] LIU Z Y, SUN M S, LIN Y K, et al. Research Progress of Knowledge Representation Learning [J]. Computer research and development, 2016, 53(2): 247–261.
- [12] Nickel M, Rosasco L, Poggio T. Holographic embeddings of knowledge graphs [C]. Proceedings of the AAAI Conference on Artificial Intelligence. Phoenix: AAAI press, 2016: 30–41.
- [13] Dettmers T, Minervini P, Stenetorp P, et al. Convolutional 2D Knowledge Graph Embeddings [C]. national conference on artificial intelligence. New Orleans: AAAI press, 2018: 1811–1818.
- [14] Guo S, Wang Q, Wang L, et al. Knowledge Graph Embedding with Iterative Guidance from Soft Rules [EB/OL]. arXiv preprint arXiv: 1711.11231, 2017.
- [15] Erxleben F, Günther M, Krötzsch M, et al. Introducing Wikidata to the Linked Data Web [C]. International Semantic Web Conference. Springer. Cham: Spring, 2014: 50–65.
- [16] Jiang T, Liu T, Ge T, et al. Encoding Temporal Information for Time-aware Link Prediction [C]. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Texas: ACL, 2016: 2350–2354.
- [17] Bordes A, Usunier N, Garcia-Duran A, et al. Translating Embeddings for Modeling Multi-relational Data [C]. Advances in neural information processing systems. Cancouver: MIT press, 2013: 2787–2795.
- [18] Wang Z, Zhang J, Feng J, et al. Knowledge graph embedding by translating on hyperplanes [C]. national conference on artificial intelligence. Vancouver: AAAI press, 2014: 1112–1119.
- [19] Lin Y, Liu Z, Sun M, et al. Learning entity and relation embeddings for knowledge graph completion [C]. national conference on artificial intelligence. Texas: AAAI press, 2015: 2181–2187.
- [20] Ji G, He S, Xu L, et al. Knowledge Graph Embedding via Dynamic Mapping Matrix [C]. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics. Beijing: ACL, 2015: 179–188.

- ciation for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Beijing: ACL, 2015: 687–696.
- [21] Nguyen D Q, Sirts K, Qu L, et al. Stranse: A Novel Embedding Model of Entities and Relationships in Knowledge Bases [EB/OL]. arXiv preprint arXiv:1606.08140, 2016.
- [22] Ji G, Liu K, He S, et al. Knowledge graph completion with adaptive sparse transfer matrix [C]. national conference on artificial intelligence. Phoenix: AAAI press, 2016: 985–991.
- [23] Bordes A, Glorot X, Weston J, et al. A Semantic Matching Energy Function for Learning with Multi-relational Data [J]. Machine Learning, 2014, 94(2): 233–259.
- [24] Nickel M, Murphy K, Tresp V, et al. A Review of Relational Machine Learning for Knowledge Graphs [J]. Proceedings of the IEEE, 2015, 104(1): 11–33.
- [25] Yang B, Yih W, He X, et al. Embedding Entities and Relations for Learning and Inference in Knowledge Bases [EB/OL]. arXiv preprint arXiv:1412.6575, 2014.
- [26] Trouillon T, Welbl J, Riedel S, et al. Complex embeddings for simple link prediction [C]. international conference on machine learning. California: IEEE, 2016: 2071–2080.
- [27] Chen D, Socher R, Manning C D, et al. Learning New Facts from Knowledge Bases with Neural Tensor Networks and Semantic Word Vectors [EB/OL]. arXiv preprint arXiv:1301.3618, 2013.
- [28] Nguyen D Q, Sirts K, Qu L, et al. Neighborhood Mixture Model for Knowledge Base Completion [J]. arXiv preprint arXiv:1606.06461, 2016.
- [29] Toutanova K, Lin X V, Yih W, et al. Compositional Learning of Embeddings for Relation Paths in Knowledge Base and Text [C]. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Berlin: ACL, 2016: 1434–1444.
- [30] Garcia Duran A, Niepert M. Kblrn: End-to-end Learning of Knowledge Base Representations with Latent, Relational, and Numerical Features [EB/OL]. arXiv preprint arXiv:2017.1709.04676.
- [31] Nguyen D Q. An Overview of Embedding Models of Entities and Relationships for Knowledge Base Completion [EB/OL]. arXiv preprint arXiv:1703.08098, 2017.
- [32] Toutanova K, Chen D. Observed Versus Latent Features for Knowledge Base and Text Inference [C]. Proceedings of the 3rd Workshop on Continuous Vector Space Models and their Compositionality. Beijing: ACL, 2015: 57–66.
- [33] Toutanova K, Chen D, Pantel P, et al. Representing Text for Joint Embedding of Text and Knowledge Bases [C]. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Pennsylvania: ACL, 2015: 1499–1509.
- [34] Wang Z, Li J Z. Text-Enhanced Representation Learning for Knowledge Graph [C]. Proceedings of International Joint Conference on Artificial Intelligent. New York: IJCAI, 2016: 1293–1299.
- [35] Kim Y. Convolutional Neural Networks for Sentence Classification [EB/OL]. arXiv preprint arXiv:1408.5882, 2014.
- [36] Nguyen D Q, Nguyen D Q, Nguyen T D, et al. A Convolutional Neural Network-based Model for Knowledge Base Completion and Its Application to Search Personalization [J]. Semantic Web, 2019, 10(5): 947–960.
- [37] Wang Y, Zhu M, Qu L, et al. Timely YAGO: Harvesting, Querying, and Visualizing Temporal Knowledge from Wikipedia [C]. Proceedings of the 13th International Conference on Extending Database Technology. New York: ACM, 2010: 697–700.
- [38] UzZaman N, Llorens H, Derczynski L, et al. Semeval-2013 Task 1: Tempeval-3: Evaluating Time Expressions, Events, and Temporal Relations [C]. Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013). Atlanta: ACL, 2013: 1–9.
- [39] Trivedi R, Dai H, Wang Y, et al. Know-evolve: Deep Temporal Reasoning for Dynamic Knowledge Graphs [C]. Proceedings of the 34th International Conference on Machine Learning-Volume 70. Sydney: ACM, 2017: 3462–3471.
- [40] Leblay J, Chekol M W. Deriving Validity Time in

- Knowledge Graph[C]. Companion Proceedings of the The Web Conference 2018. Cham: Spring, 2018:1771–1776.
- [41] 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. Doha: ACL, 2018:2001–2011.
- [42] Mirza P, Tonelli S. Catena: Causal and Temporal Relation Extraction from Natural Language Texts [C]. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Osaka: ACL, 2016:64–75.
- [43] McDowell B, Chambers N, Ororbia A G, et al. Event Ordering with a Generalized Model for Sieve Prediction Ranking [C]. international joint conference on natural language processing. Taipei: IJCAI, 2017:843–853.
- [44] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization [EB/OL]. arXiv preprint arXiv:1412.6980, 2014.
- [45] Nathani D, Chauhan J, Sharma C, et al. Learning Attention-based Embeddings for Relation Prediction in Knowledge Graphs [J]. arXiv preprint arXiv:1906.01195, 2019.
- [46] Mahdisoltani F, Biega J, Suchanek F M. Yago3: A Knowledge Base from Multilingual Wikipedias [C]. 7th biennial conference on innovative data systems research. California: CIDR Conference, 2014. 13–20.

Knowledge base Completion based on Temporal Mapping and CNN

CHEN Xinyuan^{1,2}, XIE Shengyi³, CHEN Qingqiang⁴, LIU Yu⁵

(1. College of Computer and Control Engineering, Minjiang University, Fuzhou 350121, China; 2. Department of Information Engineering, Fuzhou Melbourne Polytechnic, Fuzhou 350121, China; 3. Experimental Training Center, Fujian Vocational College of Agriculture, Fuzhou 350300, China; 4. Information Science and Engineering College, Fujian University of Technology, Fuzhou 350118, China; 5. Modern Education Technical Center, Fuzhou Melbourne Polytechnic, Fuzhou 350121, China)

Abstract: There are a lot of missing facts in the existing knowledge inventory, and the facts often carry time information. Aiming at the problem that the time dimension is often ignored in the knowledge completion of mainstream embedded representation methods, a time sensitive knowledge graph completion method (TSKGC) is designed. By assigning hyperplanes to the timestamp, the time series information is merged into the entity relation space, and the three-column k-dimensional matrix representation of the mapped triple is further used as the input of the convolution neural network (CNN). In parallel processing in multi-channels corresponding to different hyperplanes, and triplet features are extracted for knowledge completion. Experiments on YAGO11k and Wikidata12k datasets show that TSKGC has a certain time prediction ability, and can effectively use time information to improve the performance of link prediction, especially in 1-M, M-1 and M-M complex relationship types have certain advantages compared with mainstream models.

Keywords: knowledge representation; KB completion; temporal embedding; CNN; link prediction