

# 融合结构与属性注意力机制的实体对齐<sup>①</sup>

李忠阳, 王淑莹, 蒋 敏

(西南交通大学 计算机与人工智能学院, 成都 611756)

通信作者: 王淑莹, E-mail: [zy\\_l546@my.swjtu.edu.cn](mailto:zy_l546@my.swjtu.edu.cn)



**摘 要:** 在实现不同来源的图谱数据融合过程中, 实体对齐是关键步骤, 其目的在于确定不同图谱间等价的实体对. 现有实体对齐方法大多基于图嵌入方式, 通过考虑图谱的结构和属性信息进行对齐, 但并未很好处理二者之间的交互关系, 同时忽略对关系及多阶邻居信息的利用. 为解决上述问题, 提出一种融合结构与属性注意力机制模型 (fused structural and attribute attention mechanism model, FSAAM) 的实体对齐方法. 该模型首先根据图谱数据特征划分为属性和结构通道数据, 其次使用属性注意力机制实现对属性信息的学习, 在实现对结构信息的学习中增加对关系信息的学习, 利用图注意力机制寻找对于实体对齐有益的邻居特征, 引入 Transformer 编码器更好的关联实体之间的信息, 并通过 Highway 网络减少可能学习到噪声信息的影响, 最后对学习到的结构通道和属性通道信息的相似度矩阵利用 LS-SVM 网络, 得到集成相似度矩阵从而实现实体对齐. 所提模型在公开数据集 DBP15K 的 3 个子数据集上进行验证. 实验结果表明, 相较于基线模型中效果最好的结果, 其 *Hits@1* 分别提高了 2.7%, 4.3% 和 1.7%, 且 *Hits@10* 和 *MRR* 也均有提升, 表明本模型能够有效提高实体对齐的准确性.

**关键词:** 知识图谱; 实体对齐; 图注意力网络; Transformer 编码器; Highway 网络

引用格式: 李忠阳, 王淑莹, 蒋敏. 融合结构与属性注意力机制的实体对齐. 计算机系统应用, 2024, 33(6): 58–69. <http://www.c-s-a.org.cn/1003-3254/9542.html>

## Entity Alignment Integrating Structure and Attribute Attention Mechanism

LI Zhong-Yang, WANG Shu-Ying, JIANG Min

(School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China)

**Abstract:** For fusing map data from different sources, entity alignment is a key step, and its purpose is to determine equivalent entity pairs between different maps. Most of the existing entity alignment methods are based on graph embedding, which aligns by considering the structure and attribute information of the graph. However, they do not handle the interactive relationship between the two well and ignore the use of relationships and multi-order neighbor information. To solve the above problems, this study proposes an entity alignment method based on the fused structural and attribute attention mechanism model (FSAAM). The model first divides the graph data characteristics into attribute and structural channel data and then uses the attribute attention mechanism to learn attribute information. It adds the learning of relationship information to that of structural information and uses the graph attention mechanism to find the entities aligned with beneficial neighbor features. The Transformer encoder is introduced to better correlate information between entities, and the Highway network is utilized to reduce the impact of noise information that may be learned. Finally, the model applies the LS-SVM network to the similarity matrix of the learned structural channel and the attribute channel information, obtaining the integrated similarity matrix to achieve entity alignment. The proposed model is verified on three sub-datasets of the public data set DBP15K. Experimental results show that compared with the best results in the

① 基金项目: 国家重点研发计划 (2020YFB1708000); 四川省重大科技专项 (2022ZDZX0003)

收稿时间: 2023-12-14; 修改时间: 2024-01-17; 采用时间: 2024-02-23; csa 在线出版时间: 2024-04-28

CNKI 网络首发时间: 2024-05-07

baseline model, its  $Hits@1$  has increased by 2.7%, 4.3%, and 1.7% respectively, and  $Hits@10$  and  $MRR$  have also improved, indicating that this model can effectively improve the alignment accuracy of entities.

**Key words:** knowledge graph; entity alignment; graph attention network (GAT); Transformer encoder; Highway network

由于知识图谱具有语义表示和推理等功能,因此近年来相关技术快速发展,出现了大量知识图谱<sup>[1]</sup>.不同图谱通常会包含许多互补信息,将这些图谱进行融合有助于提高对知识的利用率.然而不同来源的知识图谱之间存在着异构性和数据冗余等问题.需要利用知识融合技术,对图谱中冗余信息进行对齐、合并,形成全局统一的知识表示和关联<sup>[2]</sup>.实体对齐(entity alignment, EA)是知识融合过程中关键技术,旨在关联不同来源知识图谱中指向真实世界同一对象的等价/匹配实体<sup>[3]</sup>,实现知识图谱间信息共享,从而提供一个信息较为完备的知识库.

早期,研究人员使用字符串的各种特征进行实体对齐.近年来随着知识表示学习技术的快速发展,研究人员提出许多基于表示学习的实体对齐方法,利用机器学习技术将图谱数据表示为低维稠密的向量,通过向量之间的距离来反映对象之间的语义关系<sup>[4]</sup>.

## 1 相关工作

早期的知识图谱实体对齐技术主要采用相似度计算<sup>[5]</sup>,人工特征提取<sup>[6]</sup>等方法,但是这些方法难以有效地匹配异构知识图谱,限制其应用范围.目前,随着表示学习的兴起,根据知识图谱建模方式的不同,可以分为:(1)基于三元组表征模型的翻译模型;(2)基于图结构特征的图神经网络模型.

### 1.1 翻译模型

以 TransE<sup>[7]</sup>为代表的翻译模型,对于每个三元组( $h, r, t$ ),将关系看作头尾实体的连接,并且关系存在方向,来学习实体的向量表示,该方法孤立学习每个三元组,无法感知更多的全局信息. MTransE<sup>[8]</sup>模型创造性地将 TransE 这一经典的翻译模型运用于知识图谱实体对齐任务; BootEA<sup>[9]</sup>模型针对缺乏足够的训练数据,迭代地将可能对齐的数据作为训练数据来增强模型效果,实现实体对齐; SEA<sup>[10]</sup>中指出图谱实体的度差异会对具有不同度数的实体造成不同的匹配难度,因此通过结合翻译模型和对抗训练方法提升对齐效果.翻译模型存在依赖三元组的局部语义信息,缺乏对知识图

谱全局信息的有效捕获的问题,因此研究人员将研究重点转移到基于图神经网络的实体对齐方法上.

### 1.2 图神经网络模型

基于图神经网络的方法主要利用图卷积网络(graph convolutional network, GCN)<sup>[11]</sup>和图注意力网络(graph attention network, GAT)<sup>[12]</sup>对节点邻居信息进行聚合,获得对齐实体之间的邻居相似度从而实现实体对齐,但是此类方式并未考虑真实世界中不同关系的影响,不能满足异构图的需要. HGCN<sup>[13]</sup>通过获取基于头尾实体的关系谓词嵌入来进行辅助对齐; RDGCN<sup>[14]</sup>通过探索原始图谱和对偶关系图之间丰富的关系特征进行实体对齐; HyperKA<sup>[15]</sup>通过引入基于层次结构的双曲关系图神经网络(hyperbolic relational GNN),使得能够以更低的实体嵌入向量实现卓越的匹配效果; HMAN<sup>[16]</sup>在利用 GCN 嵌入结构信息的同时加入了实体标签、拓扑信息、关系类型和文本描述等信息综合进行实体嵌入; AttrGNN<sup>[17]</sup>按照属性和结构信息进行通道划分,分别进行嵌入学习,其取得的效果要优于一些联合使用实体属性和结构信息进行对齐的模型,但是该模型在实体对齐过程中忽略了对关系信息和多阶邻居信息的考虑.因此本文在 AttrGNN 模型的基础上提出了 FSAAM 模型,本文的主要贡献如下.

(1) 利用双通道模块实现对实体属性和结构信息的联合学习.

(2) 考虑了在实体对齐任务中不同类型的关系信息对实体的重要度并不相同的影响.

(3) 利用图注意力机制与 Transformer 编码器共同编码的方式实现了多阶邻居信息聚合从而获得更佳的实体嵌入表示用于实体对齐任务.

因此本文具有的实际应用价值主要在于实现了对图谱中结构、关系、属性等多种信息的学习,从而在真实公开数据集上取得较好的效果.

## 2 问题定义

定义 1. 知识图谱可以被形式化定义为一个六元组  $G = (E, R, A, V, T^r, T^a)$ , 其中  $E$ 、 $R$ 、 $A$ 、 $V$  分别表示图

谱中的实体 (entities)、关系 (relations)、属性 (attributes)、属性值 (values) 的集合;  $T^r = \{(h, r, t) | h, t \in E, r \in R\}$  表示关系三元组的集合;  $T^a = \{(e, a, v) | e \in E, a \in A, v \in V\}$  表示属性三元组的集合<sup>[17]</sup>.

定义 2. 实体对齐就是找到源知识图谱 (Gs) 和目标知识图谱 (Gt) 之间的映射, 即为  $\psi = \{(e, e') | e \in E, e' \in E'\}$ , 其中的  $e$  与  $e'$  为两个图谱中等价的实体. 等价实体的种子集  $\psi^s$  被用作训练数据<sup>[17]</sup>.

### 3 融合结构与属性注意力机制的实体对齐模型

FSAAM 模型框架如图 1 所示, 它主要由 3 个模块

组成: (1) 属性通道模块: 通过属性注意力机制实现对图谱中属性信息的利用, 从而得到属性通道相似度矩阵; (2) 结构通道模块: 通过 GraphSAGE-RelationAware 聚合层实现对图谱中邻居和关系信息的利用, 再通过 GAT-Transformer 多阶邻居聚合层实现对图谱中多阶邻居信息的利用, 从而得到结构通道相似度矩阵; (3) 双通道特征融合模块: 是为了同时利用上述两个通道模块学习到的有用信息, 为两个相似度矩阵分配不同的比重, 从而得到集成相似度矩阵实现实体对齐.

#### 3.1 结构通道模块

结构通道实现流程步骤如图 2 所示.

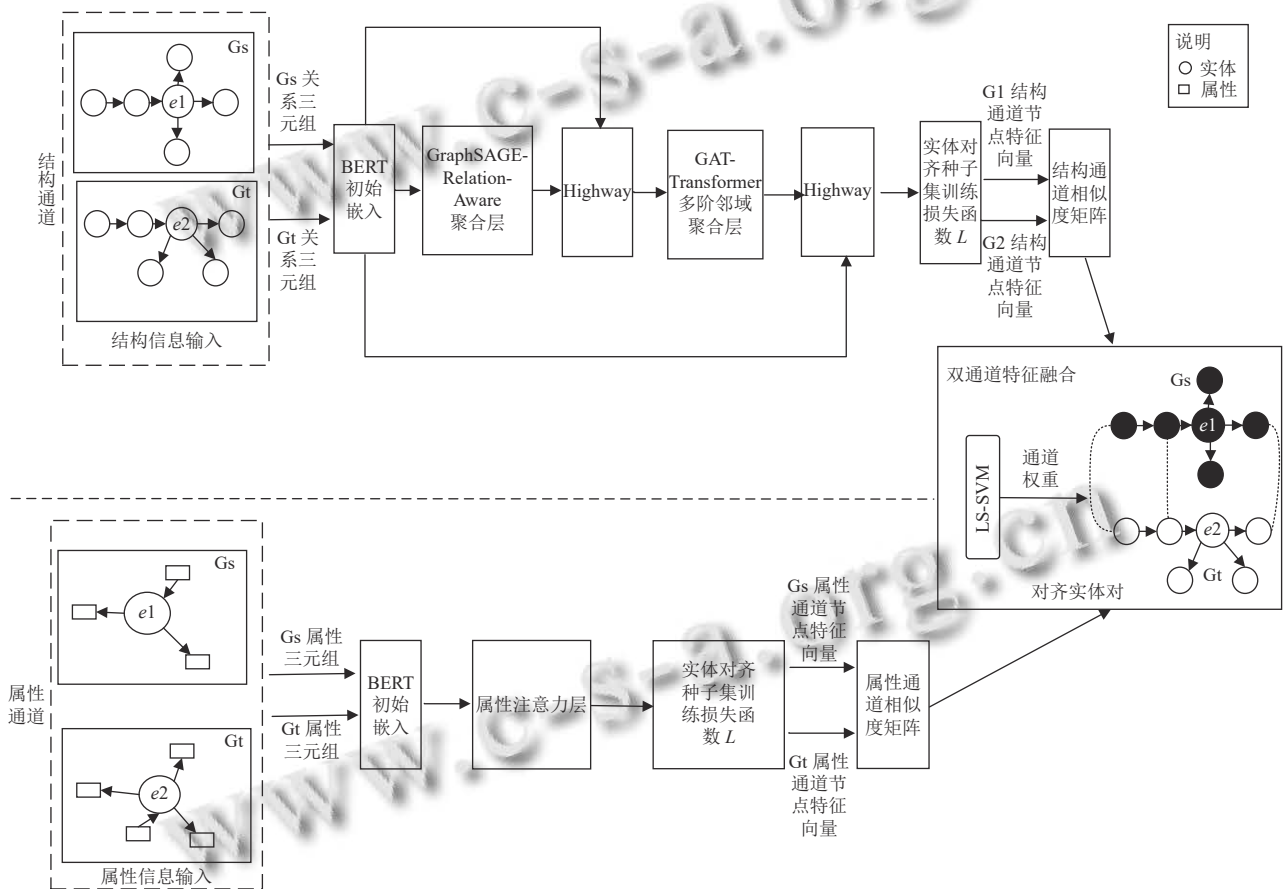


图 1 FSAAM 模型框架

输入层: 实体名称对于实体对齐是非常重要的信息, 使用预训练的词嵌入模型对实体的语义信息进行编码, 使得初始特征不再是随机初始化, 减少模型训练成本. BERT<sup>[18]</sup>在大规模的多语言数据集上进行训练, 已经捕捉了大规模文本数据中的语言特征, 能够得到较好的初始实体嵌入表示, 故本文利用 BERT 实现对实体名称进行编码获得初始实体特征表示  $h_e^{(0)}$  (表示

第 0 层的实体表示).

GraphSAGE-RelationAware 聚合层: 该层搭建的主要就是为了学习知识图谱中的邻居以及关系信息.

为了利用邻居的特征来生成实体特征表示, 依据 Alinet<sup>[19]</sup>中所得结论, 一阶邻居有着更小的异构性, 因此采用均值聚合器 (graph sample and aggregate, GraphSAGE)<sup>[20]</sup>方式实现对信息聚合. 其计算公式如下:

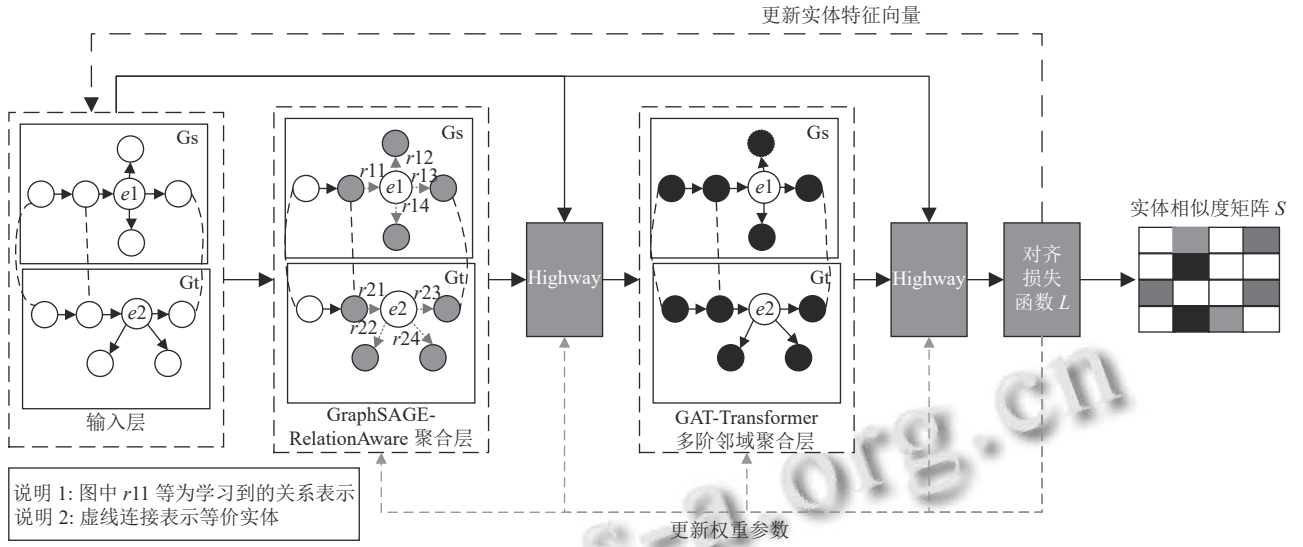


图2 结构通道实现流程步骤

$$h_e^l = \text{ReLU}(W^l \text{MEAN}(\{h_e^{l-1}\} \cup \{h_j^{l-1}, \forall j \in N(e)\})) \quad (1)$$

其中,  $W^l$  表示可学习的权重矩阵;  $\text{MEAN}(\cdot)$  表示对输入向量求平均值;  $h_e^l$  表示顶点为  $e$  的第  $l$  层的嵌入表示。

$N(e) = \{j | \forall (j, r, e) \in T^r \text{ or } \forall (e, r, j) \in T^r, \forall r \in R\}$  表示  $e$  的邻居实体。

在实体对齐任务中, 关系反映了实体间特殊的作用关系, 不同类型的关系对于同一实体而言其贡献度不同, 因此在实体的表示学习过程中增加对实体间关联关系的学习. 受 RAGA<sup>[21]</sup> 启发, 模型首先使用实体注意力机制计算关系在头实体作用下的关系表示  $r_k^h$  以及在尾实体作用下的关系表示  $r_k^t$  (二者计算过程类似), 通过将  $r_k^h$  和  $r_k^t$  相加得到关系的向量表示  $r_k$ . 其计算公式如下:

$$a_{ijk} = \frac{\exp(\text{LeakyReLU}(\beta_1^T [h_i W^{h1}; h_j W^{t1}]))}{\sum_{e_i \in H} \sum_{e_j \in T} \exp(\text{LeakyReLU}(\beta_1^T [h_i W^{h1}; h_j W^{t1}]))} \quad (2)$$

$$r_k^h = \text{ReLU} \left( \sum_{e_i \in H} \sum_{e_j \in T} a_{ijk} h_i W^{h1} \right) \quad (3)$$

$$r_k = r_k^h + r_k^t \quad (4)$$

其中,  $a_{ijk}$  表示基于头实体  $e_i$  和尾实体  $e_j$  作用下从头实体  $e_i$  到关系  $r_k$  的注意力系数;  $H$  表示关系  $r_k$  连接的头实体的集合,  $T$  表示头实体  $e_i$  和关系  $r_k$  对应的尾实体的集合;  $\beta_1$  表示用于降维的向量;  $W^{h1}$  表示头实体作用下关系的权重矩阵,  $W^{t1}$  表示尾实体作用下关系的权重矩阵。

在获得关系的向量表示  $r_k$  后需要将关系信息重新聚合到实体表示中, 因此利用关系注意力机制分别计算实体节点作为关系的头节点和作为关系的尾节点的嵌入表示. 计算公式如下所示:

$$a_{ik} = \frac{\exp(\text{LeakyReLU}(\beta_2^T [h_i; r_k]))}{\sum_{e_j \in T_{e_i}} \sum_{r_k \in R_{e_i e_j}} \exp(\text{LeakyReLU}(\beta_2^T [h_i; r_k]))} \quad (5)$$

$$h_i^{\text{in}} = \text{ReLU} \left( \sum_{e_j \in T_{e_i}} \sum_{r_k \in R_{e_i e_j}} a_{ik} r_k \right) \quad (6)$$

其中,  $a_{ik}$  表示从关系  $r_k$  到实体  $e_i$  的注意力系数;  $T_{e_i}$  表示头实体  $e_i$  对应的尾实体集合,  $R_{e_i e_j}$  表示头实体  $e_i$  和尾实体  $e_j$  之间的关系集合;  $\beta_2$  表示用于降维的向量;  $h_i^{\text{in}}$  表示实体  $e_i$  作为关系的头节点的嵌入表示 (说明: 实体  $e_i$  作为关系的尾节点的嵌入表示  $h_i^{\text{out}}$  的计算过程类似)。

在计算不同方向关系作用下的实体表示后, 将输入层得到的初始实体特征表示  $h_e^{(0)}$ 、 $h^{\text{in}}$ 、 $h^{\text{out}}$ 、 $h_e$  (表示经过均值聚合器得到的输出) 进行拼接得到新的实体表示, 公式如下所示:

$$h_{r_e} = [h_e^{(0)}; h^{\text{in}}; h^{\text{out}}; h_e] \quad (7)$$

**Highway 层:** 为了控制 GraphSAGE-RelationAware 聚合层中可能学习到的噪声信息传播的影响, 受 RDGCN 启发, 使用 Highway 网络<sup>[22]</sup> 来控制初始实体特征表示  $h_e^{(0)}$  与前阶段学习到的特征表示  $h_{r_e}$  之间的信息平衡. 计算公式如下所示:



$$h_{\text{highway}1} = G(h_e^{(0)}) \cdot h_{r\_e} + (1 - G(h_e^{(0)})) \cdot h_e^{(0)} \quad (8)$$

$$G(h_e^{(0)}) = \text{Sigmoid}(h_e^{(0)} W_{h1} + b_{h1}) \quad (9)$$

其中,  $h_{\text{highway}1}$  表示第 1 次通过 Highway 网络的输出,  $W_{h1}$  表示 Highway 网络所需权重矩阵,  $b_{h1}$  表示偏置向量.

GAT-Transformer 多阶邻居聚合层: 为进一步增强实体的特征信息, 需要实现对多阶邻居信息的考虑.

因此首先使用图注意力机制寻找到对于对齐实体有积极作用的邻居. 计算公式如下:

$$a_{ij} = \frac{\exp(\text{LeakyReLU}(\beta_3^T [Wh_i; Wh_j]))}{\sum_{j' \in N_i} \exp(\text{LeakyReLU}(\beta_3^T [Wh_i; Wh_{j'}]))} \quad (10)$$

$$h_{ner_i} = \left[ h_{\text{highway}1_i}; \text{ReLU} \left( \sum_{j \in N_i} a_{ij} \beta_3^T h_j \right) \right] \quad (11)$$

其中,  $a_{ij}$  表示从实体  $e_j$  到实体  $e_i$  的注意力系数;  $\beta_3$  表示用于降维的向量;  $W$  表示图注意力机制所需权重矩阵;  $h_{ner_i}$  表示实体  $e_j$  经过图注意力机制与经过 Highway 层输出的拼接的特征向量.

针对图注意力机制过度关注局部信息, 无法充分利用图的整体拓扑结构的缺点, 在图注意力机制网络后添加 Transformer 编码器对图中信息再次进行编码. 计算公式如下所示:

$$h_{te} = \text{LayerNorm}(h_{ner} + A(h_{ner} W_q, h_{ner} W_k, h_{ner} W_v) + \text{FFN}(h_{ner})) \quad (12)$$

$$A(h_{ner} W_q, h_{ner} W_k, h_{ner} W_v) = \text{Softmax} \left( \frac{h_{ner} W_q \cdot (h_{ner} W_k)^T}{\sqrt{d_k}} \right) \cdot (h_{ner} W_v) \quad (13)$$

$$\text{FFN}(h_{ner}) = \max(0, h_{ner} W_{f1} + b_{f1}) W_{f2} + b_{f2} \quad (14)$$

其中,  $\text{LayerNorm}$  表示 Layer Normalization 对输入进行归一化处理;  $W_k, W_v, W_q$  表示自注意力机制所需权重矩阵;  $d_k$  表示注意力头中的特征维度 (解释: 是通过经图注意力机制处理得到的  $h_{ner}$  维度处以注意力的数量计算得到的);  $W_{f1}, W_{f2}$  表示前馈神经网络所需权重矩阵,  $b_{f1}, b_{f2}$  是对应偏置.

Highway 层: 同前一个阶段 Highway 层作用一样是为了减少学习到的噪声数据, 计算得到  $h_{\text{highway}2}$ .

对齐损失函数: 根据预对齐的实体种子集训练损失函数, 将结构通道的两个子图中的实体嵌入到同一向量空间, 并在训练过程中更新 GraphSAGE-Relation-

Aware 聚合层、GAT-Transformer 多阶邻居聚合层以及 Highway 层中的模型参数, 获取更新后的实体特征向量, 基于更新的实体特征向量得到结构通道相似度矩阵  $s^1$ . 计算公式如下所示:

$$L = \sum_{(e, e') \in \psi^s} \sum_{(e_{\text{negative}}, e'_{\text{negative}}) \in \psi^{s'}} \max(\text{dis}(e, e') - \text{dis}(e_{\text{negative}}, e'_{\text{negative}}) + \lambda, 0) \quad (15)$$

$$\text{dis}(e, e') = \sum_{i=1}^n |x'_e - x'_{e'}| = \|x'_e - x'_{e'}\|_1 \quad (16)$$

其中,  $(e, e')$  表示预对齐的种子实体对;  $\psi^s$  表示实体预对齐的种子对集合;  $(e_{\text{negative}}, e'_{\text{negative}})$  表示负采样的负例种子对;  $\psi^{s'}$  表示通过负采样得到的实体对的集合;  $\text{dis}(e, e')$  表示曼哈顿距离 (又称作 L1 范数), 是用来求向量中元素的绝对值之和;  $x'_e$  和  $x'_{e'}$  表示结构通道中两个子图中实体的最终向量表示;  $\lambda$  为超参数.

### 3.2 属性通道模块

属性通道实现流程步骤如图 3 所示.

输入层: 同结构通道一样, 给定属性三元组  $\{(e, a_1, v_1), \dots, (e, a_n, v_n)\}$  作为某个实体  $e$  的输入, 利用 BERT 对实体名称、属性序列以及属性值序列进行初始学习, 从而分别得到实体的初始特征表示  $h_e^{(0)}$ , 该实体对应的属性特征序列为  $\{\text{attr}_1, \dots, \text{attr}_n\}$  和属性值特征序列为  $\{\text{value}_1, \dots, \text{value}_n\}$ .

属性注意力层: 为了学习到对于实体特征表示更加有意义的属性信息, 通过将  $\{a_1, \dots, a_n\}$  作为边,  $\{v_1, \dots, v_n\}$  作为节点, 连接  $\{v_1, \dots, v_n\}$  和  $e$ , 那么就可以将实体的注意力计算应用到属性信息的计算, 使用注意力分数来计算不同属性和属性值的加权平均值. 计算公式如下所示:

$$h_e^1 = \text{ELU} \left( h_e^{(0)} + \sum_{j=1}^n a_j W_1 [\text{attr}_j; \text{value}_j] \right) \quad (17)$$

其中,  $h_e^1$  表示经属性注意力层输出的特征表示,  $W_1 \in \mathbb{R}^{D_{h1} \times (D_a + D_v)}$  是可学习的参数矩阵,  $D_{h1}, D_a, D_v$  分别表示  $h_e^1, a_i, v_i$  的向量维度,  $a_j$  表示公式如下:

$$a_j = \frac{\exp(\text{LeakyReLU}(\beta^T [h_e^{(0)}; \text{attr}_j]))}{\sum_{k=1}^n \exp(\text{LeakyReLU}(\beta^T [h_e^{(0)}; \text{attr}_k]))} \quad (18)$$

其中,  $n$  表示与实体相关的属性数量,  $\beta \in \mathbb{R}^{(D_e + D_a) \times 1}$  是可学习的参数矩阵,  $D_e$  表示  $h_e^{(0)}$  的向量维度.

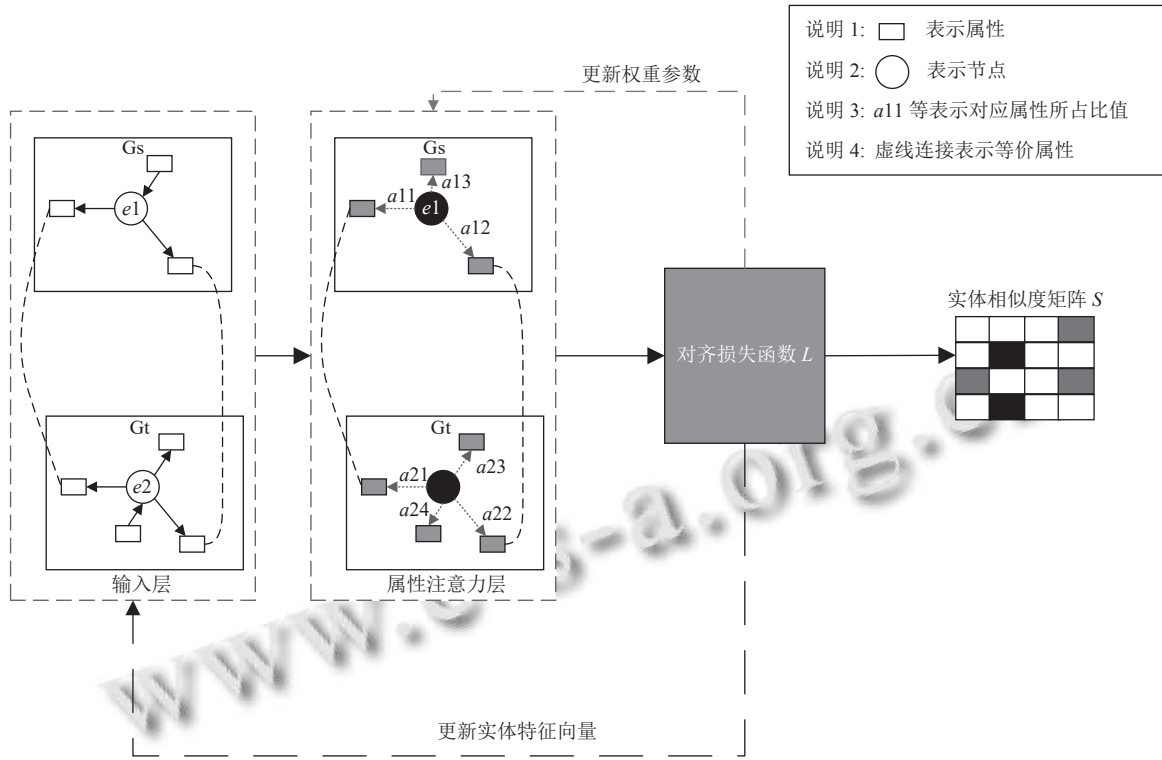


图3 属性通道实现流程步骤

然后在获得实体属性特征表示  $h_e^l$  后利用式 (15) 中的损失函数计算损失更新网络, 得到属性通道实体相似度矩阵  $s^2$ 。

### 3.3 双通道特征融合模块

分别从结构通道和属性通道获取各自对应的实体相似度矩阵  $S^k$  ( $k \in \{1, 2\}$ ) 后, 为同时利用前文学习到的结构和属性信息, 受 AttrGNN 模型启发搭建 LS-SVM 网络得到集成的相似度矩阵  $S^* = w_1 S^1 + w_2 S^2$ , 其中  $w = [w_1, w_2]$  通过如下公式学习得到:

$$L_{svm} = C \sum_{l=1}^m [y_l \cdot \max(0, 1 - w^T x_l) + (1 - y_l) \cdot \max(0, 1 + w^T x_l)] + \frac{1}{2} w^T w \quad (19)$$

其中,  $C$  是正则化参数, 用于控制拟合误差和模型复杂性之间的平衡,  $x_l = [S_{e,e'}^1, S_{e,e'}^2]$  是相似度分数组成的向量, 如果  $(e, e') \in \phi_s$ , 即为  $G_s$  和  $G_t$  中等价的实体对时  $y_l$  取 1, 否则取 0。

具体实现是通过调用 Scikit-learn 库中 SVC 函数, 采用线性核, 视为二分类问题 (解释: 将  $G_t$  和  $G_s$  中等价的实体对标签设置为 1, 随机生成固定数量的不等价的实体标签设置为 0) 来搭建 LS-SVM 网络进行训练,

优化式 (19) 中目标函数  $L_{svm}$ , 从而求出结构通道和属性通道各自相似度矩阵所占比重信息。

## 4 实验结果及分析

### 4.1 数据集

本文采用跨语言数据集 DBP15K<sup>[23]</sup> 作为实验数据集, 该数据集是大规模知识图谱 DBPedia<sup>[24]</sup> 的子集. 其中包含 ZH-EN (中-英)、JA-EN (日-英)、FR-EN (法-英), 每个数据集在两个不同语言的知识图谱中包含 15 000 个对齐实体对, 作为本文实验的训练、验证和测试集. DBP15K 中每个数据集的具体统计信息如表 1 所示。

表1 数据集统计

数据集	关系三元组	文字三元组	数值三元组	
ZH-EN	汉语	153k	290k	177k
	英语	237k	292k	203k
JA-EN	日语	164k	227k	152k
	英语	233k	268k	171k
FR-EN	法语	192k	313k	162k
	英语	278k	323k	227k

### 4.2 基线模型

为了评估提出的模型有效性, 本文选择一些具有竞争性的实体对齐模型进行比较。

BootEA<sup>[9]</sup>: 2018年发表于IJCAI国际会议, 针对训练集有限导致的嵌入表示不够准确的问题, 该模型使用一种迭代的策略标记训练数据, 从而提高实体嵌入的准确性。

HGCN<sup>[13]</sup>: 2019年发表于EMNLP国际会议, 该模型根据实体表示得出关系的表示, 在实体表示中融入关系的信息, 利用关系的信息促进实体对齐。

RDGCN<sup>[14]</sup>: 2019年发表于IJCAI国际会议, 探索通过对原始图和对偶关系图之间的交互进行建模, 利用Highway门控GCN层捕获邻居结构, 以学习更好的实体表示。

AttrGNN<sup>[17]</sup>: 2020年发表于EMNLP国际会议, 基于图神经网络方法联合学习实体结构和属性嵌入, 来实现实体对齐。

RNM<sup>[25]</sup>: 2021年发表于AAAI国际会议上, 利用邻居匹配和从连接关系中探索有用的信息实现实体对齐。

DATTI<sup>[26]</sup>: 2022年发表于ACL会议上, 提出了一种基于三阶张量同构(DATTI)的高效实体对齐解码算法, 有效地利用了知识图谱中的邻接和自身现象来增强实体对齐的解码过程。

NAMN<sup>[27]</sup>: 2022年发表于计算机工程, 提出依据每跳邻居对中心实体重要性不同的特点, 采用分层的思想区别处理每跳邻居信息, 通过门控机制进行聚会以学习图结构的表征。

STEA<sup>[28]</sup>: 2023年发表于ACM, 提出利用实体之间的依赖关系, 来进行自我训练实现实体对齐。

### 4.3 评价指标

本文采用实体对齐模型中常用的评价指标:  $Hits@N$  以及平均倒数排名 ( $MRR$ ) 作为评价指标。 $Hits@N$  是通过排名在前  $N$  个的正确对齐实体的比例来进行计算,  $MRR$  是指所有正确实体的平均倒数排名。这两个指标越高, 表明实体对齐模型效果越好。其公式分别计算如下:

$$Hits@N = \frac{\sum_{e \in Seed} (rank_e \leq N ? 1 : 0)}{Seed} \quad (20)$$

$$MRR = \frac{1}{|N|} \sum_{i=1}^{|N|} \frac{1}{rank_e} \quad (21)$$

其中,  $Seed$  表示测试实体对;  $rank_e$  表示与实体  $e$  对齐的实体预测排名;  $N$  为对齐实体对。

## 4.4 实验设置与环境

### 4.4.1 实验设置

为利用实体名称信息, 首先使用谷歌翻译将数据集中的中文、法语以及日文的实体信息都翻译成英文, 再实现节点特征预嵌入。

模型训练过程中使用 30% 的种子实体对作为训练集, 其余 70% 实体对作为测试集。使用 BERT 构建初始实体嵌入从而使得结构通道和属性通道得到的嵌入维度为 768 的向量。采用 Adagrad 优化器, 通过网格搜索寻找结构和属性通道实体对齐的最佳参数配置。具体参数设置如表 2 所示。

表 2 参数设置

模型参数	参数值
Highway网络层数	2
图注意机制的层数	1
Transformer encoder的层数	1
学习率	{0.0001, 0.0005, 0.001, 0.004, 0.007}
L2正则化	{ $0, 10^{-3}, 10^{-4}$ }
随机数种子	3407
通道训练epoch数	200

### 4.4.2 实验环境

实验环境基于 Windows 10 专业版, 采用的语言为 Python, 具体环境如表 3 所示。

说明 1: 在 torch-geometric 库中封装了包括图注意力机制在内的各种图神经网络算法;

说明 2: 利用 Scikit-learn 搭建 LS-SVM 网络。

表 3 实验环境

项目	环境
操作系统	Windows 10专业版
CPU	Intel(R) Xeon(R) Gold 6338 CPU @ 2.00 GHz 2.00 GHz (2 个处理器)
GPU	NVIDIA A40-48Q
Python版本	3.8.16
PyTorch版本	1.12.0
torch-geometric版本	2.3.1
Scikit-learn	1.2.2

## 4.5 实验结果

### 4.5.1 参数设置与影响

本文在结构通道和属性通道模块的实验通过网格搜索方式最佳的 L2 正则化和学习率参数的值对评价指标  $Hits@1$  的影响。图 4—图 6 展示了结构通道模块中 DBP15K 的 3 个子集  $Hits@1$  的变化情况。

取 3 个子集对应的  $Hits@1$  效果最佳时的学习率

和 L2 正则化参数, 其结构通道模块中 DBP15K 的 3 个子集损失收敛曲线对应如图 7 所示。

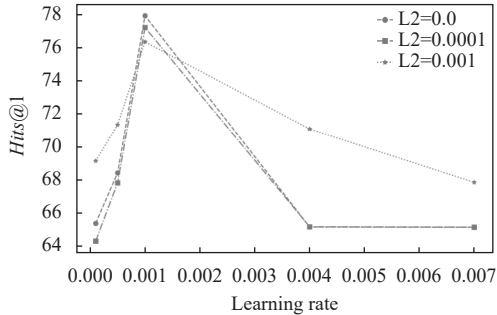


图 4 ZH-EN 数据集结构通道 Hits@1 变化

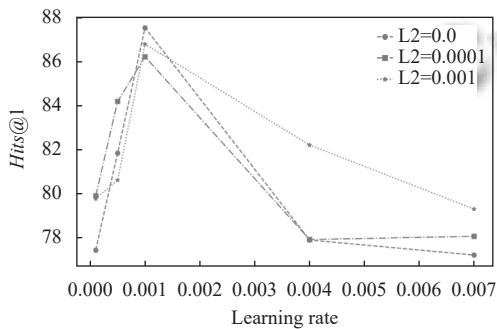


图 5 JA-EN 数据集结构通道 Hits@1 变化

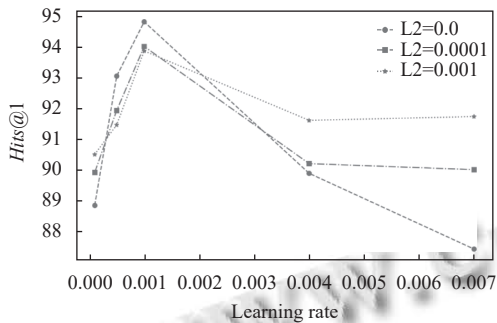


图 6 FR-EN 数据集结构通道 Hits@1 变化

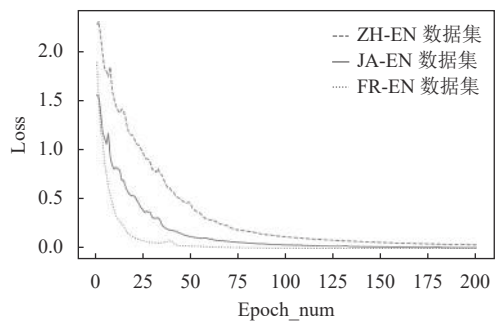


图 7 DBP15K 子集结构通道损失收敛曲线图

图 8-图 10 展示了属性通道模块中 DBP15K 的 3 个子集 Hits@1 的变化情况。

取 3 个子集对应的 Hits@1 效果最佳时的学习率和 L2 正则化参数, 其属性通道模块中 DBP15K 的 3 个子集损失收敛曲线对应如图 11 所示。

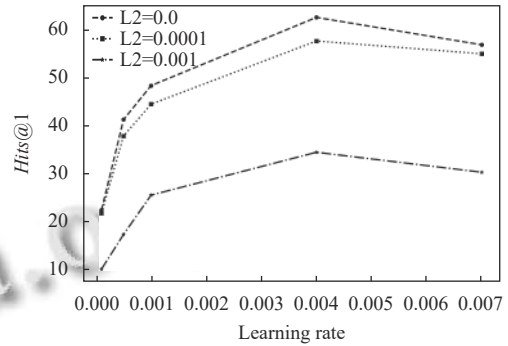


图 8 ZH-EN 数据集属性通道 Hits@1 变化

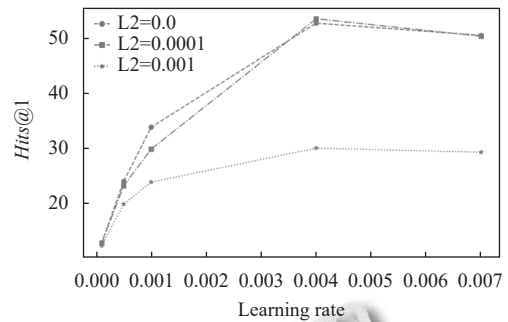


图 9 JA-EN 数据集属性通道 Hits@1 变化

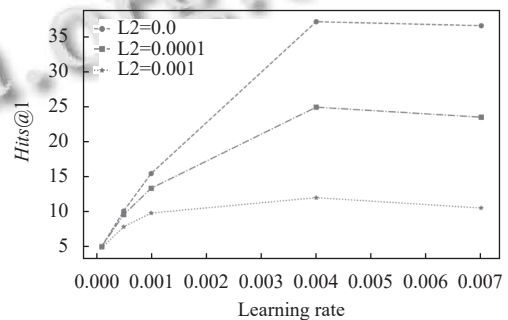


图 10 FR-EN 数据集属性通道 Hits@1 变化

对上述实验现象进行分析: 对于结构通道其 Hits@1 最佳的学习率为 0.001 和正则化参数 L2 为 0 时, 且 3 个子集数据在训练过程中损失函数的值都较为平滑的下降并最终保持稳定; 而对于属性通道而言学习率为 0.004, 正则化参数 L2 为 0 时对应的 Hits@1 最佳, 并 3 个子集数据在训练过程的中间阶段其损失函数都有较大的波动情况, 在训练轮数达到 175 次后趋于稳



定. 因此后续第 4.5.2 节以及第 4.5.3 节均采用最佳学习率及正则化参数进行实验.

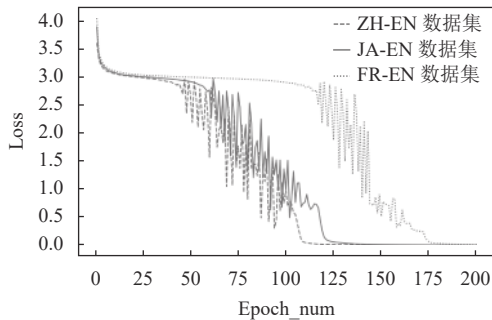


图 11 DBP15K 子集属性通道损失收敛曲线图

#### 4.5.2 对比实验

本文所提模型和各个基线模型对比结果如表 4 所示, 其中 BootEA、RDGCN、HGNC 均来自 RNM 论文中实验结果, AttrGNN 通过实验验证得到, DATTI、NAMN 以及 STEA 分别来自其各自对应论文中实验结果, 并将各个基线模型中效果最好的进行加粗标注.

BootEA 利用迭代的策略标记训练数据进行实体对齐, 但是并未考虑不同图谱间存在多对一, 多对多以及一对多实体的情况; RDGCN 通过与对偶关系图之间

的交互实现对关系的考虑, HGNC 利用实体名称嵌入表示作为实体初始嵌入并融入关系, 但是二者都未能考虑多阶邻居信息以及属性信息的影响; AttrGNN 通过通道划分实现了对结构信息和属性信息的考虑, 但是未能引入关系信息和多阶邻居信息; RNM 通过使用关系感知邻居匹配和实体感知匹配模块进行关系和实体对齐, 但未能考虑属性信息对于异构图谱的影响; DATTI 专注于改进实体对齐解码过程, 没有对现有实体编码过程进行改进; NAMN 通过对一阶邻居和多阶邻居信息的考虑并采用邻居匹配的方式进行实体对齐, 但未能考虑关系信息和属性信息的作用; STEA 利用实体之间的依赖关系进行实体对齐, 这种半监督学习方式其准确率往往不佳. 本文针对上述模型中存在的问题进行优化, 实验结果表明, 本文所提模型其 *Hits@1* 分别在 3 个数据集上相较于表格中基线模型加粗标注的结果分别提升 2.7%, 4.3%, 1.7%, 且 *Hits@10* 和 *MRR* 也均有一定提升. 证明本文所提模型实现了对图谱中属性、邻居和关系等多信息的综合学习, 利用双通道学习图谱中不同信息, 从而有助于从潜在的实体集合中找出正确的实体, 提高实体对齐的命中率.

表 4 不同模型的实体对齐结果对比

模型	ZH-EN			JA-EN			FR-EN		
	<i>Hits@1</i> (%)	<i>Hits@10</i> (%)	<i>MRR</i>	<i>Hits@1</i> (%)	<i>Hits@10</i> (%)	<i>MRR</i>	<i>Hits@1</i> (%)	<i>Hits@10</i> (%)	<i>MRR</i>
BootEA	62.9	84.8	0.703	62.2	85.4	0.701	65.3	87.4	0.731
RDGCN	70.8	84.6	0.746	76.7	89.5	0.812	88.6	95.7	0.911
HGNC	72.0	85.7	0.768	76.6	89.7	0.813	89.2	96.1	0.917
AttrGNN	79.60	92.93	0.845	78.33	92.08	0.834	91.85	97.77	0.910
RNM	84.0	91.9	0.870	<b>87.2</b>	94.4	<b>0.899</b>	<b>93.8</b>	<b>98.1</b>	0.954
DATTI	83.5	<b>95.3</b>	0.880	83.6	<b>96.9</b>	0.884	87.3	97.9	0.913
NAMN	76.8	89.4	0.821	79.2	93.6	0.894	92.9	97.4	<b>0.964</b>
STEA	<b>85.4</b>	94.2	<b>0.885</b>	84.3	95.0	0.882	87.5	96.8	0.909
FSAAM	<b>88.1</b>	<b>97.1</b>	<b>0.915</b>	<b>91.5</b>	<b>98.3</b>	<b>0.940</b>	<b>95.5</b>	<b>99.0</b>	<b>0.970</b>

#### 4.5.3 消融实验

为了证明本文搭建的结构通道实现了对邻居信息和关系信息的有效利用, 以及将图谱信息进行通道划分的有效性, 本文通过消融实验进行对比分析.

##### (1) 结构通道消融实验

为了验证本文搭建的结构通道对实体对齐结果的影响, 本文使用不同的策略来设计 FSAAM 的变体. FSAAM-0 表示结构通道模块保持第 3.1 节中所有网络层的情况; FSAAM-1 表示在引入图注意力机制后不添加 Transformer 编码器; FSAAM-2 表示在一阶邻居信息和关系信息拼接后不添加 Highway 网络; FSAAM-3

表示在经过 Transformer 编码器学习后不添加 Highway 网络; FSAAM-4 表示不引入图注意力机制; FSAAM-5 表示不考虑关系信息. 上述变体在进行实验时, 均保持其他实验条件不改变, 其实验结果对比如表 5、表 6 所示, 其中表 5 为结构通道消融实验结果对比, 表 6 为双通道特征融合消融实验结果对比.

FSAAM-1 在表 5 中的 3 个数据集的 *Hits@1* 平均下降 3.4%, 在表 6 中的 3 个数据集的 *Hits@1* 平均下降 1.1%, 证明了本文在利用图注意力机制实现引入图注意力机制后添加 Transformer 编码器实现对全局特征的考虑对于实体对齐结果的有效性; FSAAM-2 在

表5中的3个数据集的  $Hits@1$  平均下降 0.97%, 在表6中的3个数据集的  $Hits@1$  平均下降 0.83%, FSAAM-3在表5中的3个数据集的  $Hits@1$  平均下降 32%, 在表6中的3个数据集的  $Hits@1$  平均下降 19%, 证明为了避免误差累积引入 Highway 网络能够很好平衡实体初始特征和学习了关系邻居信息的特征信息, 并通过 FSAAM-2 与 FSAAM-3 实验结果对比, 搭建的网络模型深度越深误差累积可能会越多, 引入 Highway 网络

效果越明显; FSAAM-4 在表5中的3个数据集的  $Hits@1$  平均下降 7.4%, 在表6中的3个数据集的  $Hits@1$  平均下降 0.83%, 证明为缓解结构异质性从而利用注意力机制实现对更有益于实体对齐任务邻居信息考虑的有效性; FSAAM-5 在表5中的3个数据集的  $Hits@1$  平均下降 3.6%, 在表6中的3个数据集的  $Hits@1$  平均下降 3.1%, 证明在实体对齐任务中增加对关系信息的考虑对于最终实体对齐效果的提升。

表5 结构通道消融实验结果对比

模型 (结构通道结果作为最终输出结果)	ZH-EN			JA-EN			FR-RN		
	$Hits@1$ (%)	$Hits@10$ (%)	MRR	$Hits@1$ (%)	$Hits@10$ (%)	MRR	$Hits@1$ (%)	$Hits@10$ (%)	MRR
FSAAM-1	72.9	85.6	0.775	84.1	93.1	0.874	93.0	97.7	0.948
FSAAM-2	76.7	88.9	0.811	87.1	95.3	0.901	93.5	98.6	0.960
FSAAM-3	38.8	59.7	0.460	61.8	83.9	0.695	63.4	83.0	0.705
FSAAM-4	68.3	78.9	0.721	79.7	88.2	0.828	90.0	94.8	0.918
FSAAM-5	73.3	86.6	0.781	83.5	93.1	0.869	92.6	98.0	0.947
FSAAM-0	<b>77.9</b>	<b>90.3</b>	<b>0.824</b>	<b>87.5</b>	<b>95.7</b>	<b>0.905</b>	<b>94.8</b>	<b>98.9</b>	<b>0.963</b>

表6 双通道特征融合消融实验结果对比

模型	ZH-EN			JA-EN			FR-RN		
	$Hits@1$ (%)	$Hits@10$ (%)	MRR	$Hits@1$ (%)	$Hits@10$ (%)	MRR	$Hits@1$ (%)	$Hits@10$ (%)	MRR
FSAAM-1	86.1	95.7	0.907	90.7	97.84	0.934	95.0	98.9	0.966
FSAAM-2	87.0	97.2	0.914	90.5	98.2	0.940	95.1	99.1	0.967
FSAAM-3	70.4	89.8	0.774	75.9	93.64	0.823	71.7	91.5	0.789
FSAAM-4	86.8	96.4	0.904	90.9	97.5	0.933	94.9	98.8	0.964
FSAAM-5	84.7	95.5	0.887	87.9	96.7	0.912	93.2	98.5	0.952
FSAAM	<b>88.1</b>	<b>97.1</b>	<b>0.915</b>	<b>91.5</b>	<b>98.3</b>	<b>0.940</b>	<b>95.5</b>	<b>99.0</b>	<b>0.970</b>

## (2) 通道划分消融实验

为证明采用通道划分的方式相较于直接将实体结构和属性信息拼接利用的作用, 提出相应变体. FSAAM-6 表示只考虑结构通道信息进行实体对齐, FSAAM-7 表示只考虑属性通道信息进行实体对齐, FSAAM-8 表示不通过双通道特征融合模块, 直接将结构通道获取的结构实体表示和属性通道获取的属性实体表示进行向量拼接, 得到增强实体表示进行实体对齐. 上述变体在进行实验时, 均保持其他实验条件不改变, 其实验结果对比如表7所示。

通过表7可知 FSAAM-8 直接将属性和结构信息

进行拼接得到的实体对齐效果相比与 FSAAM-6 只考虑结构通道信息在3个数据集上的  $Hits@1$  平均下降 6.37%, 证明直接将不同类型的信息拼接进行实体对齐的效果并不理想, 相反不恰当的使用信息会带来噪声干扰导致实体对齐效果下降; 同时 FSAAM-6 在表7中的3个数据集相较于 FSAAM 的  $Hits@1$  平均下降 4.97%, FSAAM-7 在表7中的3个数据集相较于 FSAAM 的  $Hits@1$  平均下降 39.7%, 证明单独考虑图谱中的属性或结构信息获得的实体对齐的结果并不佳, 需要充分考虑图谱中的属性、结构和关系等多种信息从而获取更佳的实体表示用于实体对齐任务。

表7 通道划分消融实验结果对比

模型	ZH-EN			JA-EN			FR-RN		
	$Hits@1$ (%)	$Hits@10$ (%)	MRR	$Hits@1$ (%)	$Hits@10$ (%)	MRR	$Hits@1$ (%)	$Hits@10$ (%)	MRR
FSAAM-6	77.9	90.3	0.824	87.5	95.7	0.905	94.8	98.9	0.963
FSAAM-7	62.5	82.3	0.694	56.3	73.1	0.622	37.3	54.8	0.433
FSAAM-8	71.9	84.7	0.764	80.2	89.4	0.836	89.0	94.13	0.909
FSAAM	<b>88.1</b>	<b>97.1</b>	<b>0.915</b>	<b>91.5</b>	<b>98.3</b>	<b>0.940</b>	<b>95.5</b>	<b>99.0</b>	<b>0.970</b>

## 5 结束语

针对现有实体对齐模型并未很好地处理结构和属

性之间的交互关系, 同时忽略对关系信息及多阶邻居信的利用的问题, 本文提出一种融合结构与属性注意

力机制的实体对齐方法. 该方法通过对属性和结构信息的划分实现了对图谱中属性、邻居和关系信息等多方面信息的利用. 本文模型的  $Hits@1$ ,  $Hits@10$  以及  $MRR$  指标均优于上述基线模型, 证明了本文所提模型的有效性. 但由于现有实体对齐效果依赖于已对齐的实体种子集, 而在实际生产中已有的实体种子集数量较少, 使得现有实体对齐模型效果下降, 因此在下一步研究中将采用一些无监督或者半监督方式来实现实体对齐任务.

### 参考文献

- 1 赵晓娟, 贾焰, 李爱平, 等. 多源知识融合技术研究综述. 云南大学学报(自然科学版), 2020, 42(3): 459–473. [doi: 10.7540/j.ynu.20190481]
- 2 Hogan A, Blomqvist E, Cochez M, *et al.* Knowledge graphs. ACM Computing Surveys, 2022, 54(4): 71.
- 3 高云君, 葛丛丛, 郭宇翔, 等. 面向关系型数据与知识图谱的数据集成技术综述. 软件学报, 2023, 34(5): 2365–2391. [doi: 10.13328/j.cnki.jos.006808]
- 4 张富, 杨琳艳, 李健伟, 等. 实体对齐研究综述. 计算机学报, 2022, 45(6): 1195–1225. [doi: 10.11897/SP.J.1016.2022.01195]
- 5 Suchanek FM, Abiteboul S, Senellart P. PARIS: Probabilistic alignment of relations, instances, and schema. Proceedings of the VLDB Endowment, 2011, 5(3): 157–168. [doi: 10.14778/2078331.2078332]
- 6 Mahdisoltani F, Biega J, Suchanek FM. YAGO3: A knowledge base from multilingual Wikipedias. Proceedings of the 7th Biennial Conference on Innovative Data Systems Research. Asilomar, 2015.
- 7 Bordes A, Usunier N, Garcia-Duran A, *et al.* Translating embeddings for modeling multi-relational data. Proceedings of the 26th International Conference on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc., 2013. 2787–2795.
- 8 Chen MH, Tian YT, Yang MH, *et al.* Multilingual knowledge graph embeddings for cross-lingual knowledge alignment. Proceedings of the 26th International Joint Conference on Artificial Intelligence. Melbourne: AAAI Press, 2017. 1511–1517.
- 9 Sun ZQ, Hu W, Zhang QH, *et al.* Bootstrapping entity alignment with knowledge graph embedding. Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm: AAAI Press, 2018. 4396–4402.
- 10 Pei SC, Yu L, Hoehndorf R, *et al.* Semi-supervised entity alignment via knowledge graph embedding with awareness of degree difference. The World Wide Web Conference. San Francisco: ACM, 2019. 3130–3136.
- 11 Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. Proceedings of the 5th International Conference on Learning Representations. Toulon: OpenReview.net, 2017.
- 12 Veličković P, Cucurull G, Casanova A, *et al.* Graph attention networks. Proceedings of the 6th International Conference on Learning Representations. Vancouver: OpenReview.net, 2018.
- 13 Wu YT, Liu X, Feng YS, *et al.* Jointly learning entity and relation representations for entity alignment. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Hong Kong: ACL, 2019. 240–249.
- 14 Wu YT, Liu X, Feng YS, *et al.* Relation-aware entity alignment for heterogeneous knowledge graphs. Proceedings of the 28th International Joint Conference on Artificial Intelligence. Macao: IJCAI.org, 2019. 5278–5284.
- 15 Sun ZQ, Chen MH, Hu W, *et al.* Knowledge association with hyperbolic knowledge graph embeddings. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). ACL, 2020. 5704–5716.
- 16 Yang HW, Zou YY, Shi P, *et al.* Aligning cross-lingual entities with multi-aspect information. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Hong Kong: ACL, 2019. 4431–4441.
- 17 Liu ZY, Cao YX, Pan LM, *et al.* Exploring and evaluating attributes, values, and structures for entity alignment. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). ACL, 2020. 6355–6364.
- 18 Devlin J, Chang MW, Lee K, *et al.* BERT: Pre-training of deep bidirectional Transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Minneapolis: ACL, 2019. 4171–4186.
- 19 Sun ZQ, Wang CM, Hu W, *et al.* Knowledge graph alignment network with gated multi-hop neighborhood aggregation. Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington: AAAI Press, 2020.

- 222–229.
- 20 Hamilton WL, Ying Z, Leskovec J. Inductive representation learning on large graphs. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 1025–1035.
- 21 Zhu RB, Ma M, Wang P. RAGA: Relation-aware graph attention networks for global entity alignment. Proceedings of the 25th Pacific-Asia Conference on Knowledge Discovery and Data Mining. Cham: Springer, 2021. 501–513.
- 22 Srivastava RK, Greff K, Schmidhuber J. Highway networks. arXiv:1505.00387, 2015.
- 23 Sun ZQ, Hu W, Li CK. Cross-lingual entity alignment via joint attribute-preserving embedding. Proceedings of the 16th International Semantic Web Conference. Vienna: Springer, 2017. 628–644.
- 24 Lehmann J, Isele R, Jakob M, *et al.* DBpedia—A large-scale, multilingual knowledge base extracted from Wikipedia. Semantic Web, 2015, 6(2): 167–195. [doi: [10.3233/SW-140134](https://doi.org/10.3233/SW-140134)]
- 25 Zhu Y, Liu HZ, Wu ZH, *et al.* Relation-aware neighborhood matching model for entity alignment. Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington: AAAI Press, 2021. 4749–4756.
- 26 Mao X, Ma MR, Yuan H, *et al.* An effective and efficient entity alignment decoding algorithm via third-order tensor isomorphism. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Dublin: ACL, 2022. 5888–5898.
- 27 谭元珍, 李晓楠, 李冠宇. 基于邻域聚合的实体对齐方法. 计算机工程, 2022, 48(6): 65–72.
- 28 Liu B, Lan TC, Hua W, *et al.* Dependency-aware self-training for entity alignment. Proceedings of the 16th ACM International Conference on Web Search and Data Mining. Singapore: ACM, 2023. 796–804.

(校对责编: 张重毅)