第八章 知识图谱

戴洪良

计算机科学与技术学院/人工智能学院

hongldai@nuaa.edu.cn

知识图谱推理

- 目前大部分的开放知识图谱，例如WikiData、DBpedia都是由人工或者半自动的方式构建。
- 这些知识图谱比较稀疏，大量实体之间隐含的关系没有被充分地挖掘出来。
- 有些关系并不适合被显示存储下来
  - 如校友关系、同事关系
- 通过知识推理把知识图谱中蕴含了，但没有显示表示的知识推理出来

知识图谱推理方法

- 基于逻辑规则的方法
  - 可基于谓词逻辑、描述逻辑等
  - AMIE (Galárraga et al., 2013)可为知识图谱挖掘Horn rules
  - (Krötzsch et al., 2018) 提出Attributed Description Logics，可在KG中节点或边带属性的情况下推理
- 基于知识图谱嵌入的方法
- 基于神经网络的方法
- 规则与神经网络的结合方法

Galárraga, Luis Antonio, Christina Teflioudi, et al. "AMIE: association rule mining under incomplete evidence in ontological knowledge bases." In Proceedings of WWW. 2013.

基于逻辑规则的方法

- AMIE (Galárraga et al., 2013)的规则挖掘方法
  1. 生成可能的规则（Horn rules）
  2. 用设计的指标评估生成的规则，留下得分高的高质量规则

规则形式： $[r(x,y), B_1, B_2, \ldots, B_n]$

含义： $B_1 \wedge B_2 \wedge \ldots \wedge B_n \Rightarrow r(x,y)$

例： $hasChild(p,c) \wedge isCitizenOf(p,s) \Rightarrow isCitizenOf(c,s)$

基于逻辑规则的方法

- AMIE (Galárraga et al., 2013)的规则挖掘方法

规则形式： $[r(x,y), B_1, B_2, \ldots, B_n]$

含义： $B_1 \wedge B_2 \wedge \ldots \wedge B_n \Rightarrow r(x,y)$

- 通过对空列表[]持续添加原子命题得到候选规则
  - 如对空列表[]添加isCitizenOf(c,s)、hasChild(p,c)、isCitizenOf(c,s)，可得到规则

$$[isCitizenOf(c,s), hasChild(p,c), isCitizenOf(c,s)]$$

即

$$hasChild(p,c) \wedge isCitizenOf(p,s) \Rightarrow isCitizenOf(c,s)$$

基于逻辑规则的方法

- AMIE (Galárraga et al., 2013)的规则挖掘方法

规则形式： $[r(x,y), B_1, B_2, \ldots, B_n]$

含义： $B_1 \wedge B_2 \wedge \ldots \wedge B_n \Rightarrow r(x,y)$

- 提出了不同的评估规则质量的指标

如Standard Confidence:

$$conf(\vec{B} \Rightarrow r(x,y)) := \frac{supp(\vec{B} \Rightarrow r(x,y))}{\#(x,y) : \exists z_1, \ldots, z_m : \vec{B}}$$

满足$\vec{B}$的情况下，$r(x,y)$为真的(x,y)实体对数

满足$\vec{B}$的总(x,y)实体对数

基于知识图谱嵌入的方法

- 各种知识图谱嵌入方法：
  - TransE, NIPS2013, Translating embeddings for modeling multi-relational data
  - TransH, AAAI2014, Knowledge graph embedding by translating on hyperplanes
  - TransR, AAAI2015, Learning Entity and Relation Embeddings for Knowledge Graph Completion
  - TransD, ACL2015, Knowledge graph embedding via dynamic mapping matrix
  - TransA, arXiv2015, An adaptive approach for knowledge graph embedding
  - TranSparse, AAAI2016, Knowledge graph completion with adaptive sparse transfer matrix
  - TransG, arXiv2015, A Generative Mixture Model for Knowledge Graph Embedding
  - KG2E, CIKM2015, Learning to represent knowledge graphs with gaussian embedding

PTransE方法

- PTransE在TransE的基础上加入了实体对间的路径信息

如有路径：$h \xrightarrow{\texttt{BornInCity}} e_1 \xrightarrow{\texttt{CityInState}} e_2 \xrightarrow{\texttt{StateInCountry}} t$

则能推理出：$(h, Nationality, t)$

- 打分函数：

$$f(h,r,t) = E(h,r,t) + E(h,P,t)$$

其中P为h到t的路径集合

$$E(h,r,t) = \|\boldsymbol{h} + \boldsymbol{r} - \boldsymbol{t}\|$$

$$E(h,P,t) = \frac{1}{Z} \sum_{p \in P(h,t)} R(p|h,t) E(h,p,t)$$

$R(p|h,t)$为路径p对实体对(h,t)的可靠度

$E(h,p,t)$为路径p下(h,r,t)的合理性（plausibility）

PTransE方法

$$E(h,P,t)=\frac{1}{Z}\sum_{p\in P(h,t)}R(p|h,t)E(h,p,t)$$

- $R(p|h,t)$为路径p对实体对(h,t)的可靠度（Reliability）
- 并非所有路径都有学习的价值

$$h \xrightarrow{\texttt{Friend}} e_1 \xrightarrow{\texttt{Friend}} e_2 \xrightarrow{\texttt{Profession}} t$$

- PTransE根据一种资源分配算法，基于从h经路径p，t可获得的资源量计算$R(p|h,t)$
- 总的来说，如果从h经路径p可以达到很多不同实体节点，则$R(p|h,t)$值小

PTransE方法

$$E(h,p,t) = \|\boldsymbol{h} + \boldsymbol{p} - \boldsymbol{t}\| = \|\boldsymbol{p} - (\boldsymbol{t} - \boldsymbol{h})\| = \|\boldsymbol{p} - \boldsymbol{r}\|$$

- 需获取路径对应的向量表示$\boldsymbol{p}$
- 通过将路径上的关系对应的向量表示结合得到

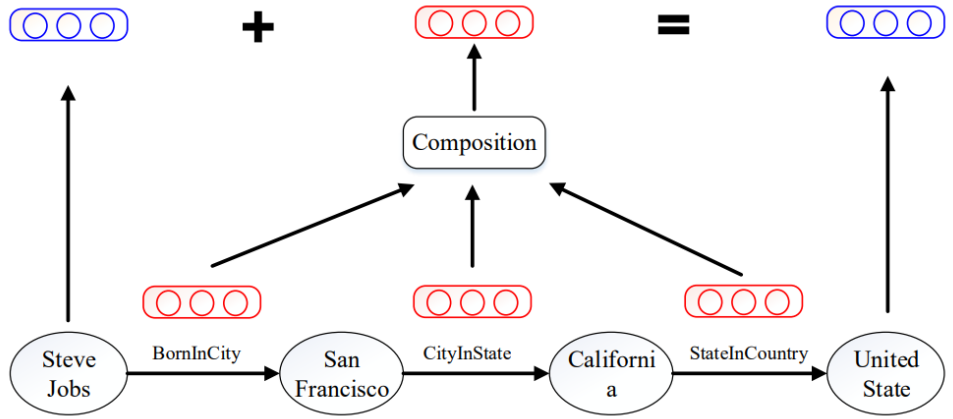
Addition (ADD):

$$\mathbf{p} = \mathbf{r}_1 + \ldots + \mathbf{r}_l$$

Multiplication (MUL):

$$\mathbf{p} = \mathbf{r}_1 \cdot \ldots \cdot \mathbf{r}_l$$

Recurrent Neural Network (RNN): 使用RNN神经网络得到



PTransE方法

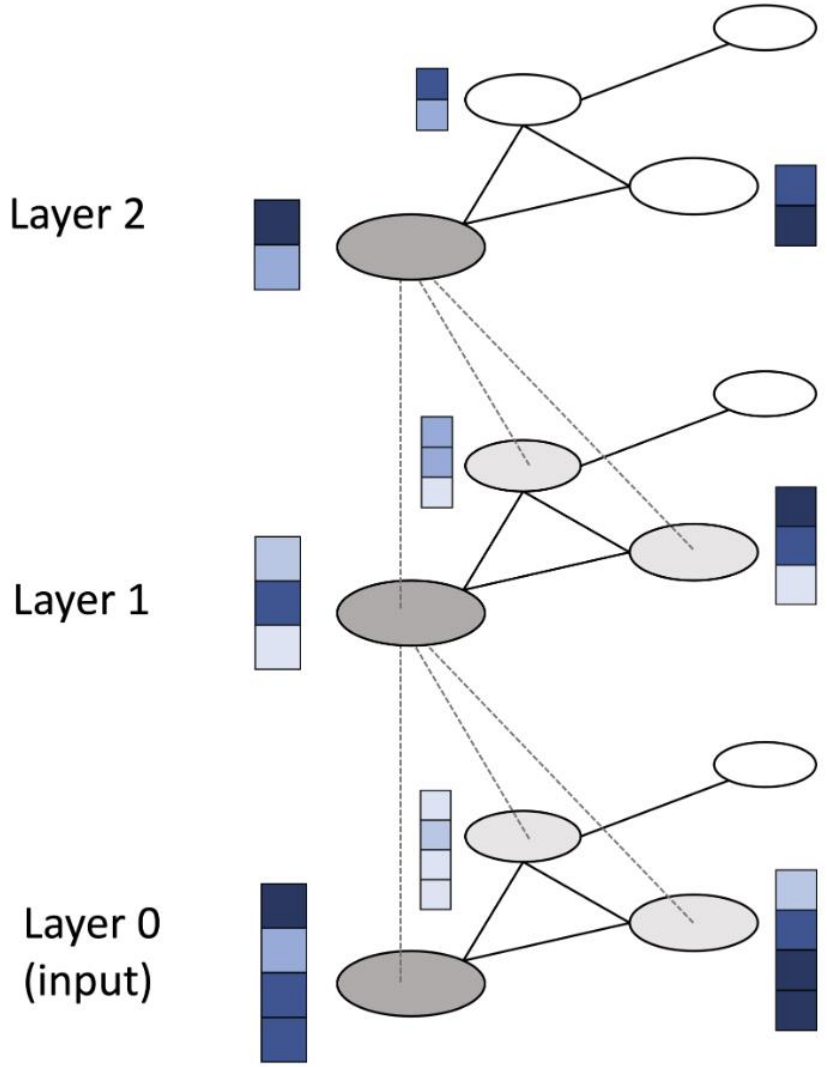
- PTransE实验效果

Metric	Mean Rank		Hits@10 (%)	
	Raw	Filter	Raw	Filter
RESCAL	828	683	28.4	44.1
SE	273	162	28.8	39.8
SME (linear)	274	154	30.7	40.8
SME (bilinear)	284	158	31.3	41.3
LFM	283	164	26.0	33.1
TransE	243	125	34.9	47.1
TransH	212	87	45.7	64.4
TransR	**198**	77	48.2	68.7
TransE (Our)	205	63	47.9	70.2
PTransE (ADD, 2-step)	200	**54**	**51.8**	83.4
PTransE (MUL, 2-step)	216	67	47.4	77.7
PTransE (RNN, 2-step)	242	92	50.6	82.2
PTransE (ADD, 3-step)	207	58	51.4	**84.6**

Lin, Yankai, et al. "Modeling Relation Paths for Representation Learning of Knowledge Bases." Proceedings of EMNLP. 2015.

使用图神经网络的方法

- R-GCN模型
- 基于图卷积神经网络（GCN）
  - 图神经网络（GNN）基于图结构为节点获取向量表示，再将向量表示用于具体任务
  - 在神经网络的第l层，为每个节点基于它本身及它的邻居的第$l-1$层向量表示计算一个新的向量



Schlichtkrull, Michael, et al. "Modeling relational data with graph convolutional networks." The semantic web: 15th international conference, ESWC 2018.

使用图神经网络的方法

- R-GCN模型

第l+1层向量表示计算：

$$h_i^{(l+1)} = \sigma\left(\sum_{r\in\mathcal{R}}\sum_{j\in\mathcal{N}_i^r}\frac{1}{c_{i,r}}W_r^{(l)}h_j^{(l)} + W_0^{(l)}h_i^{(l)}\right)$$

$\mathcal{N}_i^r$：节点i通过关系r可到达的邻居节点

其中$c_{i,r}$用于正则化 $c_{i,r} = |\mathcal{N}_i^r|$

$W_r^{(l)}$：第l层对应关系r的参数矩阵

使用图神经网络的方法

- R-GCN模型
- 使用最后一层的节点向量作为实体的最终向量表示 $e_i = h_i^{(L)}$
- 三元组打分函数（DistMult）：

$$f(s,r,o) = e_s^T R_r e_o$$

使用图神经网络的方法

- R-GCN效果

	FB15k					WN18				
	MRR		Hits @			MRR		Hits @		
Model	Raw	Filtered	1	3	10	Raw	Filtered	1	3	10
LinkFeat		0.779			0.804		0.938			0.939
DistMult	0.248	0.634	0.522	0.718	0.814	0.526	0.813	0.701	0.921	0.943
R-GCN	0.251	0.651	0.541	0.736	0.825	0.553	0.814	0.686	0.928	0.955
R-GCN+	**0.262**	**0.696**	**0.601**	**0.760**	**0.842**	0.561	0.819	0.697	0.929	**0.964**
CP*	0.152	0.326	0.219	0.376	0.532	0.075	0.058	0.049	0.080	0.125
TransE*	0.221	0.380	0.231	0.472	0.641	0.335	0.454	0.089	0.823	0.934
HolE**	0.232	0.524	0.402	0.613	0.739	**0.616**	0.938	0.930	**0.945**	0.949
ComplEx*	0.242	0.692	0.599	0.759	0.840	0.587	**0.941**	**0.936**	**0.945**	0.947

结合规则和神经网络的方法

- Neural Logic Programming (Neural LP)
- 为学习谓词逻辑规则集合设计了一个完全可导的系统
- 从而使得可以用基于梯度的优化方法来学习到逻辑规则

规则集合中每条规则形式：

$$\alpha \quad \mathrm{query}(Y, X) \leftarrow R_n(Y, Z_n) \wedge \cdots \wedge R_1(Z_1, X)$$

其中α为对该规则的confidence

Yang, Fan, Zhilin Yang, and William W. Cohen. "Differentiable learning of logical rules for knowledge base reasoning." Advances in NIPS (2017).

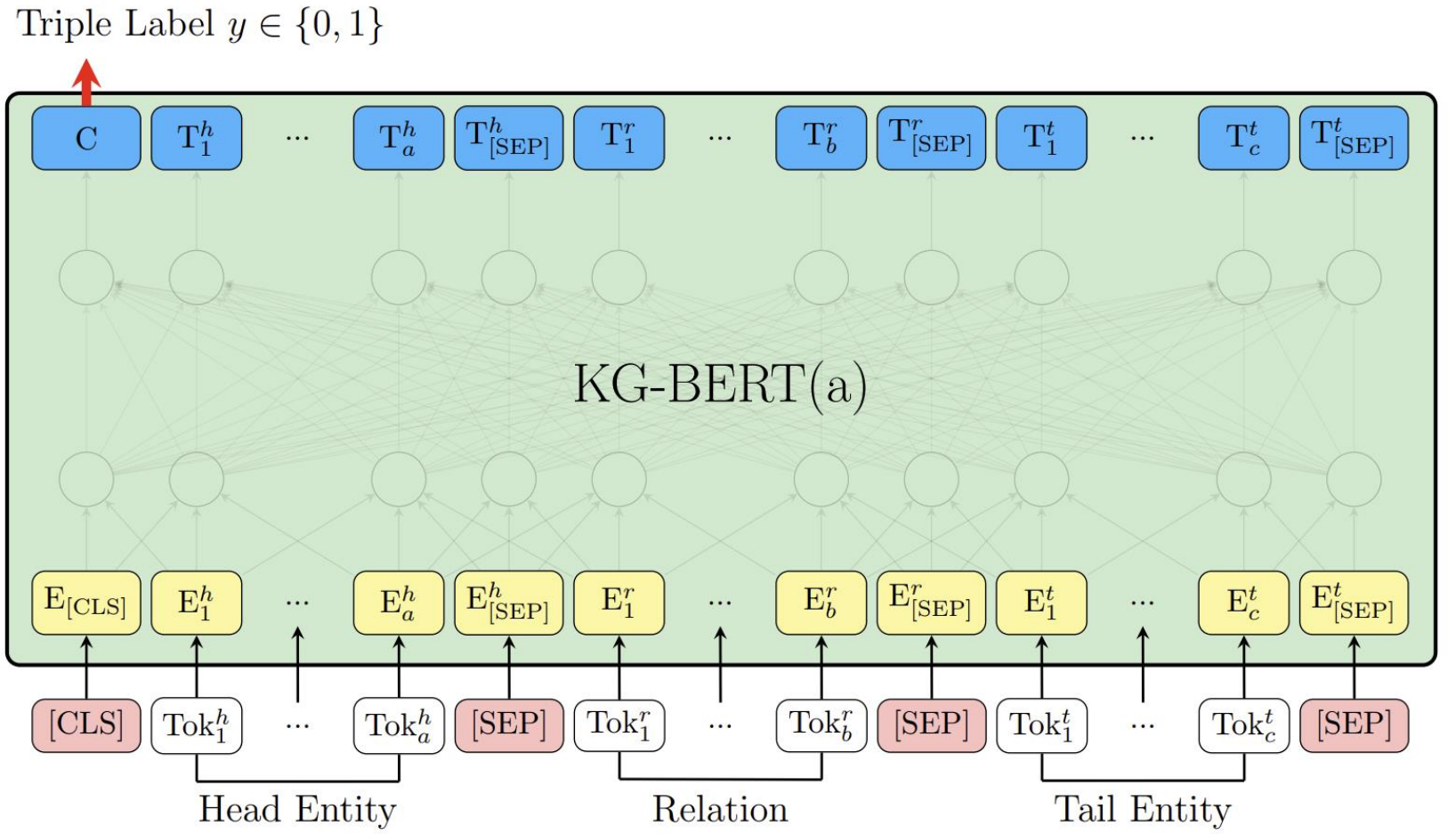
结合规则和神经网络的方法

- Neural Logic Programming (Neural LP)

	WN18		FB15K		FB15KSelected	
	MRR	Hits@10	MRR	Hits@10	MRR	Hits@10
Neural Tensor Network	0.53	66.1	0.25	41.4	-	-
TransE	0.38	90.9	0.32	53.9	-	-
DISTMULT [29]	0.83	94.2	0.35	57.7	**0.25**	**40.8**
Node+LinkFeat [25]	0.94	94.3	**0.82**	87.0	0.23	34.7
Implicit ReasoNets [23]	-	**95.3**	-	**92.7**	-	-
Neural LP	**0.94**	94.5	0.76	83.7	0.24	36.2

基于预训练模型的方法

- 基于BERT的方法



Yao, Liang, Chengsheng Mao, and Yuan Luo. "KG-BERT: BERT for knowledge graph completion." arXiv preprint arXiv:1909.03193 (2019).

基于预训练模型的方法

- 效果

Method	WN18RR		FB15k-237		UMLS	
	MR	Hits@10	MR	Hits@10	MR	Hits@10
TransE (our results)	2365	50.5	223	47.4	1.84	98.9
TransH (our results)	2524	50.3	255	48.6	1.80	**99.5**
TransR (our results)	3166	50.7	237	51.1	1.81	99.4
TransD (our results)	2768	50.7	246	48.4	1.71	99.3
DistMult (our results)	3704	47.7	411	41.9	5.52	84.6
ComplEx (our results)	3921	48.3	508	43.4	2.59	96.7
ConvE (Dettmers et al. 2018)	5277	48	246	49.1	–	–
ConvKB (Nguyen et al. 2018a)	2554	52.5	257	51.7	–	–
R-GCN (Schlichtkrull et al. 2018)	–	–	–	41.7	–	–
KBGAN (Cai and Wang 2018)	–	48.1	–	45.8	–	–
RotatE (Sun et al. 2019)	3340	**57.1**	177	**53.3**	–	–
KG-BERT(a)	**97**	52.4	**153**	42.0	**1.47**	99.0

END