计 算 机 研 究 与 发 展
Journal of Computer Research and Development
https://crad.ict.ac.cn
63(3): 812−823, 2026

基于双仿射配对和层级标注的联合实体关系抽取

张清华[1,2,3] 周 雄[1,2] 廖 伟[1,2] 黄帅帅[1,2] 秦徐婷[1,2]
[1](计算智能重庆市重点实验室(重庆邮电大学) 重庆 400065)
[2](大数据智能计算重点实验室(重庆邮电大学) 重庆 400065)
[3](旅游多源数据感知与决策技术文化和旅游部重点实验室(重庆邮电大学) 重庆 400065)
(zhangqh@cqupt.edu.cn)

Joint Entity and Relation Extraction Based on Biaffine Pairing and Cascade Tagging

Zhang Qinghua[1,2,3], Zhou Xiong[1,2], Liao Wei[1,2], Huang Shuaishuai[1,2], and Qin Xuting[1,2]
[1](*Chongqing Key Laboratory of Computational Intelligence* (*Chongqing University of Posts and Telecommunications*), *Chongqing* 400065)
[2](*Key Laboratory of Big Data Intelligent Computing* (*Chongqing University of Posts and Telecommunications*), *Chongqing* 400065)
[3](*Key Laboratory of Tourism Multisource Data Perception and Decision Ministry of Culture and Tourism* (*Chongqing University of Posts and Telecommunications*), *Chongqing* 400065)

Abstract Joint entity and relation extraction is a foundational task for knowledge graph construction, which aims at extracting relational triples from unstructured text. To address the issues of the redundant annotations in corresponding matrix and insufficient interactions between subjects and objects, a joint model based on biaffine entity pairing and cascade tagging scheme is proposed. First, the potential relations in the sentence are predicted by performing a multi-label classification task, which eliminates the redundant relations in the relation-specific entity identification module. Then, the integrated candidate entity representations are passed through a biaffine network to enhance the interactions between subjects and objects, resulting in an entity pairing matrix that contains only candidate entities and reduces redundant annotations in the entity pairing phase. Next, a cascade tagging scheme is applied to identify relation-specific entities, and the identified entities are combined with the entity pairing matrix to form relational triples. Finally, the effectiveness of the proposed model is validated by comparative experiments and ablation study conducted on four public datasets. The experimental results demonstrate the proposed model achieves significant performance improvements over current baseline methods in terms of standard evaluation metrics, while effectively mitigating both the redundant annotation problem and the issue of insufficient entity interactions.

Key words relation extraction; potential relation; biaffine network; cascade tagging; joint entity and relation extraction

摘 要 联合实体关系抽取作为知识图谱构建的基础任务之一,旨在从非结构化文本中提取出关系三元组。针对联合模型中存在的相关矩阵冗余标注和头尾实体交互不足2个问题,提出了一种基于双仿射实体配对和层级标注策略的联合实体关系抽取模型。首先,通过一个多标签分类任务来预测句子中的潜在关系,从而减少特定关系实体识别阶段的冗余关系。其次,将整合的候选实体表示通过一个双仿射网络以增强头尾实体的交互并形成双仿射实体配对矩阵,从而减少实体配对阶段的冗余标注。然后,使用层

收稿日期: 2024-04-26; **修回日期**: 2025-12-10
基金项目: 国家自然科学基金项目(62276038); 重庆市自然科学基金创新发展联合基金项目(CSTB2023NSCQ-LZX0164); 重庆英才名家名师项目(CQYC20210202215)
This work was supported by the National Natural Science Foundation of China (62276038), the Joint Fund of Chongqing Natural Science Foundation for Innovation and Development (CSTB2023NSCQ-LZX0164), and the Chongqing Talent Program (CQYC20210202215).

级标注策略识别出特定关系的实体，并结合实体配对矩阵形成关系三元组。最后，通过在4个公共数据集上进行对比实验和消融实验，验证了所提模型的有效性。

关键词 关系抽取；潜在关系；双仿射网络；层级标注；联合实体关系抽取

中图法分类号 TP389.1

DOI: 10.7544/issn1000-1239.202440289 **CSTR:** 32373.14.issn1000-1239.202440289

实体关系抽取是自然语言处理的基础任务，同时也是信息抽取的关键任务，在知识图谱构建[1]、搜索引擎优化[2]以及生物信息挖掘[3]中都有广泛应用。其主要目标是从非结构化文本中抽取出关系三元组，通常表示为(头实体，关系，尾实体)。早期的实体关系抽取主要采取流水线方法[4-5]，即先识别句子中存在的所有实体，再通过关系分类判别实体对之间的关系。然而，流水线方法普遍存在2个缺陷，即误差传播和实体关系交互弱。因此，近年来的工作主要围绕联合抽取模型进行研究，其主要思想是将实体抽取和关系抽取在训练阶段联合学习，通过共享编码器参数增强2个阶段的交互，从而提升抽取效果。近几年来，实体关系抽取模型主要采用预训练语言模型作为句子编码器，比如先通过BERT[6]预训练语言模型得到句子的嵌入表示，再设计下游模块进行微调。尽管联合实体关系抽取的研究已经取得不错的效果，但仍然存在3个方面的问题：

1)关系三元组重叠问题。根据关系三元组中实体和关系的重叠方式[7-8]可以将关系三元组重叠问题分为3类，如表1所列。正常(Normal)三元组中的头尾实体和其他关系三元组的头尾实体是不同的，不存在重叠；单个实体重叠(single entity overlap, SEO)三元组中的头实体或尾实体和其他关系三元组的头实体或尾实体存在重叠；实体对重叠(entity pair overlap, EPO)中的头尾实体和句子中其他关系三元组的头尾实体存在重叠。关系三元组重叠问题是联合实体关系抽取研究的重点难点之一。

Table 1 Cases of Normal, Single Entity Overlap and Entity Pair Overlap Triples

表1 正常、单个实体重叠以及实体对重叠三元组示例

模式	句子	三元组
Normal	[上海]位于[中国]	(上海，位于，中国)
SEO	[特朗普]出生于[纽约][皇后区]	(特朗普，出生于，纽约) (特朗普，出生于，皇后区) (皇后区，位于，纽约)
EPO	[北京]是[中国]的首都	(北京，首都，中国) (北京，位于，中国)

2)头尾实体识别交互弱的问题。该问题指在抽取三元组过程中头实体和尾实体识别过程的语义交互不强。现有的联合抽取模型采取多阶段联合训练的方式，通常是先识别出头实体，再将头实体语义嵌入和关系语义嵌入融合到句子嵌入以标注出相应的尾实体，即头实体和尾实体需要分别抽取。因此，头尾实体的语义交互不足会导致前一阶段的误差传播到后一阶段，最终影响三元组抽取精度。

3)冗余标注问题。该问题指在实体识别或头尾实体交互过程中，由于模型设计缺陷导致的大量无意义标注。首先，一些联合关系抽取模型在识别尾实体的过程中，需要遍历预定义关系以标注对应的尾实体，即无论句子中是否存在此关系，都要遍历关系标注特定于关系的实体。其次，在构建头尾实体全局相关矩阵中，大量非实体的词也需要进行标签标注。

为了解决上述问题，本文提出了一种基于双仿射实体配对和层级标注策略的联合实体关系抽取模型BEPCTS。模型整体分为四大模块，即潜在关系预测、候选实体识别、双仿射实体配对和特定于关系的实体识别，采用多阶段联合训练的方式进行实体关系联合抽取。首先，使用多标签分类的方式得到句子中的潜在关系。其次，进行候选实体识别得到句子中的命名实体，并将每个实体整合嵌入表示。通过双仿射网络增强头尾实体的交互，建立双仿射实体配对矩阵，实体配对矩阵仅标注候选实体之间是否相关，不涉及句子中无关词的标注，因此减少了冗余标注。最后，将潜在关系的嵌入表示融合到句子嵌入表示以标注特定关系的实体，并结合双仿射实体配对矩阵得到最终的关系三元组。

本文的主要贡献包括3个方面：

1)设计了一个端到端联合关系抽取模型。将关系三元组抽取转化为层级标注问题和实体配对问题，可以有效解决关系三元组重叠问题。

2)提出了一种双仿射网络的实体配对策略。使用双仿射网络增强了头尾实体的交互，并仅针对候选实体建立实体配对矩阵，减少了实体配对过程中的冗余标注。

3)在4个公开数据集上的实验结果表明，所提出的模型不仅在三元组整体抽取中具有更好的性能表现，并且在缓解关系三元组重叠、冗余标注、头尾实体交互弱和减少冗余关系等问题上也有一定的优势。

1 相关工作

在早期研究中，关系三元抽取任务被划分为2个独立的阶段，即流水线方法。这种方法首先进行命名实体识别(named entity recognition, NER)，以提取所有可能的头实体和尾实体；然后进行关系分类(relation classification, RC)，以识别每对实体提及之间的关系。然而，流水线方法存在2个明显的缺点：1)误差传播问题，由于命名实体识别任务和关系分类任务是2个独立的任务，命名实体识别阶段的错误累积无法得到纠正，而被传递到影响RC阶段。2)忽略了2个任务之间的相互作用，即流水线方法中，实体识别和关系分类之间的语义交互不足。

为解决流水线方法的2个缺点，学术界提出了联合模型并受到了广泛关注。传统基于特征的联合模型需要复杂的特征工程并通过自然语言处理工具包进行手动特征提取。近年来，神经网络已广泛应用于联合关系三元抽取的方法中，减少了特征提取的手动构建过程。基于神经网络的联合模型可分为基于序列标注的联合模型[9-11]、基于片段抽取的联合模型[12-14]、基于表格填充的联合模型[15-19]和基于机器阅读理解的联合模型[19]。

1.1 基于序列标注的联合模型

Zheng等人[9]提出了一种新的标注框架(NovelTagging)进行联合抽取，该框架实现了联合解码并将三元抽取任务转化为序列标记任务；然而，该模型无法抽取重叠关系三元组。Wei等人[10]提出了一种层级二分类标注框架(CasRel)。该标注方法不同于以往命名实体使用的位置标注方法，其放弃了识别实体的种类，转而只识别实体的开始和结束位置。Zheng等人[11]提出了联合抽取模型(PRGC)，利用BERT的整体句子表示[CLS]来预测句子中的潜在关系。在实体识别阶段，只识别特定于潜在关系的候选头实体和尾实体；同时，为了学习候选头实体和尾实体是否存在相关关系，该模型建立了一个全局相关矩阵。尽管基于标注的联合模型取得了不错的效果，但仍然忽略了2个关键问题：1)在预测潜在关系时是否充分利用了句子的语义信息；2)头实体和尾实体的语义交互是否充足。

1.2 基于片段抽取的联合模型

基于片段的联合模型的主要思想是先罗列句子中所有可能的跨度组合，再通过构建合理的跨度嵌入表示进行分类。Dixit等人[12]提出了一个基于片段抽取的联合模型，其首先列举文本中所有可能的跨度集合，将跨度开始、跨度结束以及跨度长度的表示拼接作为跨度嵌入表示，然后将此嵌入表示通过分类器预测该跨度是否为一个实体。在关系分类阶段，将提取出的所有跨度实体两两配对，通过一个多标签分类模块判断关系是否存在和识别关系类别。Eberts等人[13]将预训练语言模型和跨度抽取思想结合，提出了SpERT模型，该模型使用BERT作为文本编码器，将跨度范围内的嵌入表示池化之后和BERT整体句子表示[CLS]拼接作为最终的跨度表示。在分类2个跨度实体之间是否存在关系时，将跨度对之间的表示池化之后和将跨度对表示拼接作为分类器的输入。Yang等人[14]基于片段抽取，提出了一个融合双向关系语义和知识引导的关系注意力网络。

1.3 基于表格填充的联合模型

基于表格填充的模型需要为每个预定义关系建立一个表，表的每一个行列组合表示一个单词对之间是否存在该关系。因此，该抽取范式的主要目标是填充每一个关系表，再根据关系表构建出关系三元组。Wang等人[15]将关系三元组抽取构建为一个单阶段的单词链接问题，提出了TPLinker。该模型将填表描述为一种握手标记方案，并缓解了联合模型中的误差传播和曝光偏差问题。Dai等人[16]认为现有的表格填充模型忽略了头尾实体的信息交互和关系三元组的内在联系，因此提出了一个包含头尾实体交互和推理路径抽取的表格填充模型。Shang等人[17]提出了一种基于填表的多模块单步解码模型，缓解了信息冗余和误差传播问题。Ning等人[18]将表格填充视为一项对象检测任务，提出了一种基于单阶段对象检测的关系三元组抽取模型。基于填表的方法在结构上相对基于标注的方法要复杂许多，且基于填表的联合抽取模型一般计算复杂度都比较高。

1.4 基于机器阅读理解的联合模型

为了利用句子中包含的丰富语义信息，一些联合抽取模型结合了机器阅读理解(machine reading comprehension, MRC)的思想。基于MRC的模型首先构建一个与头实体类型相关的查询，并同时从句子上下文中提取头实体。然后，通过将头实体上下文与关系上下文结合，构建提取尾实体的查询并在后续步骤中重复使用句子上下文。Li等人[19]将关系抽

取任务转化为多轮问答(multi-turn question answering, QA)问题，进而将特定于关系的实体识别转化为根据问题在句子中标注合适的片段。其主要思想是根据实体类别和关系类型构建合理的查询模板序列，并和文本序列结合，通过编码器表示之后进行答案标注。然而，与基于标注的模型相比，基于 MRC 的模型更为复杂，这是因为抽取任务转化为了一个 QA 问题，为了识别头尾实体需要多次进行查询。此外，基于 MRC 的模型在输入数据集包含大量预定义的实体类型和关系类型时，需要更多的查询来提取实体和关系。因此，这类模型不适用于关系类型较多的数据集。

2 模型设计

本文提出了一种基于双仿射实体配对和层级标注策略的联合实体关系抽取模型 BEPCTS。模型分为 4 个模块，即潜在关系预测、候选实体识别、双仿射实体配对和特定于关系的实体识别，采用多阶段联合训练进行实体关系联合抽取。

BEPCTS 的整体架构如图 1 所示，在图 1 的最下方是文本序列编码器，文本通过 BERT 预训练语言模型获得词嵌入表示。在图 1 的左下方是关系检测模块，通过多标签分类任务识别句子中的潜在关系，减少了在特定关系的实体识别阶段的冗余关系。在图 1 的右上方是基于双仿射网络的实体配对矩阵，通过双仿射网络可以增强头实体和尾实体的语义交互。在图 1 的左上方是特定于关系的实体识别，通过该模块标注特定关系的实体，然后和双仿射实体配对矩阵结合抽取出关系三元组。

2.1 编码器和输入层

给定一个句子 $S=\{t_1,t_2,\cdots,t_n\}$ 以及预定义关系集合 $R=\{r_1,r_2,\cdots,r_k\}$，其中 n 是句子 S 的词个数，k 是预定义关系个数。关系三元组抽取任务的目标是输出关系三元组：$T=\{(s,r,o)|s\in E_s,o\in E_o,r\in R\}$，其中 E_s 和 E_o 分别表示头实体集合和尾实体集合。

在 BEPCTS 模型中，句子的上下文表示通过 BERT 预训练语言模型获得。具体来讲，首先在长度为 n 的句子 S 前面加上 [CLS] 标识词，用于学习句子的整体语义嵌入。然后将词序列输入 BERT 得到句子的上下文表示，具体过程如式(1)所示。

$$\boldsymbol{H}=BERT(S)=[\boldsymbol{h}_{\text{cls}};\boldsymbol{h}_1;\boldsymbol{h}_2;\cdots;\boldsymbol{h}_n], \tag{1}$$

其中 $\boldsymbol{H}\in\mathbb{R}^{n\times d}$，$n$ 表示词序列的长度，d 表示通过 BERT 得到的词嵌入维度。

2.2 潜在关系预测

大多数联合模型都是遍历数据集中的预定义关系集合进行特定关系的实体识别，导致了关系冗余问题。因此，BEPCTS 模型首先预测句子中的潜在关系。通常，预训练语言模型都会在句子的开头处设置一个预留位置来学习句子的整体语义表示，因此

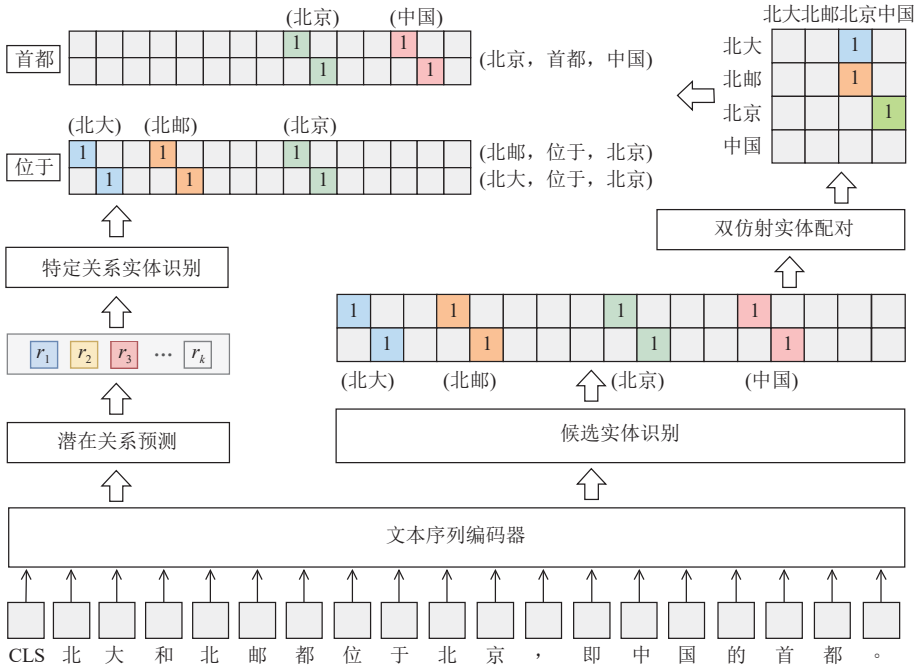


Fig. 1 Overall architecture of BEPCTS model

图 1 BEPCTS 模型整体架构

通过这个特殊的词嵌入表示可以进行相应的分类任务。具体来讲, 在 BEPCTS 中使用 [CLS] 表示进行多标签分类。首先使用潜在关系预测模块获得一个属于预定义关系集合的关系子集, 表示这个关系子集是当前句子可能存在的关系类型集合。比如, 在句子“北大和北邮都位于北京, 即中国的首都。”中, 预测的关系子集理论上可以包含“位于”和“首都”两个关系类型。在预测出相应的潜在关系之后, 后续特定于关系的实体识别仅针对潜在关系进行, 即减少了后续阶段的冗余关系。潜在关系预测的具体计算过程如式(2)所示。

$$p_{\text{rel}} = \sigma(\boldsymbol{W}_r \boldsymbol{h}_{\text{cls}} + \boldsymbol{b}_r), \tag{2}$$

其中 $\boldsymbol{W}_r \in \mathbb{R}^{d\times 1}$, $\boldsymbol{b}_r \in \mathbb{R}^d$ 分别表示可训练的参数矩阵和偏差, $\sigma(\cdot)$ 表示 sigmoid 函数。

潜在关系预测在 BEPCTS 模型中是一个多标签二分类任务, 因此, 该任务的损失函数可以使用二分类交叉熵损失进行优化, 如式(3)所示。

$$\mathcal{L}_{\text{rel}} = -\frac{1}{k}\sum_{i=1}^{k}\left(y_i \log p_{\text{rel}}^i + (1-y_i)\log\left(1-p_{\text{rel}}^i\right)\right), \tag{3}$$

其中 k 表示预定义关系集合中的关系个数, $y_i \in \{0,1\}$ 表示预定义集合中的某关系是否是当前句子的潜在关系。

2.3 双仿射实体配对

BEPCTS 模型的核心是使用双仿射网络来增强头尾实体的交互, 双仿射实体配对的主要作用是学习实体之间是否有关联关系。候选实体识别和双仿射实体配对的关系是: 候选实体识别先从句子中识别出所有候选实体, 在整合候选实体的嵌入表示后, 实体序列会通过双仿射实体配对网络学习实体间的关联关系。特定关系实体识别和双仿射实体配对的关系是: 特定关系实体识别首先融合某个潜在关系确定与此关系相关的实体, 然后根据实体配对矩阵特定关系的实体对是否相关, 从而抽取出关系三元组。候选实体识别使用层级标注策略, 即使用二分类标注方案来标注实体的开始位置和结束位置, 具体的计算过程如式(4)(5)所示。

$$p_i^{\text{start}} = \sigma(\boldsymbol{W}_{\text{start}} \boldsymbol{h}_i + \boldsymbol{b}_{\text{start}}), \tag{4}$$

$$p_i^{\text{end}} = \sigma(\boldsymbol{W}_{\text{end}} \boldsymbol{h}_i + \boldsymbol{b}_{\text{end}}), \tag{5}$$

其中 $\boldsymbol{W}_{\text{start}} \in \mathbb{R}^{d\times 1}$, $\boldsymbol{W}_{\text{end}} \in \mathbb{R}^{d\times 1}$, $\boldsymbol{b}_{\text{start}} \in \mathbb{R}^d$, $\boldsymbol{b}_{\text{end}} \in \mathbb{R}^d$, 其中 $\boldsymbol{W}$ 和 $\boldsymbol{b}$ 分别表示可训练的参数矩阵和偏差。p_i^{start} 和 p_i^{end} 分别表示用于分类句子中第 i 个词是否为一个实体的开始或结束的概率分布。

候选实体识别任务为一个二分类任务, 因此使用二分类交叉熵损失函数进行优化, 如式(6)所示。

$$\mathcal{L}_{\text{ent}} = -\frac{1}{n}\left(\log p_\theta^{\text{start}}(s|\boldsymbol{H}) + \log p_\theta^{\text{end}}(s|\boldsymbol{H})\right), \tag{6}$$

其中 $p_\theta(s|\boldsymbol{H}) = \prod_{i=1}^{n} p_i^{\mathbb{I}\{y_i=1\}}(1-p_i)^{\mathbb{I}\{y_i=0\}}$, $\mathbb{I}\{\cdot\}$ 表示一个指示函数, 如果“·”为真, 该指示函数为 1, 否则为 0。

当识别出句子中存在的候选实体后, 由于实体一般是句子序列中的多个词, 也称为实体跨度, 因此需要整合实体的嵌入表示。通常可以采用对跨度池化表示或对实体开始和结束的表示取平均得到实体跨度表示, 在 BEPCTS 模型中采用对头尾词表示取平均的方式。具体步骤为: 首先获得候选实体的开始位置和结束位置; 然后利用句子嵌入表示将该实体的开始词表示和结束词表示平均; 最后依次将所有候选实体的表示整合为一个序列嵌入表示。具体的计算步骤如式(7)(8)所示。

$$\boldsymbol{h}_i^{\text{e}} = \frac{\boldsymbol{h}_{\text{start}} + \boldsymbol{h}_{\text{end}}}{2}, \tag{7}$$

$$\boldsymbol{H}^{\text{e}} = \left[\boldsymbol{h}_1^{\text{e}}; \boldsymbol{h}_2^{\text{e}}; \cdots; \boldsymbol{h}_{n_{\text{e}}}^{\text{e}}\right], \tag{8}$$

其中 n_{e} 代表句子中候选实体的个数, $\boldsymbol{H}^{\text{e}}$ 表示整合后的候选实体序列嵌入表示。

接下来, 将形成序列的候选实体嵌入表示通过双仿射网络[20]形成实体配对矩阵。如图 2 所示, 首先, 候选实体序列嵌入表示通过 2 个前馈神经网络进行特征维度转换; 然后, 将左边的维度变换后的序列嵌入表示进行双仿射变换; 最后, 再与右边的维度变换后的序列嵌入表示相乘得到实体配对矩阵, 对实体配对矩阵进行分类判断两两实体之间是否相关。具体的计算步骤如式 (9)~(11)所示。

$$\boldsymbol{h}_{\text{e}}^{i'} = FFN\left(\boldsymbol{h}_{\text{e}}^{i}\right) = \boldsymbol{W}_{\text{sub}} \boldsymbol{h}_{\text{e}}^{i} + \boldsymbol{b}_{\text{sub}}, \tag{9}$$

$$\boldsymbol{h}_{\text{e}}^{j'} = FFN\left(\boldsymbol{h}_{\text{e}}^{j}\right) = \boldsymbol{W}_{\text{obj}} \boldsymbol{h}_{\text{e}}^{j} + \boldsymbol{b}_{\text{obj}}, \tag{10}$$

$$\boldsymbol{p}_{i_{\text{sub}}, j_{\text{obj}}} = \sigma\left(\boldsymbol{h}_{\text{e}}^{i'} \boldsymbol{U}_{\text{e}} (\boldsymbol{h}_{\text{e}}^{j'})^{\text{T}}\right), \tag{11}$$

其中 $\boldsymbol{W}_{\text{sub}} \in \mathbb{R}^{d\times 128}$, $\boldsymbol{W}_{\text{obj}} \in \mathbb{R}^{d\times 128}$, $\boldsymbol{b}_{\text{sub}}$ 和 $\boldsymbol{b}_{\text{obj}}$ 是可学习的

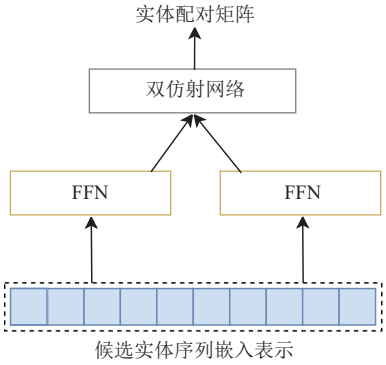


Fig. 2 Construction of biaffine entity pairing matrix

图 2 双仿射实体配对矩阵构建

参数矩阵和偏差，$\boldsymbol{U}_{\mathrm{e}} \in \mathbb{R}^{128\times1\times128}$是可学习的仿射参数矩阵。

双仿射实体配对矩阵的每个单元格表示2个候选实体之间是否相关，因此是个二分类任务，同样采取二分类交叉熵来优化实体配对阶段的损失函数，如式(12)所示。

$$\mathcal{L}_{\mathrm{pairing}} = -\frac{1}{n_{\mathrm{e}}^2}\sum_{i=1}^{n_{\mathrm{e}}}\sum_{j=1}^{n_{\mathrm{e}}}\left(y_{i,j}\log p_{i_{\mathrm{sub}},j_{\mathrm{obj}}} + \left(1-y_{i,j}\right)\log\left(1-p_{i_{\mathrm{sub}},j_{\mathrm{obj}}}\right)\right), \tag{12}$$

其中n_{e}表示候选实体的个数，$y_{i,j} \in \{0,1\}$表示第i个候选实体和第j个候选实体是否相关。

2.4 特定于关系的实体识别

由于在潜在关系预测阶段识别了句子的关系，因此在特定于关系的实体识别阶段只需针对潜在关系进行实体标注，识别出在当前关系下存在的实体。类似于候选实体标注，在该模块中同样采用层级标注策略进行实体标注。首先，将潜在关系和句子嵌入连接在一起，然后通过一个分类器标注实体的开始和结束。具体计算过程如式(13)(14)所示。

$$p_{i,j}^{\mathrm{start}} = \sigma\left(\boldsymbol{W}_{\mathrm{start}}\left[\boldsymbol{h}_i;\boldsymbol{u}_j\right] + \boldsymbol{b}_{\mathrm{start}}\right), \tag{13}$$

$$p_{i,j}^{\mathrm{end}} = \sigma\left(\boldsymbol{W}_{\mathrm{end}}\left[\boldsymbol{h}_i;\boldsymbol{u}_j\right] + \boldsymbol{b}_{\mathrm{end}}\right), \tag{14}$$

其中$\boldsymbol{W}_{\mathrm{start}} \in \mathbb{R}^{2d\times1}$，$\boldsymbol{W}_{\mathrm{end}} \in \mathbb{R}^{2d\times1}$，$\boldsymbol{b}_{\mathrm{start}} \in \mathbb{R}^{2d}$，$\boldsymbol{b}_{\mathrm{end}} \in \mathbb{R}^{2d}$；$\boldsymbol{u}_j \in \mathbb{R}^d$表示第$j$个潜在关系$r_j$在可学习嵌入矩阵$\boldsymbol{U}_r \in \mathbb{R}^{n_r\times d}$下的嵌入表示，$n_r$表示潜在关系的个数。

特定于潜在关系的实体识别模块的损失函数采用二分类交叉熵计算，如式(15)所示。

$$\mathcal{L}_{\mathrm{ent'}} = -\frac{1}{n\times n_r}\sum_{j=1}^{n_r}\left(\log p_\theta^{\mathrm{start}}(s|\boldsymbol{H}_r) + \log p_\theta^{\mathrm{end}}(s|\boldsymbol{H}_r)\right), \tag{15}$$

其中$p_\theta(s|\boldsymbol{H}_r) = \prod_{i=1}^{n}\prod p_{i,j}^{\mathbb{I}\{y_{i,j}=1\}}\left(1-p_{i,j}\right)^{\mathbb{I}\{y_{i,j}=0\}}$，$\mathbb{I}\{\cdot\}$表示一个指示函数，如果“·”为真，该指示函数为1，否则为0。$\boldsymbol{H}_r$表示融合了潜在关系表示的句子嵌入表示。

最后，对4个模块采取联合训练共享编码器更新参数以缓解流水线的误差传播问题。根据4个模块的任务复杂度，联合训练损失函数的各个模块的权重占比在BEPCTS模型中都设置为一样，如式(16)所示。

$$\mathcal{L} = \alpha\mathcal{L}_{\mathrm{rel}} + \beta\mathcal{L}_{\mathrm{ent}} + \lambda\mathcal{L}_{\mathrm{pairing}} + \eta\mathcal{L}_{\mathrm{ent'}}, \tag{16}$$

其中$\alpha = \beta = \lambda = \eta = 1$。

3 实　　验

3.1 实验数据集

为了验证BEPCTS模型的有效性，本文在4个公共数据集和一个拓展数据集上进行了广泛的实验。NYT数据集是使用远距离监督技术生成的，通过将《纽约时报》的新闻文章语料库与Freebase对齐以获得自动标注的数据集。WebNLG数据集起初是为自然语言生成任务设计的，并由Zeng等人[7]用于关系三元组抽取任务。根据以往研究工作的划分方式[14-15]，将NYT和WebNLG数据集按照重叠模型和三元组数量进行了划分，以对比模型在复杂场景下的三元组抽取性能。除此之外，本文还在NYT10-HRL和NYT11-HRL数据集[21]上进行了实验，4个数据集的详细统计信息如表2所示。

3.2 评价指标和基线模型

按照关系抽取领域研究工作的惯例，当一个抽取的关系三元组与标注的三元组完全匹配时，认为该关系三元组抽取正确。为了进行公平比较，采用精确率(precision, *Prec*)、召回率(recall, *Rec*)以及*F*1值3个评价指标来评估模型的性能，其计算方式如式(17)~(19)所示。

$$Prec = \frac{correctTriples}{predictTriples}, \tag{17}$$

$$Rec = \frac{correctTriples}{goldTriples}, \tag{18}$$

$$F1 = \frac{2\times Prec\times Rec}{Prec+Rec}, \tag{19}$$

Table 2　Experimental Datasets

表2　实验数据集

数据集	关系类别数	句子数量			重叠模式			三元组数量	
		训练集	验证集	测试集	正常	SEO	EPO	$N=1$	$N>1$
NYT	24	56 195	4 999	5 000	3 266	1 297	978	3 244	1 756
WebNLG	171	5 019	500	703	246	457	26	266	437
NYT10-HRL	29	70 339	352	4 006	-	-	-	-	-
NYT11-HRL	12	62 648	313	369	-	-	-	-	-

注：“-”表示该数据集无复杂场景的三元组。

其中*correctTriples*表示正确抽取的三元组总数，*predictTriples*表示预测的三元组总数以及*goldTriples*表示真实标注的三元组总数。

BEPCTS将与以下基线模型进行对比：

1) CopyRE[7]。一种基于复制机制的联合学习模型，并在模型中提出了关系三元组重叠问题。

2) NovelTagging[9]。一种基于标注框架的联合模型，将关系三元组抽取任务转化为序列标注任务。

3) CasRel[10]。一种基于层级标注策略的联合抽取模型，增强了实体和关系的交互。

4) PRGC[11]。使用多标签分类任务预测潜在关系并使用全局相关矩阵学习头尾实体相关关系。

5) TPLinker[15]。一种基于实体链接的联合模型，有效缓解了误差传播问题。

6) HRL[21]。一种基于强化学习的联合抽取模型，将实体识别和关系分类通过强化学习互相引导。

7) GraphRel[22]。一种关系权重的图神经网络来建模实体和关系之间的联系。

8) ERFD-RTE[23]。一种基于关系检测优先的联合关系三元组抽取模型。

9) EmRel[24]。一个将实体、关系、上下文进行统一建模的联合抽取模型。

10) ERGM[25]。应用了关系注意力机制和全局相关矩阵来抽取关系三元组。

11) SPTree[26]。基于LSTM和依存分析树的端到端关系抽取模型，通过LSTM捕获词的序列信息。

3.3 实验环境和参数设置

本节模型的训练和推理均在带有GPU的服务器上进行，实验所使用的软件和硬件环境如下：操作系统为Ubuntu18.04，CPU为Intel® Xeon® Gold 5218 CPU@2.30GHz，显卡为NVIDIA Tesla V100-SXM2，显存为32.0 GB，服务器内存为16.0 GB，实验使用语言为Python 3.7.9，深度学习开发框架为Pytorch 1.10.1，使用的预训练语言模型库为Hugging Face开源的Transformers库，版本为4.20.0。BEPCTS模型实验的具体参数如表3所示。

Table 3 Experimental Hyper-Parameter Settings

表3 实验超参数设置

实验参数	属性值
预训练语言模型	BERT
词嵌入维度	768
最大句子长度	512
双仿射变换维度	128
微调初始学习率	1E−5
下游任务初始学习率	5E−5
分类阈值	0.5
训练集批量大小	16
测试集批量大小	8
最大迭代次数	100
优化器	AdamW
正则化参数	0.3
权重衰减	0.01
梯度累计计算步骤	2
梯度剪裁阈值	2

3.4 对比实验结果

为了验证BEPCTS模型抽取关系三元组的性能，在4个常用的公开数据集NYT，WebNLG，NYT10-HRL和NYT11-HRL上进行了实验并与基线模型进行了实验结果对比，对比实验结果如表4和表5所示。

Table 4 Comparative Experimental Results of BEPCTS Model on NYT and WebNLG Datasets

表4 BEPCTS模型在NYT和WebNLG数据集上的对比实验结果 %

编码器	模型	NYT			WebNLG		
		Prec	*Rec*	*F*1	*Prec*	*Rec*	*F*1
基于Bi-RNN	CopyRE	61.0	56.6	58.7	37.7	36.4	37.1
基于Bi-LSTM-GCN	GraphRel	63.9	60.0	61.9	44.7	41.1	42.9
基于BERT	CasRel	89.7	89.5	89.6	93.4	90.1	91.8
	TPLinker	91.3	92.5	91.9	91.8	92.0	91.9
	PRGC	93.4	89.7	90.8	92.9	92.4	92.6
	EmRel	91.7	**92.5**	92.1	92.7	**93.0**	92.9
	ERFD-RTE	**94.0**	91.5	92.7	93.8	92.0	92.9
	ERGM	93.3	91.5	92.4	94.2	91.2	92.7
	BEPCTS（本文）	93.1	**92.5**	**92.8**	**94.7**	92.4	**93.5**

注：黑体数值表示在对应指标上表现最优。

从表4和表5可以看出，BEPCTS模型在NYT数据集的召回率达到了92.5%，*F*1值达到了92.8%，并且3个评价指标的平衡性更好。在WebNLG数据集上，由于该数据集是通过远程监督构建，因此数据集中存在错误标注，导致模型一般在召回率上比精确率偏低，但BEPCTS模型仍然达到了93.5%的*F*1值。在数据集NYT10-HRL上，BEPCTS相比CasRel提升较大，分别在精确率、召回率和*F*1值上提升了2.3个百分点、0.2个百分点、1.1个百分点。在数据集NYT11-HRL上，BEPCTS相比CasRel在精确率和*F*1值上分别提升了3.8个百分点和0.9个百分点。

图3是BEPCTS模型在4个公开数据集上的实

Table 5 Comparative Experimental Results of BEPCTS Model on NYT10-HRL and NYT11-HRL Datasets

表 5 BEPCTS 模型在 NYT10-HRL 和 NYT11-HRL 数据集上的对比实验结果 %

编码器	模型	NYT10-HRL			NYT11-HRL		
		Prec	*Rec*	*F*1	*Prec*	*Rec*	*F*1
基于 Bi-LSTM	NovelTagging	59.3	38.1	46.4	46.9	48.9	47.9
	HRL	71.4	58.6	64.4	53.8	53.8	53.8
基于 Bi-RNN	CopyRE	56.9	45.2	50.4	34.7	53.4	42.1
基于 LSTM-RNNs	SPTree	49.2	55.7	52.2	52.2	54.1	53.1
基于 BERT	CasRel	77.7	68.8	73.0	50.1	**58.4**	53.9
	BEPCTS（本文）	**80.0**	**69.0**	**74.1**	**53.9**	55.7	**54.8**

注：黑体数值表示在对应指标上表现最优。

验结果变化曲线。从图 3 中可以看出，BEPCTS 模型在 4 个公开数据集上的收敛速度也较快。一方面，这是由于 BEPCTS 先预测潜在关系，再基于潜在关系进行实体识别，因此在实体识别阶段减少了句子中的冗余关系。另一方面，BEPCTS 采取了联合学习共享编码器参数更新的训练方式，因此三元组中各个模块收敛速度相差不大，最终模型可以在较早的迭代轮次收敛到较好的结果。

综合来看，在 4 个公开数据集上 BEPCTS 的模型性能优于对比的基线模型。首先，BEPCTS 模型进行了潜在关系的预测，减少了实体标注阶段的冗余关系，因此抽取三元组的精确率较高。其次，在实体配对阶段引入双仿射网络，进一步从模型结构上增强头实体和尾实体的交互，从而更好地学习实体间的相关关系。最后，在实体标注阶段采用了层级标注策略，使得模型的实体标注标签减少为 2 个，只需要标注实体的开始位置和结束位置。

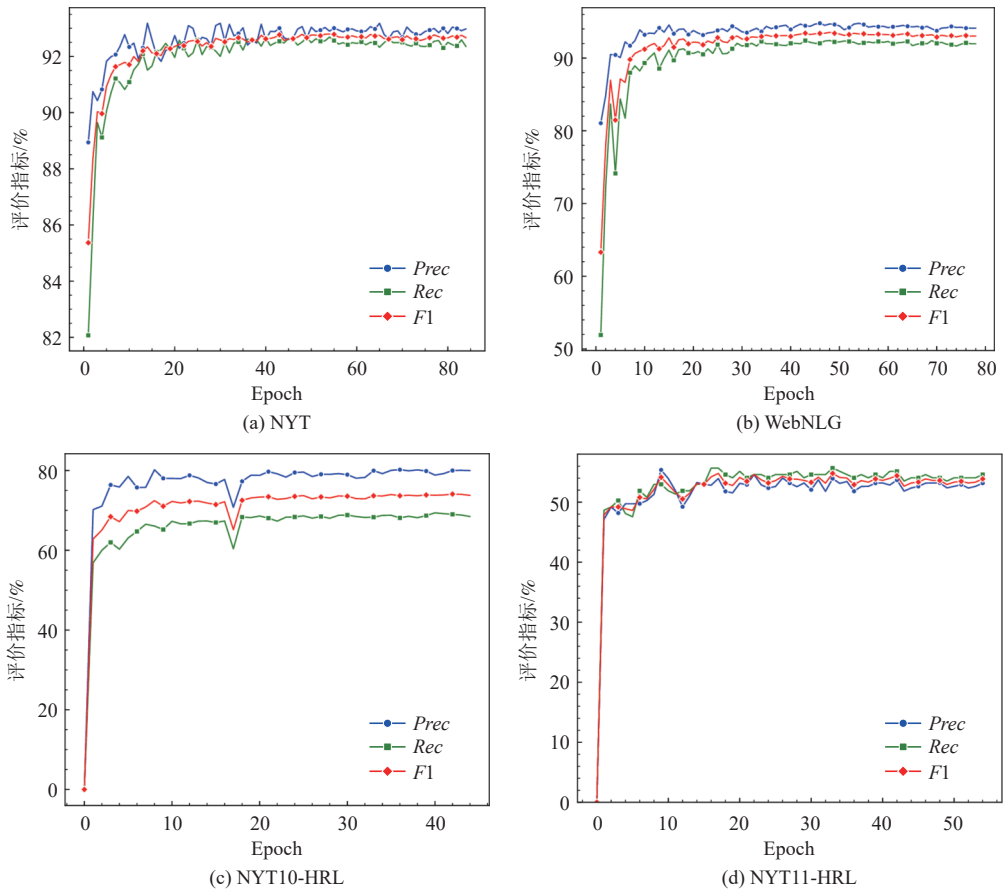


Fig. 3 Change curves of experimental results for BEPCTS model on four public datasets

图 3 BEPCTS 模型在 4 个公开数据集上的实验结果变化曲线

3.5 消融实验结果

为了验证 BEPCTS 模型中潜在关系预测、双仿射实体配对和层级标注策略 3 个模块对模型的影响，本节对 BEPCTS 模型进行了消融实验，实验结果如表 6 所示。

首先，为了验证潜在关系预测模块的有效性，BEPCTS 模型将关系预测模块替换为预定义关系集合。此时模型在特定关系的实体识别阶段将依次遍

Table 6 Experimental Results of the Ablation Study

表 6 消融实验结果 %

模型	NYT			WebNLG		
	Prec	*Rec*	*F*1	*Prec*	*Rec*	*F*1
BEPCTS	93.1	92.5	92.8	94.7	92.4	93.5
−潜在关系预测	92.3	91.1	91.7	93.9	91.8	92.8
−双仿射实体配对	92.8	91.5	92.1	94.1	91.9	93.0
−层级标注策略	92.5	92.2	92.3	93.7	92.1	92.9

历预定义关系集合中的每一个关系。从表 6 可以看出，在去掉了潜在关系预测模块后，BEPCTS 模型在 NYT 和 WebNLG 这 2 个数据集上的 *F*1 值分别下降了 1.1 个百分点和 0.7 个百分点。虽然模型在抽取性能上没有出现大幅度下降，但潜在关系预测模块的主要作用是减少冗余关系对后续模块的影响，因此遍历整个预定义关系集合对于关系类型较多的数据集将带来大量无意义的实体标注。

其次，为了验证双仿射实体配对模块的有效性，模型将双仿射网络去掉，参考 PRGC 中使用全局相关矩阵的方式进行实体配对。表 6 实验结果表明，在不使用双仿射网络进行实体配对的情况下，BEPCTS 模型在 NYT 数据集上的精确率下降了 0.3 个百分点，召回率下降了 1.0 个百分点以及 *F*1 值下降了 0.7 个百分点，在 WebNLG 数据集上的精确率下降了 0.6 个百分点，召回率下降了 0.5 个百分点以及 *F*1 值下降了 0.5 个百分点。相较于使用全局相关矩阵，使用双仿射实体配对增强了头尾实体间的语义交互性，因此使得模型拥有更强的泛化能力。

最后，为了验证层级标注策略相较于传统命名实体标注策略的优势，模型将实体标注策略替换为传统标注方案。从表 6 看出，模型在 NYT 和 WebNLG 两个数据集上的 *F*1 值分别下降了 0.5 个百分点和 0.6 个百分点。层级标注策略主要体现在候选实体标注和特定于关系的实体标注模块，模型使用层级标注策略只需标注实体的开始位置和结束位置，相较于传统标注方案减少了实体标签的个数，简化了模型结构的同时也提升了实体识别的精确率。

3.6 复杂场景分析

为了进一步验证 BEPCTS 在复杂场景下抽取关系三元组的性能，本节将对比 BEPCTS 在抽取重叠三元组和多个三元组时的 *F*1 值。通常，在句子中存在 SEO 和 EPO 模式的三元组重叠以及一个句子包含多个三元组时，模型的性能会出现下降，因此模型在这 2 个复杂场景下的性能表现决定了模型的鲁棒性。如图 4 和表 7 所示，BEPCTS 和 CasRel, PRGC, ERGM 这 3 个基线模型进行了复杂场景下的对比实验。

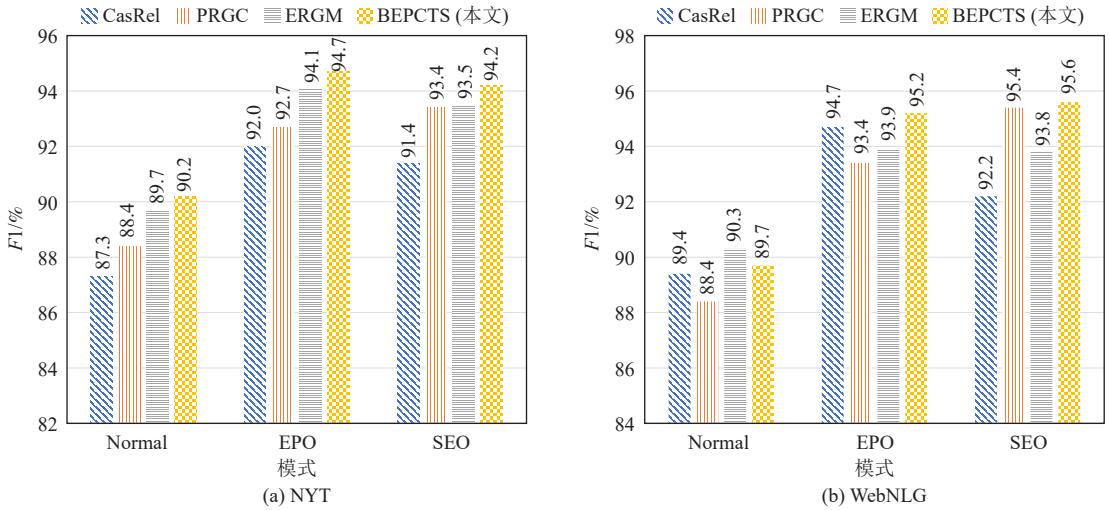


Fig. 4 Comparative experimental results on extracting overlapping triples of BEPCTS model

图 4 BEPCTS 模型抽取重叠三元组的对比实验结果

首先，在抽取重叠三元组时，BEPCTS 在 NYT, WebNLG 这 2 个数据集上都表现出了不错的性能。如图 4 所示，在 NYT 数据集上，模型在 EPO 和 SEO 这 2 个重叠模式上分别达到了 94.7% 和 94.2% 的 *F*1 值，相比 ERGM 分别提升了 0.6 个百分点和 0.7 个百分点。在 WebNLG 数据集上，模型在 EPO 和 SEO 这 2 个重叠模式上分别达到了 95.2% 和 95.6%，相比 ERGM 模型分别提升了 1.3 个百分点和 1.8 个百分点。进一步验证了 BEPCTS 模型可以有效缓解三元组重叠问题。

其次，针对每个句子包含多个三元组的复杂场景，模型也进行了对比实验。如表 7 所示，当句子的

Table 7 Experimental Results of *F*1 on BEPCTS Model Extracting Multiple Triples

表 7 BEPCTS 模型抽取多个三元组的 *F*1 实验结果 %

模型	NYT					WebNLG				
	$N=1$	$N=2$	$N=3$	$N=4$	$N\geqslant 5$	$N=1$	$N=2$	$N=3$	$N=4$	$N\geqslant 5$
CasRel	88.2	90.3	91.9	94.2	83.7	89.3	90.8	94.2	92.4	90.9
PRGC	89.0	91.9	92.5	95.6	86.2	88.4	91.9	94.0	94.8	92.9
ERGM	90.9	93.4	93.1	95.7	90.1	90.3	91.5	95.1	94.0	92.5
BEPCTS（本文）	90.2	93.4	93.8	95.8	92.9	89.3	92.7	94.9	95.1	95.0

三元组个数属于$N=3$, $N=4$, $N\geqslant 5$这 3 种场景时，BEPCTS 模型在 NYT 数据集上分别达到了 93.8%, 95.8%, 92.9% 的 *F*1 值，相比 ERGM 分别提升了 0.7 个百分点、0.1 个百分点、2.8 个百分点。在 WebNLG 数据集上，BEPCTS 分别在句子三元组个数为 N=3, N=4 和 $N\geqslant 5$ 时达到了 94.9%, 95.1%, 95.0% 的 *F*1 值，其中抽取后 2 种三元组个数的 *F*1 值相比 ERGM 分别提升了 1.1 个百分点和 2.5 个百分点。可以看出，BEPCTS 在三元组个数较多的句子中也有着不错的性能表现，验证了 BEPCTS 在多三元组复杂场景中的鲁棒性。

最后，为了验证 BEPCTS 在推理效率方面的提升，分析了对比联合模型的时间复杂度并对推理时间进行了对比。如表 8 所示，在 NYT 数据集上，BEPCTS 模型相比 ERGM 推理时间降低了 2.1 ms。在 WebNLG 数据集上，BEPCTS 相比 ERGM 降低了 2.4 ms。此外，由于只针对候选实体构建双仿射实体配对矩阵，BEPCTS 的时间复杂度也更低。

Table 8 Analysis for Computational Complexity and Time Performance

表 8 计算复杂度和时间性能分析 ms

模型	时间复杂度	推理时间	
		NYT	WebNLG
CasRel	$O(n+sro)$	54.0	82.8
PRGC	$O(n^2)$	46.5	113.6
ERGM	$O(n+srn_e^2)$	52.5	79.3
BEPCTS（本文）	$O(n+rn_e^2)$	50.4	76.9

注：s 表示头实体，r 表示关系，o 表示尾实体，n_e 表示实体数量。

4 结 论

本文提出了一种基于双仿射实体配对和层级标注策略的联合实体关系抽取模型，模型使用 BERT 预训练语言模型作为句子编码器。首先，通过多标签分类任务进行潜在关系预测，过滤了与句子无关的关系，使得模型在后续的特定关系实体识别阶段减少了冗余关系。其次，通过层级二分标注策略识别句子中的候选实体集合，将候选实体集合的序列嵌入表示通过双仿射网络形成实体配对矩阵，该模块有效增强了头尾实体的交互性并减少了配对矩阵的冗余标注。最后，通过特定关系的实体识别模块，将得到的特定关系实体与实体配对矩阵结合以抽取出关系三元组。通过在 4 个公开数据集上进行实验，结果表明所提出的模型不仅在三元组抽取整体性能上有着不错的表现，并且针对重叠三元组和多三元组等复杂场景也有一定的优势。此外，与结构复杂的联合模型相比，所提出的模型有更快的收敛速度以及更快的推理速度，能在较少的迭代轮次达到较好的抽取性能。未来工作将进一步研究如何构建结构更简单的联合抽取模型。

作者贡献声明：张清华提出了算法思路和实验方案；周雄负责完成实验并撰写论文；廖伟、黄帅帅、秦徐婷提出指导意见并修改论文。

参 考 文 献

[1] Lin Yankai, Liu Zhiyuan, Sun Maosong, et al. Learning entity and relation embeddings for knowledge graph completion[C]//Proc of the 29th AAAI Conf on Artificial Intelligence. Menlo Park, CA: AAAI, 2015: 2181−2187

[2] Martinez-Rodriguez J L, Hogan A, Lopez-Arevalo I. Information extraction meets the semantic Web: A survey[J]. Semantic Web, 2020, 11(2): 255−335

[3] Zhang Xiaoming, Yu Can, Yan Rui. ParTRE: A relational triple extraction model of complicated entities and imbalanced relations in Parkinson's disease[J]. Journal of Biomedical Informatics, 2024, 152: 104624

[4] Li Dongmei, Zhang Yang, Li Dongyuan, et al. Review of entity relation extraction methods[J]. Journal of Computer Research and Development, 2020, 57(7): 1424−1448 (in Chinese)

(李冬梅, 张扬, 李东远, 等. 实体关系抽取方法研究综述[J]. 计算机研究与发展, 2020, 57(7): 1424−1448)

[5] Zhang Yangsen, Liu Shuaikang, Liu Yang, et al. Joint extraction of entities and relations based on deep learning: A survey[J]. Acta Electronica Sinica, 2023, 51(4): 1093−1116(in Chinese)
(张仰森, 刘帅康, 刘洋, 等. 基于深度学习的实体关系联合抽取研究综述[J]. 电子学报, 2023, 51(4): 1093−1116)

[6] Devlin J, Chang Mingwei, Lee K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[C]//Proc of the 17th Conf of the North American Chapter of the ACL. Stroudsburg, PA: ACL, 2019: 4171−4186

[7] Zeng Xiangrong, Zeng Daojian, He Shizhu, et al. Extracting relational facts by an end-to-end neural model with copy mechanism[C]//Proc of the 56th Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2018: 506−514

[8] Zhang Zhanjun, Zhang Haoyu, Wan Qian, et al. Entity–relation triple extraction based on relation sequence information[J]. Expert Systems with Applications, 2024, 238: 121561

[9] Zheng Suncong, Wang Feng, Bao Hongyun, et al. Joint extraction of entities and relations based on a novel tagging scheme[C]//Proc of the 55th Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2017: 1227−1236

[10] Wei Zhepei, Su Jianlin, Wang Yue, et al. A novel cascade binary tagging framework for relational triple extraction[C]//Proc of the 58th Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2020: 1476−1488

[11] Zheng Hengyi, Wen Rui, Chen Xi, et al. PRGC: Potential relation and global correspondence based joint relational triple extraction[C]//Proc of the 59th Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2021: 6225−6235

[12] Dixit K, Al-Onaizan Y. Span-level model for relation extraction[C]//Proc of the 57th Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2019: 5308−5314

[13] Eberts M, Ulges A. Span-based joint entity and relation extraction with Transformer pretraining[C]//Proc of the 24th European Conf on Artificial Intelligence. Hemweg, AMM: IOS, 2020: 2006−2013

[14] Yang Yi, Zhou Shangbo, Liu Yuxuan. Bidirectional relation-guided attention network with semantics and knowledge for relational triple extraction[J]. Expert Systems with Applications, 2023, 224: 119905

[15] Wang Yucheng, Yu Bowen, Zhang Yueyang, et al. TPLinker: Single-stage joint extraction of entities and relations through token pair linking[C]//Proc of the 28th Int Conf on Computational Linguistics. Stroudsburg, PA: ACL, 2020: 1572−1582

[16] Dai Qicai, Yang Wenzhong, Wang Liejun, et al. SOIRP: Subject-object interaction and reasoning path based joint relational triple extraction by table filling[J]. Neurocomputing, 2024, 580: 127492

[17] Shang Yuming, Huang Heyan, Mao Xianling. Onerel: Joint entity and relation extraction with one module in one step[C]//Proc of the 36th AAAI Conf on Artificial Intelligence. Menlo Park, CA: AAAI, 2022: 11285−11293

[18] Ning Jinzhong, Yang Zhihao, Sun Yuanyuan, et al. OD-RTE: A one-stage object detection framework for relational triple extraction [C]//Proc of the 61st Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2023: 11120−11135

[19] Li Xiaoya, Yin Fan, Sun Zijun, et al. Entity-relation extraction as multi-turn question answering[C]//Proc of the 57th Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2019: 1340−1350

[20] Yu Juntao, Bohnet B, Poesio M. Named entity recognition as dependency parsing[C]//Proc of the 58th Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2020: 6470−6476

[21] Takanobu R, Zhang Tianyang, Liu Jiexi, et al. A hierarchical framework for relation extraction with reinforcement learning[C]//Proc of the 33rd AAAI Conf on Artificial Intelligence. Menlo Park, CA: AAAI, 2019: 7072−7079

[22] Fu T J, Li P H, Ma Weiyun. GraphRel: Modeling text as relational graphs for joint entity and relation extraction[C]//Proc of the 57th Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2019: 1409−1418

[23] Chen Juan, Hu Jie, Li Tianrui, et al. An effective relation-first detection model for relational triple extraction[J]. Expert Systems with Applications, 2024, 238: 122007

[24] Xu Benfeng, Wang Quan, Lyu Yajuan, et al. EmRel: Joint representation of entities and embedded relations for multi-triple extraction[C]//Proc of the 20th Conf of the North American Chapter of the ACL. Stroudsburg, PA: ACL, 2022: 659−665

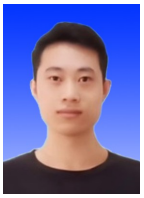
[25] Gao Chen, Zhang Xuan, Li Linyu, et al. ERGM: A multi-stage joint entity and relation extraction with global entity match[J]. Knowledge-Based Systems, 2023, 271: 110550

[26] Miwa M, Bansal M. End-to-end relation extraction using LSTMs on sequences and tree structures[C]//Proc of the 54th Annual Meeting of the ACL. Stroudsburg, PA: ACL, 2016: 1105−1116



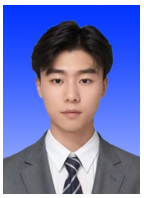
Zhang Qinghua, born in 1974. PhD, professor, PhD supervisor. His main research interests include rough set, fuzzy set, granular computing, and uncertain information processing.

张清华, 1974 年生。博士, 教授, 博士生导师。主要研究方向为粗糙集、模糊集、粒计算、不确定性信息处理。



Zhou Xiong, born in 1998. Master. His main research interests include natural language processing, information extraction, and relation extraction.

周　雄, 1998 年生。硕士。主要研究方向为自然语言处理、信息抽取、关系抽取。



Liao Wei, born in 1999. Master. His main research interests include machine learning, rough set, and uncertain information processing.

廖　伟, 1999 年生。硕士。主要研究方向为机器学习、粗糙集、不确定性信息处理。



Huang Shuaishuai, born in 1999. Master. His main research interests include knowledge graph, knowledge acquisition, and representation learning.

黄帅帅, 1999 年生。硕士。主要研究方向为知识图谱、知识获取、表示学习。



Qin Xuting, born in 2001. Master candidate. Her main research interests include graph convolutional neural networks, multi-view learning, and machine learning.

秦徐婷, 2001 年生。硕士研究生。主要研究方向为图卷积神经网络、多视图学习、机器学习。