计算机科学
COMPUTER SCIENCE

# 限定域关系抽取技术研究综述

侯景, 邓晓梅, 汉鹏武

引用本文

侯景, 邓晓梅, 汉鹏武. 限定域关系抽取技术研究综述[J]. 计算机科学, 2024, 51(1): 252-265.

HOU Jing, DENG Xiaomei, HAN Pengwu. Survey on Domain Limited Relation Extraction[J]. Computer Science, 2024, 51(1): 252-265.

---

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.230200100

# 限定域关系抽取技术研究综述

侯　景[1,2]　邓晓梅[1]　汉鹏武[1]
1 中国科学院空间应用工程与技术中心　北京 100094
2 中国科学院大学　北京 100094
(houjing21@mails.ucas.ac.cn)

**摘　要**　限定域关系抽取技术是在预定义实体类型和关系类型的前提下,从文本中捕获关键信息的技术,多采用由头尾实体和关系构成的三元组作为信息表示形式。作为信息抽取领域的重要研究方向之一,其在知识问答、信息检索等任务中被广泛应用。文中在介绍相关概念和任务范式的基础上,分析了深度学习背景下限定域关系抽取任务的研究进展,根据句中实体是否可见,分为关系分类任务和三元组抽取任务,依据任务表现特征,前者可细分为有监督条件下的关系分类任务、小样本关系分类任务和远程监督条件下的关系分类任务。文中探讨和分析了以上任务中常用的技术方法及其优缺点,最后归纳总结了关系抽取技术在低资源、多模态等更为接近真实情景下的发展潜力和现存的挑战。

**关键词**:限定域关系抽取;深度学习;关系分类;三元组;远程监督
**中图分类号**　TP391

## Survey on Domain Limited Relation Extraction

HOU Jing[1,2],DENG Xiaomei[1] and HAN Pengwu[1]
1 Technology and Engineering Center for Space Utilization,Chinese Academy of Sciences,Beijing 100094,China
2 University of Chinese Academy of Sciences,Beijing 100094,China

**Abstract**　Domain-limited relation extraction aims to capture essential text information from the text under the premise of predefined entity types and relation types,and mostly uses triples composed of head and tail entities and relations as structured information representation. As one of the important tasks of information extraction,it plays an important role in question answering and information retrieval. Based on its concepts and task paradigms,this paper systematically sorts out the technical methods in domain-limited relation extraction under the background of deep learning. Whether the entity is visible or not,it is divided into relation classification and triplet extraction. According to the performance characteristics of the task,the former can be divided into relation classification under supervised conditions,few-shot relation classification,and relation classification under distant supervision. This paper discusses and analyzes the commonly used technical methods and their advantages and disadvantages in the above tasks. Finally,we summarize the development potential and existing challenges of relation extraction technology in low-resource,multimodal and other situations that are closer to the real world.

**Keywords**　Domain-limited,Deep learning,Relation classification,Triples,Distant supervision

## 1　引言

关系抽取作为自然语言处理的关键任务之一,旨在从非结构化文本中自动识别由一对实体和联系这对实体的关系构成的相关三元组,得到文本语料的关键信息[1]。根据实体和关系预定义与否,分为开放域关系抽取和限定域关系抽取。

在限定域句子关系抽取任务中,实体类型和关系类型是预先给定的,实体是具有特定指称项的提及,如人物、地点、机构等[2]。传统方法多是基于统计机器学习的方法,将关系实例表示为高维空间的特征向量,模型性能十分依赖于特征选取的优劣。近年来,深度学习技术被引入该任务中,通过设计不同的网络架构和文本表示方法来捕捉文本特征,在SCIERC[3],ACE2005[4],ADE[5]等多个数据集上取得了良好的性能。目前,已有一些学者对相关文献做出了梳理,如Li等[6]按照传统关系抽取方法、基于机器学习、基于深度学习和基于开放领域的关系抽取方法整理了不同阶段的主流技术方法;Zhuang等[7]、Nayak等[8]从深度学习网络架构角度展开了相关技术方法的介绍;Zhang等[9]分析了三元组抽取中的联合抽取技术;Bai等[10]从技术方法视角,探讨了基于概率图的、基于矩阵补全的和基于嵌入的远程监督关系抽取方法。以上研究的视角定位于句子及文件级别的关系分类任务,按照关系抽取发展脉络或基于网络架构展开。

不同于上述研究,考虑到深度学习技术发展突飞猛进,在多个任务上不断刷榜,涌现了新的技术范式,并催生了新的

到稿日期:2023-02-15　返修日期:2023-06-29
通信作者:汉鹏武(hanpengwu@csu.ac.cn)

任务范式,本文从关系抽取任务范式视角出发,按照关系抽取任务范式研究并探讨了最新技术进展。本文第2章介绍了限定域关系抽取的概念和难点问题;第3章介绍了常用评测数据集和已有工具;第4—7章研究了在有监督条件下的关系分类、小样本关系分类、远程监督抽取任务和三元组抽取中的技术进展;最后,探讨了在低资源、多模态场景下该任务的发展潜力和研究挑战。

## 2　限定域关系抽取简介

限定域关系抽取是在关系类型预定义的前提下,从文本中提取结构化信息,即三元组。在实体给定的前提下,只需识别给定实体之间的关系,本文称之为关系分类任务,如图1所示,例句"乌克兰加入北约的请求,导致俄罗斯和乌克兰发生了战事冲突"中,关系分类任务是判断"俄罗斯"和"乌克兰"之间的关系类型为"战争"。根据模型中所用数据特征,可细分为有监督条件下的关系分类问题[11-13]、小样本下的关系分类问题[14-15]和远程监督下的关系分类问题[16-17]。有监督条件下的关系抽取是面向高质量标注数据完成关系分类任务,任务范式较为简单。小样本下的分类问题是针对人工标注数据集稀缺的现状,仅使用小批量标注数据完成关系抽取任务,一般设定N-way-K-shot形式,其中,$N$表示关系类型数量,$K$表示每种类型下的实例数量。2009年,Wu等[18]采用Infobox对wikipedia文本回标产生训练语料集,随后,Mike等[19]提出了远程监督的思想,认为知识库中包含同一实体对的句子表述了相同的关系类别。基于这一假设,根据知识库中的关系,启发式标注语料集,以构成句子包,得到伪标注数据集,其判断关系标签的任务被视为远程监督下的关系抽取。

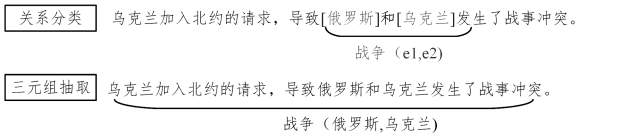



图1　分类任务和三元组抽取任务示例

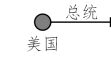
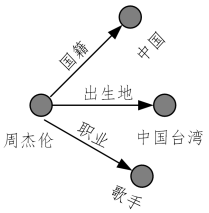
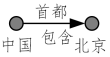
Fig. 1　Example of relation classification task and triple extraction task

在实体未给定的条件下,需同时识别句中实体和关系,称为三元组抽取任务。根据两者的识别顺序可分为串行的管道式抽取[20]方法和并行的联合抽取[21-22]方法,如图1所示,不仅需要识别出语句中的实体——"俄罗斯"和"乌克兰",并且需要判断实体对之间可能存在的关系类型,进而得到最终的三元组——〈俄罗斯,战争,乌克兰〉。

在实际语料中,句子关系情形复杂,单句中可能包含多个三元组,出现实体重叠、实体对重叠的现象,如表1所列,增加了任务难度。

表1　三元组重叠示例

Table 1　Example of triple overlay

| 重叠类型 | 样句 | 三元组 | 三元组图示 |
|---|---|---|---|
| 正常 | 美国总统拜登发表上任演讲 | 〈美国,总统,拜登〉 |  |
| 单实体重叠(Single Entity Overlap,SEO) | 周杰伦出生于中国台湾地区,是中国知名歌手 | 〈周杰伦,出生地,中国台湾〉、〈周杰伦,职业,歌手〉、〈周杰伦,国籍,中国〉 |  |
| 实体对重叠(Entity Pair Overlap,EPO) | 北京是中国的首都 | 〈中国,首都,北京〉、〈中国,包含,北京〉 |  |

## 3　相关数据集及工具

### 3.1　评测数据集

在限定域关系抽取任务中,常用的数据集如表2所列,其中,SemEval[4]和TACRED[23]常用于监督条件下的关系分类任务,Few系列[24]常用于小样本学习场景,NYT系列[25]、WebNLG[26]、SKE[27]常用于联合抽取任务,GDS多用于远程监督任务。但总体来看,已有的标注数据集并不能适用于垂直领域,大多是面向通用文本领域的数据集,并且数据集的质量不能保证,可能存在关系类型标注错误[28]、实体标注错误[29]的问题。

表2　限定域关系抽取常用评测数据集

Table 2　Common evaluation datasets in limited domain relation extraction task

| 数据集 | 关系总数 | 训练条目 | 验证条目 | 测试条目 | 语言 | 是否手工创建 | 重叠与否 | 单元组 or 多元组 | 数据集地址 |
|---|---|---|---|---|---|---|---|---|---|
| SemEval | 19(Other) | 8000 | | 2717 | 英语 | 是 | 否 | 单元组 | hpt/lmwmal fbk eulsemevala2phplocatio n=tasks #T11 |
| Few2.0 | 100 | 56000 | | 14000 | 英语 | 是 | 是 | 单元组 | https./github.com/thumlp/ferel |
| NYT-10-D | 53(NA) | 466876 | 55167 | 172448 | 英语 | 否 | 是 | 多元组 | htps//github com/thumlp/OpenNRE/tree/master/ben chmark |
| WebNLG | 246 | 5019 | 500 | 703 | 英语 | 是 | 是 | 多元组 | tp/:/acti Joiaff webnlg/stories/challenge.html |
| SKE | 50 | 172983 | 21626 | 19981 | 中文 | 否 | 是 | 多元组 | htp//ai baidu com/broa d/dowmload?dataset=sked |
| TACRED | 42(NA) | 68124 | 22631 | 15509 | 英文 | 是 | 是 | 多元组 | hts://p stanfordeduprojectsltacred/ |
| GDS | 5 | 11297 | 1864 | 5663 | 英文 | Test(部分) | 否 | 单元组 | htps//gjthub.com/SharmisthaJat/REDS-Wond-Attention-Models |
| NYT-11-D | 25(NA) | 53395 | | 368 | 英语 | test | 是 | 多元组 | htsp//lifhuu.b com/trufhless11HRL-RE |

### 3.2　常用工具

限定域关系抽取经过多年的发展,较为成熟的关系抽取模型已被开发为工具包,供研究人员快速上手使用,常用工具包如表3所列。现有的工具多是以通用公开语料集为训练语料,在垂直领域中的表现堪忧。另外,由于大多数工具中的关系是预定义的,因此难以将该工具迁移到

新的关系集合中完成关系分类任务。

表 3 关系抽取工具包

Table 3 Relation extraction toolkit

| 工具名称 | 开发机构 | 语种 | 链接 |
|---|---|---|---|
| DeepKE | 浙江大学 | 中文 | https://github.com/zjunlp/deepke |
| Jiagu | 思知机器人公司 | 中文 | https://github.com/ownthink/Jiagu |
| DeepDive | 斯坦福大学 | 中文 | http://www.openkg.cn/dataset/cn-deepdive |
| OpenNRE | 清华大学 | 中文 | https://github.com/thunlp/OpenNRE |
| ReVerb | 华盛顿大学 | 英文 | https://github.com/knowitall/reverb |
| Ollie | 华盛顿大学 | 英文 | https://github.com/knowitall/ollie |
| IEPY | UNC-FaMAF | 英文 | https://github.com/machinalis/iepy |
| ClausIE | 斯坦德大学 | 英文 | https://github.com/jeffrschneider/clausie |
| MinIE | 曼海姆大学 | 英文 | https://github.com/uma-pi1/minie |

## 4 限定域下的关系分类方法

在传统方法中形成了基于特征向量[30]、核函数为代表[31]的方法，主要依赖于人工设计并抽取文本词汇特征和句法结构特征，然而这些方法需要大量的人工工作，可能会忽略一些重要的特征，同时也存在泛化性能较差的问题。近年来，逐渐采用循环神经网络(Recurrent Neural Network，RNN)[32-33]、长短时记忆神经网络(Long Short Term Memory，LSTM)[34-36]、卷积神经网络(Convolution Neural Network，CNN)[1,37-38]、图卷积神经网络(Graph Convolutional Network，GCN)[39-41]等深度学习网络架构替代传统特征工程，以捕获文本信息特征，并在多项任务中取得了良好的性能表现，文中按照任务范式梳理了基于深度学习的相关技术方法。

### 4.1 有监督条件下的关系分类方法

将深度学习引入关系分类任务中，旨在自动获取句子的语义信息特征，提高模型的泛化能力。Zeng 等[42]提出了基于卷积神经网络的关系分类模型，如图 2 所示，利用 CNN 捕获句子级特征，在词表示上考虑了位置特征，与词向量共同构成词的语义表示，缓解了“猫吃鱼”“鱼吃猫”这类语义序列混乱的问题，在 SemEval 数据集上超越了手工设计特征下的实验结果，但 CNN 无法建模长距离依赖关系，导致句子信息前后关联性弱化，而 RNN 可以建模长序列文本。Zhang 等[32]将 CNN 替换为双向 RNN 模型，在长句语义信息学习上呈现出了较好的表现力。考虑到 LSTM 模型设计了记忆单元和门控机制，使得模型能够有效地保留和更新输入信息，能够缓解 RNN 带来的梯度消失问题，Zhou 等[35]将 LSTM 作为信息抽取的网络结构，并注意到句子中的不同位置的词对关系分类任务的贡献度不同，提出了注意力机制下的 LSTM 网络结构，模型性能提高了 1.3%。以上方法在模型架构选择上逐渐倾向于具有时序性的模型，如 RNN 和 LSTM，这使得模型能够更好地保留完整的语句信息，但只能从一个方向处理序列的局限性，导致模型无法同时考虑双向下的语义信息。学者们逐渐使用双向 RNN、双向 LSTM 结构缓解这个问题。另外，将句中所有的词视为等权重的方式可能造成关键信息丢失。为了更精确地把握关系分类任务中所需要的语义信息，注意力机制被引入到模型中，根据语句中词的位置和语境，动态地给每个词分配不同的权重，促使模型能够更专注于句子中的关键信息。

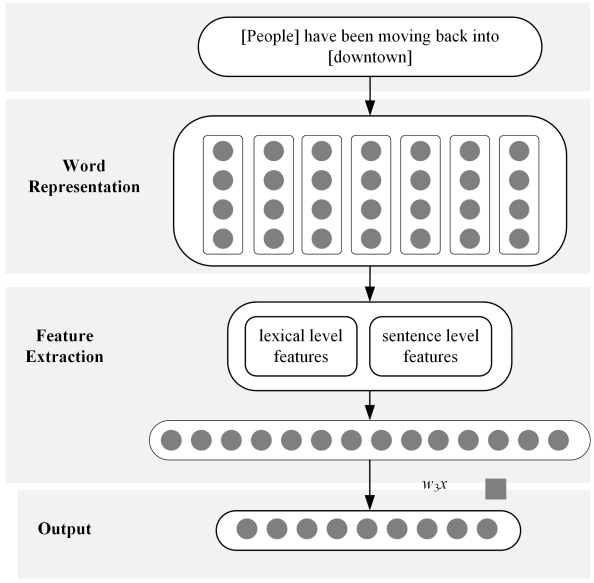



图 2 Zeng 等提出的基于 CNN 的语义学习模型

Fig. 2 CNN based semantic learning model proposed by Zeng et al.

当单句包含多个三元组时，每个三元组中的实体对表现出的关系类别存在差异，Zhou 等[43]首先研究了句子中存在 3 个实体时关系识别的方法，对所有可能的实体对关系组合进行判断，若存在关系，则对应序列位置为 1，否则为 0，如{$i$,0,0}，$i$ 表示实体 1 和实体 2 之间的关系类型，其余两个 0 分别表示实体 2 和实体 3、实体 1 和实体 3 之间没有关系，这种方式忽略了三元组中关系的方向性且可拓展性较差。进一步，当句子中包含多个实体对时，针对不同实体对进行关系分类时，句子中同一个词的重要性直觉上是不同的，Qin 等[13]提出了基于实体对信息的注意力机制，依据句子中的不同实体对，为句子分配不同的注意力权重，使得模型学习面向不同实体对的特定语义向量表示，用于关系分类。在预训练语言模型(Pretrained Language Model，PLM)强大的语言表征能力下，关系分类任务得到了快速发展，模型性能甚至超越人类，进而推动了新的任务范式产生，如小样本关系分类、零样本关系抽取等，新的任务也更为贴近于真实场景。

### 4.2 小样本条件下的关系分类方法

对于一个新的领域而言，人工标注数据集稀缺，而从头标注数据需要大量人力、财力支撑，这增加了技术落地的难度，因此，如何在小批量标注数据集上建立高精度模型具有一定的实用价值和意义。按照小样本条件下所用的预训练语言模型范式，从自回归语言模型下的关系抽取方法和自编码语言模型下的关系抽取方法两个角度梳理了相关技术。

#### 4.2.1 自回归语言模型下的关系抽取方法

自回归语言模型是在给定文本上文的条件下，对当前词进行预测，训练过程要求对数似然函数最大化，如式(1)所示。代表模型有 ELMO[44]、GPT 系列[45]、BART[46]、T5[47]等。

$$\max \log p_\theta(x)=\sum_{t=1}^{T}\log p_\theta(x_t \mid x_{<t}) \tag{1}$$

在自回归语言模型中，提示学习[48-50]方法得到了越来

越多的关注,其核心是尽可能地将下游任务建模为预训练语言模型的任务范式,以实现在低数据资源下完成下游任务,大致流程如下:

1)针对任务目标,建立提示模板[51-53];

2)答案搜索,填充模板中的掩码位置[54];

3)答案映射[55-56],将预训练语言模型预测产生的输出映射到现有任务定义的标签词上。

Fabio 等[57]采用人工构建的硬提示,如[X] was born in [MASK],在提示中添加[mask]标志,通过识别掩码位置([mask])的令牌词(token)达到关系分类的目的。但这种方式将可填充词表限制为预训练语言模型的词汇库,导致在OOV(Out of Vocabulary)情况下表现不佳。针对这个问题,引入更为丰富和全面的知识源能够提升模型的鲁棒性,提高模型性能。Zhang 等[58]利用文本标签——关系名称作为提示信息来辅助模型学习,同时为了保持训练和测试的一致性,提出了一种标签提示丢弃(Label Prompt Dropout)的方法,即在训练过程中随机删除部分标签提示信息,以增强模型对标签提示信息的泛化能力,提升了小样本关系抽取模型的性能,但将关系集合中的每种关系类型视为独立的类别,未考虑到不同关系之间可能存在的层次或语义上的联系。在未来的研究中,可以探索如何利用关系之间的内在结构来进一步提升模型性能。基于自回归的提示学习方法目前主要被应用在情感分类[59-60]、实体抽取[61-63]等简单的信息抽取任务中,在关系分类任务中应用较少,但发展前景可观。

#### 4.2.2　自编码语言模型下的关系抽取方法

自编码语言模型本质为去噪自编码模型,采用掩码方式对句子加入噪声,在训练过程中,根据上下文预测掩码的词,使其概率最大化,如式(2)所示,以达到去噪的目的,从而学习得到语言表征,代表模型有 Bert[64]等。自编码语言模型通过重建损失能够有效捕捉文本序列的语义特征,常作为原型网络的编码器。原型网络常采用每个类别下样本编码的平均值作为类表示向量,也被称为原型表示,通过测评实例和关系原型之间的相似度,来达到判断实例关系类型的目的。Han 等[65]利用基于 bert 自编码语言模型,使用原型网络构建了小样本关系分类数据集 FewRel 的测评基线。Gao 等[24]为其增加了新的测试集,同样以自编码语言模型 Bert 作为语言表征的学习模型,提出了 Bert-pair 的分类方式,具体是将测试集中的实例作为查询子集(query),将训练集中的实例作为支持子集(support),通过计算 query 和 support 配对分值,来计算两个实例表达相同关系的概率。类似地,Dong 等[66]、Liu 等[67]将输入句子和目标关系映射到一个语义空间中,然后计算它们之间的相似度,以判断句子中是否存在目标关系,但这种方法的假设前提是关系是对称的,即两个实体之间的关系不受它们在句子中的顺序影响,这可能不适用于一些非对称的关系,如"出生地"或"毕业院校"等。除使用句子和关系标签本身携带的信息外,Zhang 等[68]认为知识图谱和维基百科中包含了实体的相关信息,提出了基于融入实体描述信息的预训练语言模型增强策略,在小样本分类任务中验证了方法的有效性。Yang 等[69]认为实体描述信息过长,限制了模型捕捉有效片段的能力,提出将其替换为简短精炼的实体概念信息,在注意力机制下将其作为外部信息嵌入到句子的语义向量表示空间中,在 FewRel 数据集上模型性能得到提升。上述基于原型向量的方法,能够在少量的标注数据的情况下快速学习新的关系类别,而不需要重新训练整个模型,降低了模型的训练成本,但其在处理复杂和多样的关系类别时,单一的原型向量可能无法充分捕捉其特征,导致模型分类性能下降。在这个问题上,学者们尝试引入多种外部知识源来增强模型的语义表征能力,如实体描述类型、关系描述信息等。

$$\max \log p_\theta(\bar{x}|\hat{x}) \approx \sum_{t=1}^{T} \log m_t p_\theta(x_t|\hat{x}) \tag{2}$$

在基于自编码语言模型的提示学习方法中,多采用 prompt-tuning 的方法,即在输入中插入模板进而将分类任务转化为掩码语言建模问题,并利用预训练语言模型实现微调和推理,如 Chen 等[70]摒弃了硬提示策略,在模板中插入虚拟类型词学习实体类型信息,其作为模板中可学习的变量,随上下文动态变化,用于处理复杂的多标签场景。Han 等[71]认为关系中存在一些预定义的逻辑规则,如 person->'parsent was->person:parsent,这些由头尾实体类型和关系类型构成的先验语义逻辑规则,对判断句子语义关系能够起到辅助作用,提出了以逻辑规则为条件概率的小样本关系分类模型——PTR 模型,在提示中融入了先验知识,在小样本场景下取得了良好的性能,但逻辑规则需要手工设计,这可能需要大量的人力和时间。在未来的研究中,可以进一步使用自动或半自动的方法来生成或优化逻辑规则,例如使用知识图谱、知识库、语言推理等技术。同时,PTR 对于提示之间可能存在的相互作用或冲突未做过多考虑。针对这个问题,在之后的研究工作中,可以考虑使用动态或自适应的方法来选择或组合子提示,例如使用注意力机制、强化学习等技术。总的来看,提示学习为低资源条件下的模型学习提供了新的思路,在不显著改变预训练语言模型结构和参数的情况下,通过向输入中增加"提示信息"来完成下游任务,但提示学习模型架构复杂多变,影响因素众多,在实际训练过程中相比完全监督、微调等范式难度较大。

### 4.3　远程监督条件下的关系分类方法

远程监督方法采用知识库对齐的方式能够快速产生大量伪标注样本,但简单的对齐回标也引入了噪声,如图 3 所示,乔布斯和苹果公司之间被错标为了知识库中已有的 CEO 关系类型。在对含噪数据进行分类时产生了两种视角,一是面向包级别数据,认为给定的包标签仅对于包内的部分句子而言是正确的,这一研究的出发点多是数据降噪,提出了为包内句子分配不同的权重或移除噪声句子等方式来缓解该问题;另一种面向句级别的数据,其出发点是认为包级别降噪仍无法准确得到每一个句子的正确标签,导致无法充分利用语料资源,因此,针对句子级别的远程监督关系分类任务,展开了相关的研究。

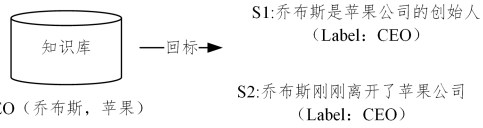



图 3　远程监督噪声实例

Fig. 3　Example of distant supervision noise

#### 4.3.1　包级别的关系分类问题

Zeng 等[72]提出了分段卷积神经网络(Piecewise Convo-

lutional Neural Network, PCNN), 在假设一个包内句子表达同一种关系的前提下, 根据句中两个实体的位置, 分为头实体前、两实体间、尾实体后 3 段分别进行平均池化操作, 得到句子表示特征, 再通过分类器预得到每个句子的类别标签, 选择概率最大的句子标签作为包的标签。该方法只注意到最有可能表达实体对关系的句子, 遗漏了包内其他句子包含的丰富信息, 浪费了大量的资源。Lin 等[73]提出了基于注意力的选择机制(Selective Attention), 认为包内的每个句子对于包关系标签的确定具有不同的影响力, 通过给句子分配不同的注意力, 在充分利用句子信息的前提下, 动态减少噪声数据的影响。类似地, Lin 等[74]提出了一种基于置信度的多实例学习的方法, 对每个实体对下的所有句子进行加权平均, 从而提高了模型的泛化能力。相比 PCNN, 该模型的性能有了进一步的提升, 但以上学者的出发点是基于一个包中存在一种关系类型展开的, 未考虑到语句中存在的一些复杂情况, 如多关系。据统计, 在 NYT 数据集中约 18.3% 的包内同时存在多个关系[17]。针对这个问题, Jiang 等[17]提出了跨句子最大池化的方式来捕获句子间的信息, 如图 4 所示, 并针对每种关系类型建模为二分类问题, 判断是否表达该关系来建模包内的多关系现象。

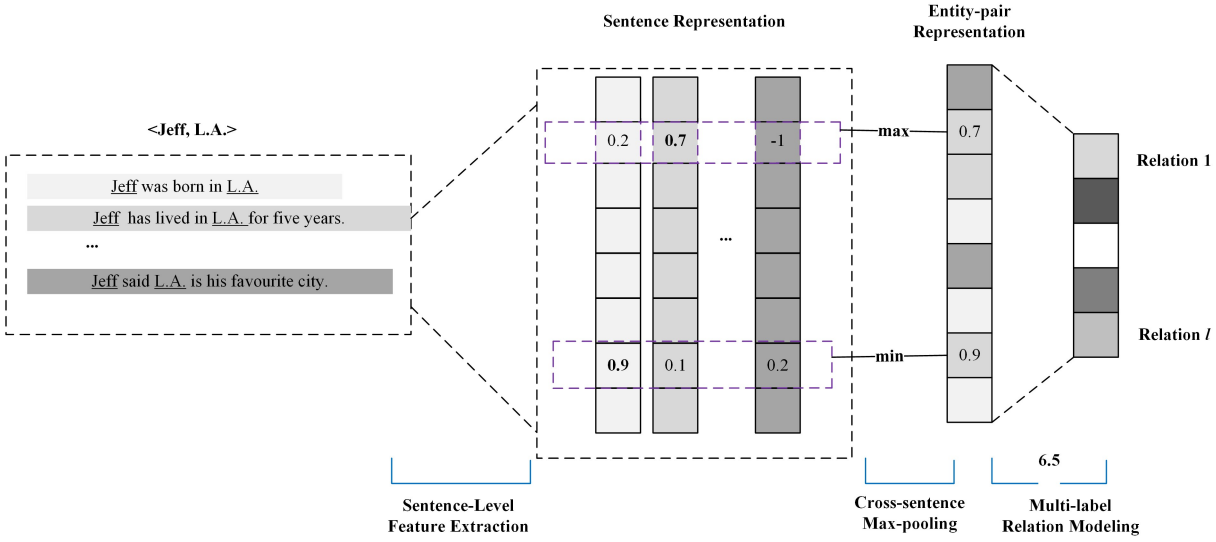



图 4 Jiang 等提出的跨句子最大池化模型

Fig. 4 Cross sentence maximum pooling model proposed by Jiang et al.

以上研究的出发点是考虑如何尽可能地利用包自身携带的有效信息, 提出了跨句子池化、选择注意力的方式来选择包内最具代表性的信息, 但这些方法可能忽略了句子之间的语义相关性和多样性, 以及句子内部的结构信息。为了更为全面地利用包内携带的语义知识, Hogan 等[75]利用知识库中的实体类型和关系类型信息, 对远程监督下的包标签进行细粒度的划分, 得到不同可信度和难度的子集用于预训练, 具体而言是使用对比学习目标函数, 让模型在不同子集之间进行对比学习, 增强模型对正负样本的区分能力和鲁棒性。通过自监督的方式, 能够有效学习数据的内在结构和语义信息, 提高模型的泛化能力和表示能力, 从而提升在关系分类任务上的性能。Hogan 等[75]只在句子级别进行对比学习的策略, 忽略了实体级别和包级别的信息交互, 可能导致模型无法充分利用数据的多样性和一致性, 也无法有效地降低噪声数据的影响。针对这个问题, Dong 等[76]使用了一个分层的对比学习框架(Hierarchical Contrastive Learning Framework), 通过随机替换、删除或添加实体生成更多的训练样本, 分别对实体、句子和包 3 个层次进行表示学习, 并利用多头自注意力机制(Multi-head Self-attention)来实现跨层次的交互, 从而提高了表示的一致性和准确性。以上这些方法都表明, 从内部数据中挖掘潜在知识和规律对提升远程监督抽取任务的性能有重要意义。

随着知识工程的不断推进, 大规模知识图谱、维基百科知识库等逐渐被作为外部信息加入到远程监督抽取任务中, 并在多项实验上得到了验证, 如 Zhang 等[68]提出了互注意力机制下的知识图谱和文本模型的联合学习策略, 在知识图谱的指导下, 实现了对远程监督噪声的削弱。Hu 等[77]认为实体描述信息对实现包降噪同样具有功效, 因此引入实体描述和知识图谱中的关系结构信息, 得到实体描述和关系向量表示, 建立二级注意力模型以筛选置信度较高的实例。Zhang 等[78]认为, 在关系层级结构图中, 具有相同上层父节点的关系相比不同的父节点而言, 关系之间的相似度较大, 而关系层级越高, 训练样本数量就越多, 因此提出了利用关系层次结构信息的降噪模型, 在底层数据中融入父级关系节点包含的语义信息, 弥补了长尾关系实例不足的缺陷。这些研究表明, 在远程监督抽取任务中, 利用外部知识和结构信息能够有效提升模型的性能。

#### 4.3.2 句级别的关系分类问题

句级别关系抽取以句子为落脚点, 目的是确定单句的标签, 常采用迭代训练的方式。Feng 等[16]提出了基于强化学习的句子选择模型, 通过实例选择器和关系分类器两个模块分别获取高质量句子和完成关系分类器的训练, 有效地降低了噪声数据对关系分类性能的影响, 并采用迭代训练的方式增加了样本量, 进而改善了模型的性能。这种方法是以每个实体对只有一种关系为假设前提, 忽略了多关系等复杂情形, 并且该方法依赖于远程监督数据中包内存在至少一个正确

描述对应关系的句子，如果一个包中所有的句子都是噪声，那么该方法无法有效地过滤掉这些噪声。不同于 Feng 等[16]采用强化学习选择实例的方式，Wei 等[79]依据每种关系下的语法模式筛选得到高质量实例输送给关系分类器，在关系分类模型中利用关系模式来约束注意力机制，使得模型在预测关系时更加专注于与关系相关的词语，在构建数据集时选择得分高的实例作为正实例，将其模式添加到关系模式中作为新的模式集，迭代筛选新的实例。这种采用注意力正则化的训练方法，使得模型更关注与关系模式相关的信息，增强了模型的可解释性。以上视角依赖于包级别的标签进行高质量句子选择，其关键在于如何清洗数据，为关系分类器选择高质量样本进而降低噪声对模型训练的影响，但这种直接采用包级别标签作为句子预定义标签的方式，也为句子关系分类引入了一定的噪声。

不同于以上基于正向训练的视角，即试图从噪声数据中筛选出高质量的实例作为模型的训练数据集，Ma 等[80]在句子级别的远程监督关系抽取任务中，从负向训练视角出发，提出为句子定义互补标签的方式构建负实例样本，即为该包内的句子分配其他包的关系标签，整个训练过程被拆解为关系分类、噪声过滤和重标签 3 步，采用迭代训练的方式，动态阈值过滤噪声句子并重标记为模型预测的关系标签，但在分配互补标签时，这种方式仍存在以低概率引入正确标签的可能性。Ma 等只考虑了数据集中的假正例错误，忽视了数据集中的其他错误类型。针对此问题，Xie 等[81]考虑了远程监督数据集的质量，发现数据集中存在两种错误：一是知识库不全导致的假负例，二是真负例。针对假负例现象，会扰乱分类器的学习，提出将部分视为无标签的正样本在一定的权值下接受该样本作为真实例数据，避免了远程监督分类任务中被错误判别的现象。可以看出，负向视角下更为关注如何减少假正例、假负例数据对模型性能的影响。

## 5　限定域下的三元组抽取方法

在三元组抽取任务中，需识别句中实体和实体对之间的关系类型。前人的研究多是基于两个角度：一是基于单任务视角[21,28,82]，即设计特殊的结构，将两个任务以统一的架构表示出来，如基于序列标注的方法、基于表填充的方法；二是基于多任务学习的视角[20]，将实体识别和关系预测作为两个独立的任务，实现实体和关系的联合抽取。

### 5.1　单任务视角下的联合抽取

#### 5.1.1　基于序列标注的方法

基于序列标注的方法是为文本中的每个词分配一个标签，根据标签解码生成实体，常用的标注体系有“BIO”(B-begin/I-in/O-others)和“BIOES”(B-begin/I-in/O-others/E-end/S-single)。其一般形式为：假设句子 $S=\{x_1,x_2,x_3,\cdots\}$，使用标签体系为“BIO”(或者“BIOES”)，该句对应的真实标签序列为 $Y=\{y_1,y_2,y_3,\cdots\}$。标注任务转化为寻找最优状态序列 $\arg\max P(y_1,y_2,\cdots,y_n|x_1,x_2,\cdots,x_n)$，根据隐马尔可夫模型，得到：

$$\arg\max P(y_1,y_2,\cdots,y_n|x_1,x_2,\cdots,x_n)=\arg\max\prod_{i=1}^{n}p(y_i|h_i;w) \tag{3}$$

其中，$h_i$ 为第 $i$ 个位置令牌的语义表征，$\omega$ 为模型参数。

Zheng 等[83]提出了层次神经网络框架，采用 LSTM 作为编码层来捕获单词之间的长距离依赖关系，实体识别(Named Entity Recognize，NER)和关系抽取模块共享编码层，在解码过程中采用序列标注的机制训练 NER 模块得到实体，每个词的标签是由表示实体类别和边界的一套符号拼接而成，如 B-Loc，B 表示实体的开始位置，Loc 表示实体的类型信息。随后 Zheng 等[84]基于此，在原有的标签上加上关系类型信息，设计了适用于关系抽取的标签体系，提出了基于序列标注的联合抽取方法，如图 5 所示，其标签体系是由实体位置(BIOS)和关系类型信息构成的，如 B-CP-1，其中 B 表示实体的首令牌，CP 为关系类型，1 表示三元组中的头实体。在为句子中的每个词分配不同的标签后，采用条件随机场(Conditional Random Field，CRF)解码的方式得到句子中的三元组。基于序列标注的方法能够实现端到端的三元组抽取，简单易实现，但这种方法默认句中令牌只能被分配一个标签，表达能力弱，导致在建模重叠关系等多标签场景时具有一定的难度。

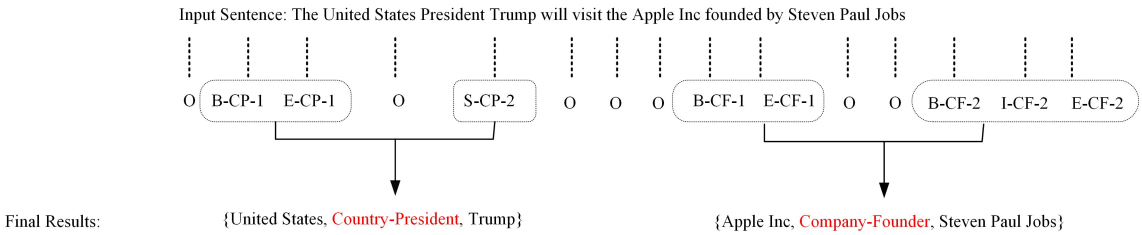



图 5　序列标注实例图

Fig. 5　Example of sequence annotation

#### 5.1.2　基于表填充的方法

基于表填充的方法是将实体和关系映射到二维矩阵中，再将矩阵中的值映射解码得到三元组的方法。Ren 等[85]提出了一种基于表格填充的关系三元组抽取模型，能够利用全局特征来提高抽取效果。该模型首先为每个关系生成一个表格特征，然后从表格特征中挖掘两种全局关联：关系之间的互斥性和实体之间的共现性。不同于 Ren 等[85]以两个独立标签空间建模实体和关系的形式，Wang 等[86]以表填充的方式，将关系和实体映射到统一的标签空间内，其基本出发点是认为实体词之间的欧氏距离应小于非实体词之间的欧氏距离，采用对角线区域表示实体、非对角线区域表示关系的方式，同时建模了重叠关系、有向关系等复杂场景，但对于嵌套实体和不连续实体的建模能力较弱，且解码过程的难度较大。为了加快模型训练和推理速度，降低模型的解码难度，Wang 等[28]提出了 TPLinker 模型，其设计了一种新的标签机制，将句子中的词两两组队，定义了 EH-to-ET(实体头部-实体尾部)、

SH-to-OH(头实体头部-尾实体头部)、ST-to-OT(头实体尾部-尾实体尾部)3 种关系连接机制,当任意一种关系机制存在时,对应位置数字变为非 0,以此确定三元组的实体边界和关系类型,提高了复杂关系和嵌套实体情景下的表现效果。类似地,Shang 等[11]使用 HB-TB,HB-TE,HE-TE 作为标签体系,其中 H 和 T 表示头尾实体,B 和 E 表示实体头部和尾部,-表示非实体,在每个关系类型对应的二维矩阵中填充对应的关系标签,以此解码得到三元组。这两种方式的本质是将关系三元组映射在关系二维表中,对语句中的每一对令牌进行分类,判断是否构成实体和存在关系,标签空间设计简单,加快了训练和推理速度,但在模型中对不同长度和类型的三元组未进行区分,而是使用了相同的分类器和阈值,这可能导致一些长距离或稀有类型的三元组性能不佳。

#### 5.1.3 其他方法

Zeng 等[87]基于对话生成的机理,提出了基于复制机制的端到端抽取模型,其模型如图 6 所示。

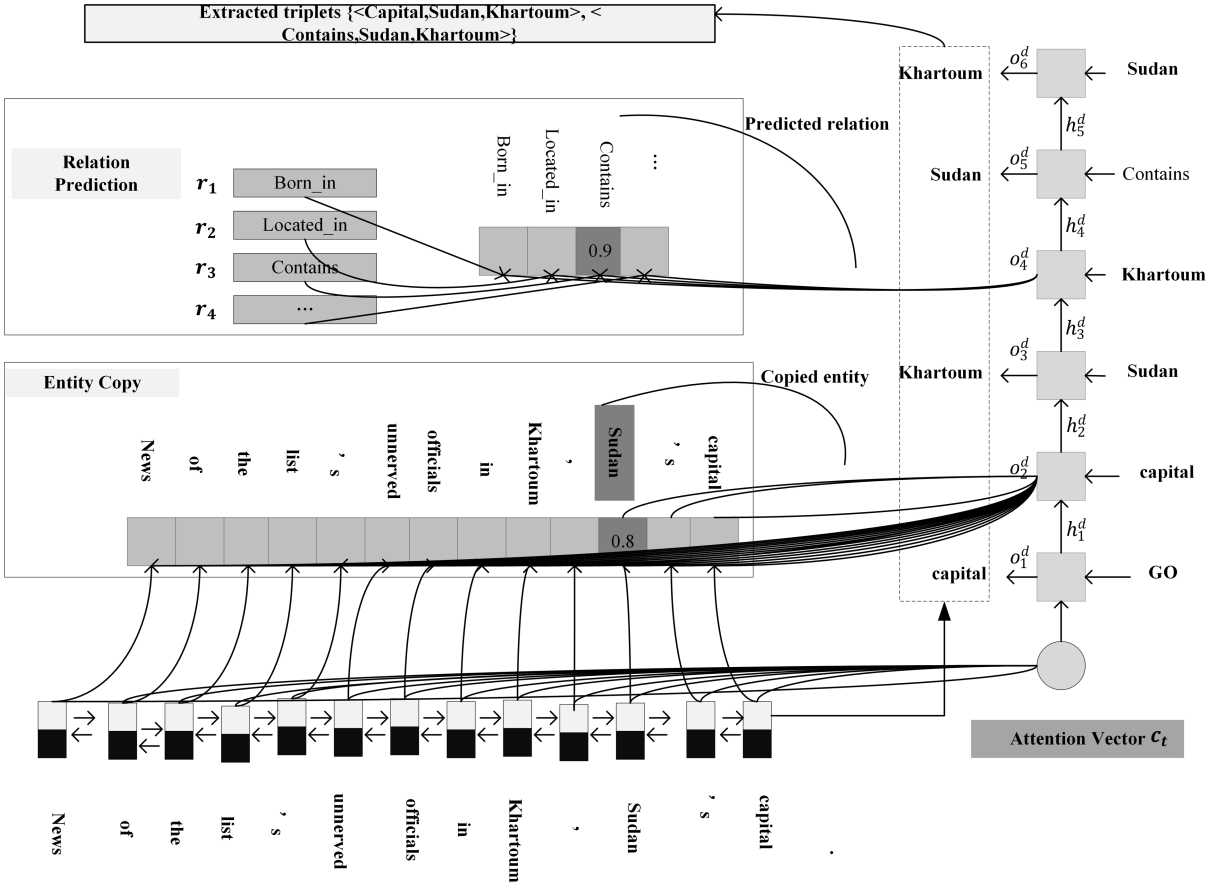



图 6 基于复制机制的三元组识别模型

Fig. 6 Triple recognition model based on replication mechanism

该模型采用编码器将自然语句编码为定长的语义向量,利用解码器解码为三元组。在解码过程中,先预测其关系再利用拷贝机制拷贝句子中的第一个实体,然后拷贝第二个实体,直至句子中不存在三元组。该模型在 NYT 数据集和 WebNLG 数据集上性能提升了 10%左右,有效地解决了句中多三元组的情景,但该方法并未考虑到实体之间的关系,没有充分利用实体信息。为了更有效地利用实体信息,Li 等[88]将三元组抽取任务建模为多轮对话形式,利用问题查询来编码重要的信息,依靠阅读理解技术,抽取得到问题的答案,解码得到三元组。其流程大致为:先判断头实体是否存在,再将第一步的头实体填充到每种关系下的问题槽中,由模型根据问题判断是否存在尾实体,从而解码得到三元组。这种方法可以灵活地处理不同领域和不同结构的数据,但需根据具体关系类型手工定义关系模板,人工参与过多,模型自主性较差。上述方法使用交叉熵作为模型参数更新的损失函数,但交叉熵损失函数对文本生成顺序过度依赖,而三元组彼此之间是无顺序的,这导致模型参数错误传播和冗余计算。针对这个问题,Sui 等[89]提出了一种基于集合表示句中三元组的方法,利用 Transformer 编码器和集合解码器,将实体和关系的预测转化为集合的预测,引入了一种基于集合的损失函数,避免了传统方法中交叉熵损失函数受文本生成顺序的影响,但模型复杂度较高,模型计算资源需求量大。

### 5.2 多任务视角下的联合抽取

相比单一框架下的三元组抽取而言,多任务学习下的三元组抽取将实体和关系作为两个子任务看待,大多数研究中共享编码层,少量研究采用了管道式方法,保持两个子任务之间的完全独立性。

Bekoulis 等[90]提出了多头选择机制的联合抽取策略,在采用 BIO 机制编码实体,由 CRF 解码得到实体的基础上,识别每一个实体对应的尾实体和关系,将关系抽取建模为多头选择问题,用于解决多元组情形,忽略了实体和关系之间的

交互信息。为了更好地捕捉语句中主语、关系和客体之间的依赖信息,Wei 等[91]提出了级联二元标签抽取框架 Casrel,其模型如图 7 所示。将关系建模为主语映射到句子中宾语的函数,在识别主语的前提下,确定每种关系类型下的宾语,整个任务表现为二分类任务的形式,实现了端到端的三元组抽取,综合考虑了句子中的嵌套实体和复杂关系。相比序列标注的方式,在 NYT 数据集和 webNLG 数据集上准确率分别提升了 27.3%和 40.9%,取得了显著的改进效果,但模型的损失函数设计是以最大化主语的识别准确率和最大化特定主语和关系的情形下宾语识别准确率为目标,和最终的评价指标——三元组准确率并非直接关联。另外模型每次只能处理句子中的一个主语,导致模型效率低下。为了缓解 Casrel 模型中的关系预测冗余问题,即对于同一个主语,需要对每个可能的客体都进行关系预测。

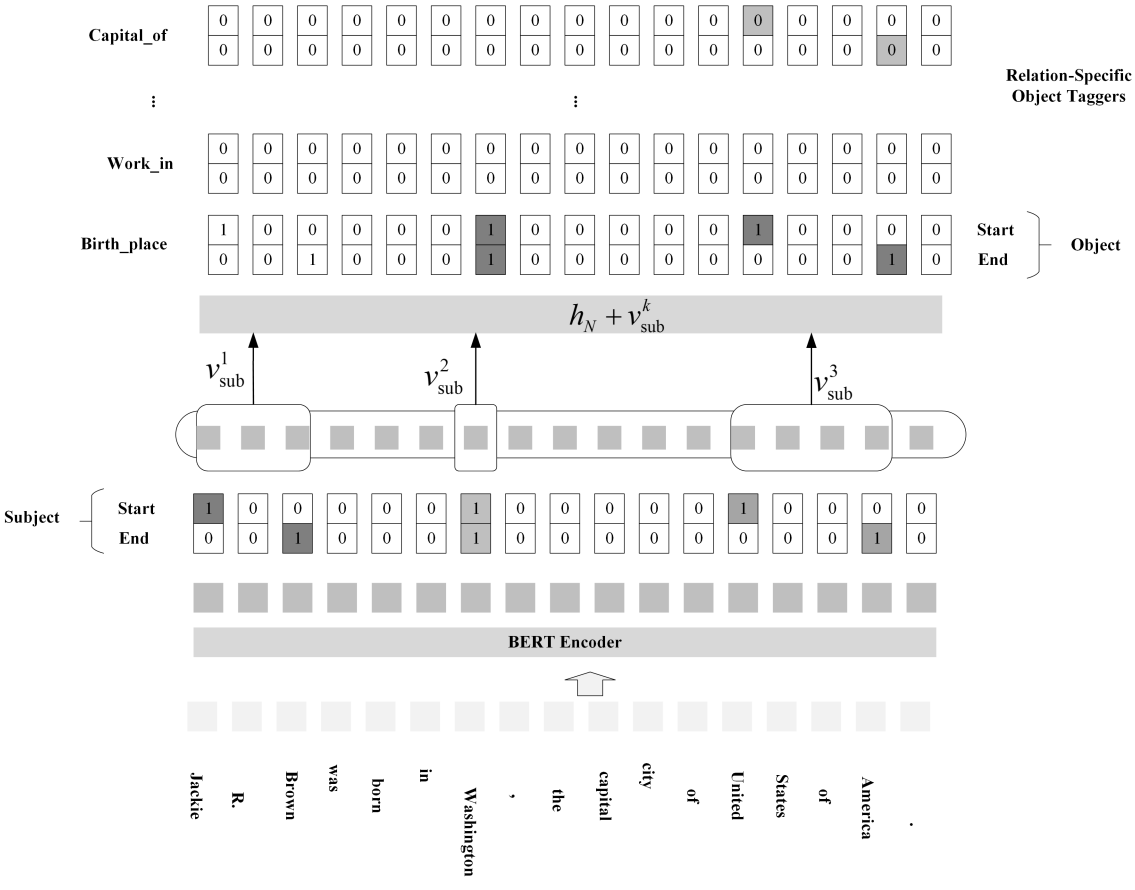



图 7　Wei 等提出的级联二元标签抽取框架

Fig. 7　Cascaded binary label extraction framework proposed by Wei et al.

Zheng 等[82]将联合抽取任务分解为关系预测、实体预测、每种关系下的头尾实体匹配问题,在关系预测阶段,生成头尾实体匹配度得分矩阵,矩阵中分值越高,则说明两个令牌之间存在某种关系;在头尾实体识别阶段,预测可能的实体;在头尾实体匹配阶段,利用前两个阶段的结果,解码生成最终的三元组。该方法根据句子语义表示向量训练得到的全局共现矩阵,为实体对匹配提供了新思路,有效解决了 SEO,EPO 问题。但模型同时训练 3 个子任务,并且使用了多头自注意力机制,增加了计算复杂度和内存消耗。Li 等[92]提出了翻译解码框架 TDEER,用于实现三元组抽取,将关系视为从主体到客体的翻译操作,即先预测所有可能的主语实体,然后将每个主语实体和关系类型相结合,用于预测其对应的客体实体。TDEER 模型可以避免 PRGC 模型中的实体检测问题,即对于同一个关系类型,只需要预测一次主体实体和客体实体的边界,而不是对每个可能的实体边界都进行检测。与 Casrel 根据主语实体来预测关系类型的方式不同,TDEER 是根据主语实体和关系类型组合来预测客体实体,但这两种方法均需要设计合适的阈值来过滤掉错误或冗余的预测结果。TDEER 通过加法操作同时建模了主体实体和关系之间的依赖关系。Ye 等[93]直接使用打包标记考虑实体对之间的依赖关系,提出了一种基于跨度(Span)表示的实体和关系抽取方法,通过在编码器中巧妙地打包标记(Marker)来实现关系抽取。具体而言,使用一个特殊的标记(Levitated Marker),将其插入到每个候选实体的起始位置和结束位置之间,形成一个新的序列。然后,使用一个打包函数(Packing Function),将每个候选实体对应的两个 levitated marker 合并为一个 packed marker,并将其放置在原始序列中最接近该实体对的位置。最后,使用一个分类器(Classifier),根据 packed marker 和原始序列中相应位置的向量来预测实体类型和关系类型。上述方法采用了文本序列模型来学习语义表征。Fu 等[94]提出了基于图神经网络的联合实体和关系抽取方法,它将文本表示为一个包含实体和关系的关系图,然后使用图神经

网络进行图结构的学习和推理。这种表示方式能够更好地建模实体和关系之间的复杂依赖关系，但也可能受限于图结构的复杂度和稀疏性，导致信息传播不充分或过度平滑。

在以上方法结构中，实体识别和关系抽取两个子任务共用编码层，但由于任务特征不同，所需语义表征可能也并不相同。为了避免子任务之间的特征相互干扰，Yan 等[95]将每个神经元分为 3 部分，其中前两部分分别单独用于实体识别任务和关系分类任务，最后是两任务共用的部分，将得到的每个子任务的语义表征用于表填充建模，若存在实体(关系)，表中对应位置由 0 更改为 1。这种特征学习方式既实现了子任务之间的特征独立，也保证了子任务间的特征交互，模型在 NYT[25]，WebNLG[26]等数据集上取得了良好的性能。

在采用管道式抽取方法解决三元组抽取问题时，实体识别和关系分类作为两个独立的过程分开训练。实体识别常用技术如表 4 所列，关系分类则可参见本文中的第 4 章。

表 4 实体识别常用技术范式

Table 4 Common technology paradigm of entity recognition

| 方法类型 | 含义 |
|---|---|
| 基于序列标注的方法[96-97] | 为文本中的词分配一个标签，常采用“BIO”和“BIOS”标注实体位置，将位置标签和类型组合作为实体词的标签，如 B-location |
| 基于跨距(span)的方法[2,98] | 将 NER 任务视为识别文本中的实体片段并对片段进行分类的过程 |
| 基于指针(Pointer)的方法[99] | 基于指针的方法通常采用 $T\times N\times N$ 的矩阵表示句子序列，$T$ 代表实体类型，$N$ 代表句子长度，以矩阵的对角线为实体片段起始位置，在矩阵的上三角中找到实体片段对应的终止位置，标注为 1，根据矩阵确定句中实体 |

**结束语** 在限定域关系抽取任务上，本文研究整理了相关数据集和工具，并探讨和分析了深度学习模型架构下具有代表性的关系分类和关系抽取方法。尽管学者们从知识增强[22,69,77,100-103]、模型架构设计等角度对以上任务提出了有效的改进策略，提升了模型性能，但在如下几个方面，本文认为尚有不足，仍具有深入研究的必要。

1)数据集的质量。在每一项任务下，数据集质量直接影响了模型的学习性能，但数据的标注质量并非总是令人满意，存在错标、漏标[29]等，这可能导致模型学习到的知识是错误的。如何针对数据集中的真假负例、丢失实例，提出有效的判别和补全策略来提高数据集知识的准确率，对于模型效果提升和鲁棒性增强具有一定的帮助。

2)在关系分类和关系抽取任务中，真实情景下的人工标注高质量数据集往往十分稀缺，如何在低数据资源的情况下，如零样本关系分类、零样本三元组抽取[104]等场景下，更好地引导大规模预训练语言模型完成任务，需要深入的研究。目前，提示学习已经崭露头角，但提示学习中的提示模板构建[59,105]、答案映射[54-55]等相关研究问题的研究尚浅。

3)在关系分类或关系抽取模型的鲁棒性上，需要深入探讨。在真实场景中，噪声是多样的，如语料库的语句存在同义字、同音字、删除[106]等噪声现象。因此，需要通过有效的模型设计和优化来提高模型的鲁棒性，以适应复杂多变的真实场景。

在关系分类和关系抽取任务中，多样的知识增强方式也是学者们探讨的热点之一。现有的手段多是注入知识图谱信息、实体相关信息、关系层次信息等文本类型的知识，将其作为模型的外部信息源。但我们的世界是多模态的，其他形式的信息，如视频、音频、图片等，对于关系分类和抽取任务仍具有一定的帮助，如何在避免噪声的条件下，引入对任务有用的多模态信息[107-108]，也是值得探究的问题。同时，在模型性能的影响因素方面，目前从关系复杂度、实体类型、数据分布等数据角度[109]进行了分析，以但在模型角度仍需要进行更为深入的分析，增强模型的可解释性。

## 参 考 文 献


[1] NGUYEN T H, GRISHMAN R. Relation Extraction: Perspective from Convolutional Neural Networks[C]// Proceedings of NAACL-HLT 2015. 2015: 39-48.

[2] FU J, HUANG X, LIU P. SpanNER: Named Entity Re-/Recognition as Span Prediction[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, 2021: 7183-7195.

[3] YI L, HE L, MARI O, et al. Multi-Task Identification of Entities, Relations, and Coreference for Scientific Knowledge Graph Construction[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium, Association for Computational Linguistics, 2018: 3219-3232.

[4] WALKER C, STRASSEL S, MEDERO J, et al. ACE 2005 multilingual training corpus[OL]. (2006) [2019-09-13]. https://www.cdc.gov/violenceprevention/aces/about.html.

[5] ZHOU B, ZHAO H, PUIG X, et al. Scene parsing through ADE20K dataset[OL]. (2017) [2020-05-23]. https://groups.csail.mit.edu/vision/datasets/ADE20K.

[6] LI D M, ZHANG Y, LI D Y, et al. A survey on entity relation extraction methods [J]. Journal of Computer Research and Development, 2020, 57(7): 1424-1448.

[7] ZHUANG C Z, JIN X L, ZHU W J, et al. A survey on relation extraction based on deep learning [J]. Journal of Chinese Information Processing, 2019, 33(12): 1-18.

[8] NAYAK T, MAJUMDER N, GOYAL P, et al. Deep Neural Approaches to Relation Triplets Extraction: a Comprehensive Survey[J]. Cognitive Computation, 2021, 13(5): 1215-1232.

[9] ZHANG S W, WANG X, CHEN Z R, et al. A survey on supervised joint entity and relation extraction methods [J]. Journal of Frontiers of Computer Science and Technology, 2022, 16(4): 1-24.

[10] BAI L, JIN X L, XI P B, et al. A survey on relation extraction based on distant supervision [J]. Journal of Chinese Information Processing, 2019, 33(10): 10-17.

[11] SHANG Y, HUANG H, MAO X. OneRel: Joint Entity and Relation Extraction with One Module in One Step[OL]. (2022-03) [2022-5-27]. https://arxiv.org/abs/2203.05412v1.

[12] ZHAO K, XU H, YANG J, et al. Consistent Representation Learning for Continual Relation Extraction[C]// Findings of the

Association for Computational Linguistics. Association for Computational Linguistics,2022:3402-3411.

[13] QIN P,XU W,GUO J. Designing an adaptive attention mechanism for relation classification[C]// 2017 International Joint Conference on Neural Networks(IJCNN) IEEE. 2017:4356-4362.

[14] SARKAR S,ARZOO K,REBECCA P,et al. CONTAINER Few-Shot Named Entity Recognition via Contrastive learning [C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics(Volume 1:Long Papers). Association for Computational Linguistics,2022:6338-6353.

[15] YU H,ZHANG N,DENG S,et al. Bridging Text and Knowledge with Multi-Prototype Embedding for Few-Shot Relational Triple Extraction[C]// Proceedings of the 28th International Conference on Computational Linguistics. International Committee on Computational Linguistics,2020:6399-6410.

[16] FENG J,HUANG M,ZHAO L,et al. Reinforcement Learning for Relation Classification from Noisy Data[C]// Association for Advancement Artificial Intelligence. New Orleans, LA, 2018:5779-5786.

[17] JIANG X,WANG Q,LI P,et al. Relation Extraction with Multi-instance Multi-label Convolutional Neural Networks [C]// The 26th International Conference on Computational Linguistics(COLING 2016). The COLING 2016 Organizing Committee,2016:1471-1480.

[18] WU F,WELD D. Autonomously semantifying wikipedia[C]// The 16th ACM International Conference on Information and Knowledge Management(CIKM 2007). Association for Computing Machinery,2007:41-50.

[19] MIKE M,STEVEN B,RION S,et al. Distant supervision for relation extraction without labeled data[C]// The Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP(ACL-IJCNLP 2009). Association for Computational Linguistics,2009:1003-1011.

[20] ZHONG Z,CHEN D. A Frustratingly Easy Approach for Entity and Relation Extraction[C]// Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics:Human Language Technologies. Association for Computational Linguistics,2021:50-61.

[21] TIAN Y,CHEN G,SONG Y,et al. Dependency-driven Relation Extraction with Attentive Graph Convolutional Networks[C]// The 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing(ACL-IJCNLP 2021). Association for Computational Linguistics,2021:4458-4471.

[22] LAI T,JI H,ZHAI C,et al. Joint Biomedical Entity and Relation Extraction with Knowledge-Enhanced Collective Inference[C]// The 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing(ACL-IJCNLP 2021). Association for Computational Linguistics,2021:6248-6260.

[23] ZHANG Y,CHEN W,LI Y,et al. The TAC Relation Extraction Dataset[OL]. (2018-12-15) [2020-02-20]. https://catalog.ldc.upenn.edu/LDC2018T24.

[24] GAO T,HAN X,ZHU H,et al. FewRel 2.0:Towards More Challenging Few-Shot Relation Classification[C]// The 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Association for Computational Linguistics,2019:6250-6255.

[25] SEBASTIAN R,YAO L,ANDREW M. The New York Times Company[OL]. (2010) [2021-01-21]. http://iesl.cs.umass.edu/riedel/ecml.

[26] GARDENT C,SHIMORINA A,NARAYAN S,et al. Creating Training Corpora for NLG Micro-Planning[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics(Volume 1:Long Papers). Association for Computational Linguistics,2017:179-188.

[27] YUAN X. Schema-based-Knowledge-Extraction [OL]. (2020-04) [2023-3-13]. https://github.com/yuanxiaosc/Schema-based-Knowledge-Extraction.

[28] WANG Y,YU B,ZHANG Y,et al. TPLinker:Single-stage Joint Extraction of Entities and Relations Through Token Pair Linking[C]// Proceedings of the 28th International Conference on Computational Linguistics Barcelona. International Committee on Computational Linguistics,2020:1572-1582.

[29] ALT C,GABRYSZAK A,HENNIG L. TACRED Revisited:A Thorough Evaluation of the TACRED Relation Extraction Task [C]// Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics,2020:1558-1569.

[30] ZHOU G Z,SU J,ZHANG J,et al. Exploring Various Knowledge in Relation Extraction[C]// Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05). Association for Computational Linguistics,2005:427-434.

[31] ZHOU G,ZHANG M,JI D H,et al. Tree Kernel-Based Relation Extraction with Context-Sensitive Structured Parse Tree Information[C]// Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Association for Computational Linguistics,2007:728-736.

[32] ZHANG D,WANG D. Relation Classification via Recurrent Neural Network[OL]. (2015-08) [2022-02-21]. https://go.exlibris.link/swcDB6f5.

[33] ZHANG C,CUI C,GAO S,et al. Multi-Gram CNN-Based Self-Attention Model for Relation Classification[J]. IEEE Access,2019,7:5343-5357.

[34] ZHANG S,ZHENG D,HU X,et al. Bidirectional Long Short-Term Memory Networks for Relation Classification[C]// Proceedings of the 29th Pacific Asia Conference on Language,Information and Computation. 2015:73-78.

[35] ZHOU P,SHI W,TIAN J,et al. ,Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification [C]// Proceedings of the 54th Annual Meeting of the Associa-

tion for Computational Linguistics (Volume 2: Short Papers). Association for Computational Linguistics, 2016: 207-212.

[36] LEE J, SEO S, CHOI Y S. Semantic Relation Classification via Bidirectional LSTM Networks with Entity-Aware Attention Using Latent Entity Typing[J]. Symmetry, 2019, 11(6): 785.

[37] SANTOS C N D, XIANG B, ZHOU B. Classifying Relations by Ranking with Convolutional Neural Networks[C]// 53rd Annual Meeting of the Association for Computational Linguistics (ACS)/7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing(IJCNLP). Assoc Computational Linguistics-Acl, 2015: 626-634.

[38] ZHANG C, CUI C, GAO S, et al. Multi-Gram CNN-Based Self-Attention Model for Relation Classification[J]. IEEE Access, 2019, 7: 5343-5357.

[39] ZHANG Y, QI P, MANNING C D. Graph Convolution over Pruned Dependency Trees Improves Relation Extraction[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2018: 2205-2215.

[40] GUO Z, ZHANG Y, LU W. Attention Guided Graph Convolutional Networks for Relation Extraction[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2019: 241-251.

[41] ZHOU L, WANG T, LIU Y. AWeighted GCN with Logical Adjacency Matrix for Relation Extraction[C]// The 24th European Conference on Artificial Intelligence. 2020: 2314-2321.

[42] ZENG D, LIU K, LAI S, et al. Relation classification via convolutional deep neural network[C]// Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers. Association for Computational Linguistics, 2014: 2335-2344.

[43] ZHOU L, WANG T, QU H, et al. A Weighted GCN with Logical Adjacency Matrix for Relation Extraction[C]// 24th European Conference on Artificial Intelligence (ECAI 2020). 2020: 2314-2321.

[44] MATTHEW E, MARK N, MOHIT I, et al. Deep contextualized word representations[C]// Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics, 2018: 2227-2237.

[45] RADFORD A, NARASIMHAN K, SALIMANS T. Improving Language Understanding by Generative Pre-Training[OL]. (2018)[2023-3-13]. https://s3-us-west-2. amazonaws. com/openai-assets/research-covers/language-unsupervised/language_understanding_paper. pdf.

[46] MIKE L, LIU Y, NAMAN G, et al. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension[C]// Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2020: 7871-7880.

[47] RAFFEL C, SHAZEER N, ROBERTS A. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer [OL]. (2019-10)[2021-5-17]. https://arxiv. org/abs/1910. 10683.

[48] HU S, DING N, WANG H, et al. Knowledgeable Prompt-tuning: Incorporating Knowledge into Prompt Verbalizer for Text Classification[C]// Findings of the Association for Computational Linguistics: ACL 2022. Association for Computational Linguistics, 2022: 2225-2240.

[49] CHIA Y K, BING L, PORIA S, et al. RelationPrompt: Leveraging Prompts to Generate Synthetic Data for Zero-Shot Relation Triplet Extraction[C]// Findings of the Association for Computational Linguistics: ACL 2022. Association for Computational Linguistics, 2022: 45-57.

[50] LU Y, LIU Q, DAI D, et al. Unified Structure Generation for Universal Information Extraction[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2022: 5755-5772.

[51] KOJIMA T, GU S S, REID M, et al. Large Language Models are Zero-Shot Reasoners[J]. arXiv: 2205. 11916, 2021.

[52] TAYLOR S, YASAMAN R, ROBERT L, et al. AUTOPROMPT: Eliciting Knowledge from Language Models with Automatically Generated Prompts[C]// Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing(EMNLP). Association for Computational Linguistics, 2020: 4222-4235.

[53] QIN G, JASON E. Learning How to Ask: Querying LMs with Mixtures of Soft Prompts[C]// Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics, 2021: 5203-5212.

[54] TIMO S, HELMUT S, HINRICH S. Automatically Identifying Words That Can Serve as Labels for Few-Shot Text Classification[C]// Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain. International Committee on Computational Linguistics, 2020: 5569-5578.

[55] OSCAR S, OIER L, GORKA L, et al. Label Verbalization and Entailment for Effective Zero- and Few-Shot Relation Extraction [C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing Online and Punta Cana, Dominican Republic. Association for Computational Linguistics, 2021: 1199-1212.

[56] CUI G, HU S, DING D, et al. Prototypical Verbalizer for Prompt-based Few-shot Tuning[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Dublin, Ireland. Association for Computational Linguistics, 2022: 7014-7024.

[57] FABIO P, TIM R, PATRICK L, et al. Language Models as Knowledge Bases? [J]. arXiv: 1909. 01066, 2019.

[58] ZHANG P, LU W. Better Few-Shot Relation Extraction with Label Prompt Dropout[C]// Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, Abu

Dhabi. Association for Computational Linguistics, 2022: 6996-7006.

[59] ZHAO T Z, WALLACE E, FENG S, et al. Calibrate Before Use: Improving Few-Shot Performance of Language Models[J]. arXiv: 2102.09690, 2021.

[60] JIANG ZB, XU F, JUN A. How Can We Know What Language Models Know[C]// Transactions of the Association for Computational Linguistics, Volume 8. MIT Press, 2020: 423-438.

[61] CUI L, WU Y, LIU J, et al. Template-Based Named Entity Recognition Using BART[C]// Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. Association for Computational Linguistics, 2021: 1835-1845.

[62] LEE D, KADAKIA A, TAN K, et al. Good Examples Make A Faster Learner: Simple Demonstration-based Learning for Low-resource NER[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2021: 2687-2700.

[63] CHEN J W, LIU Q, LIN H Y, et al. Few-shot Named Entity Recognition with Self-describing Networks[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2022: 5711-5722.

[64] JACOB D, CHANG M W, KENTON L. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [C]// 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). Association for Computational Linguistics, 2018: 4171-4186.

[65] HAN X, ZHU H, YU P, et al. FewRel: A Large-Scale Supervised Few-Shot Relation Classification Dataset with State-of-the-Art Evaluation[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2018: 4803-4809.

[66] DONG M, PAN C, LUO Z. MapRE: An Effective Semantic Mapping Approach for Low-resource Relation Extraction[C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing Online and Punta Cana, Dominican Republic. Association for Computational Linguistics, 2021: 2694-2704.

[67] LIU Y, HU J, WAN X, et al. A Simple yet Effective Relation Information Guided Approach for Few-Shot Relation Extraction [C]// Findings of the Association for Computational Linguistics: ACL 2022. Association for Computational Linguistics, 2022: 757-763.

[68] ZHANG Z, XU H, LIU Z Y, et al. ERNIE: Enhanced Language Representation with Informative Entities[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2019: 1441-1451.

[69] YANG S, ZHANG Y F, NIU G L, et al. Enhanced few shot relation extration through concept[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers). Association for Computational Linguistics, 2021: 987-991.

[70] CHEN X, ZHANG N, XIE X, et al. KnowPrompt: Knowledge-aware Prompt-tuning with Synergistic Optimization for Relation Extraction[C]// The Web Conference 2021. Association for Computing Machinery, 2022: 2778-2788.

[71] HAN X, ZHAO W, DING N, et al. PTR: Prompt Tuning with Rules for Text Classification[OL]. (2021-05) [2022-03-24]. https://arxiv.org/abs/2105.11259.

[72] ZENG D J, LIU K, CHEN Y B, et al. Distant Supervision for Relation Extraction via Piecewise Convolutional Neural Networks[C]// Distant Supervision for Relation Extraction via Piecewise Convolutional Neural Network. Association for Computational Linguistics, 2015: 1753-1762.

[73] LIN Y K, SHEN S Q, LIU Z Y, et al. Neural Relation Extraction with Selective Attention over Instances[C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2016: 2124-2133.

[74] LIN X Y, LIU T Y, JIA W J, et al. Distantly Supervised Relation Extraction using Multi-Layer Revision Network and Confidence-based Multi-Instance Learning[C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing Online and Punta Cana, Dominican Republic. Association for Computational Linguistics, 2021: 165-174.

[75] HOGAN W, LI J C, SHANG J B. Fine-grained Contrastive Learning for Relation Extraction[C]// Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing Abu Dhabi. Association for Computational Linguistics, 2022: 1083-1095.

[76] DONG Y L, ZHANG T L, NAN H, et al. HiCLRE: A Hierarchical Contrastive Learning Framework for Distantly Supervised Relation Extraction[C]// Findings of the Association for Computational Linguistics: ACL 2022. Association for Computational Linguistics, 2022: 2567-2578.

[77] HU L, ZHANG L, SHI C, et al. Improving Distantly-Supervised Relation Extraction with Joint Label Embedding[C]// Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Association for Computational Linguistics, 2019: 3821-3829.

[78] ZHANG N, DENG S, SUN Z, et al. Long-tail Relation Extraction via Knowledge Graph Embeddings and Graph Convolution Networks[C]// 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). Association for Computational Linguistics, 2019: 3016-3025.

[79] WEI J, DAI D, XIN X Y. ARNOR: Attention Regularization based Noise Reduction for Distant Supervision Relation Classification[C]// Proceedings of the 57th Annual Meeting of the As-

sociation for Computational Linguistics. Association for Computational Linguistics, 2019: 1399-1408.

[80] MA R, GUI T, LI L, et al. SENT: Sentence-level Distant Relation Extraction via Negative Training[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, 2021: 6201-6213.

[81] XIE C, LIANG J, LIU J, et al. Revisiting the Negative Data of Distantly Supervised Relation Extraction[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, 2021: 3572-3581.

[82] ZHENG H, WEN R, CHEN X, et al. PRGC: Potential Relation and Global Correspondence Based Joint Relational Triple Extraction[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, 2021: 6225-6235.

[83] ZHENG S, HAO Y, LU D, et al. Joint entity and relation extraction based on a hybrid neural network[J]. Neurocomputing, 2017, 257: 59-66.

[84] ZHENG S, WANG F, BAO H, et al. Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2017: 1227-1236.

[85] REN F L, ZHANG L H, YIN S J, et al. A Novel Global Feature-Oriented Relational Triple Extraction Model based on Table Filling[C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2021: 2646-2656.

[86] WANG Y, SUN C, WU Y, et al. UniRE: A Unified Label Space for Entity Relation Extraction[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. Association for Computational Linguistics, 2021: 220-231.

[87] ZENG X, ZENG D, HE S, et al. Extracting Relational Facts by an End-to-End Neural Model with Copy Mechanism[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2018: 506-514.

[88] LI X, YIN F, SUN Z, et al. Entity-Relation Extraction as Multi-Turn Question Answering[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2019: 1340-1350.

[89] SUI D B, CHEN Y B, LIU K. Joint Entity and Relation Extraction with Set Prediction Networks[OL]. (2020-11) [2021-03-24]. https://arxiv.org/abs/2011.01675.

[90] BEKOULIS G, DELEU J, DEMEESTER T, et al. Joint entity recognition and relation extraction as a multi-head selection problem[J]. Expert Systems with Applications, 2018, 114: 34-45.

[91] WEI Z, SU J, WANG Y, et al. A Novel Cascade Binary Tagging Framework for Relational Triple Extraction[C]// Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2020: 1476-1488.

[92] LI X M, LUO X T, DONG C H, et al. TDEER: An Efficient Translating Decoding Schema for Joint Extraction of Entities and Relations[C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Dominican Republic. Association for Computational Linguistics, 2021: 8055-8064.

[93] YE D M, LIN Y K, LI P, et al. Packed Levitated Marker for Entity and Relation Extraction[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2022: 4904-4917.

[94] FU T J, LI P H, MA W Y. GraphRel: Modeling Text as Relational Graphs for Joint Entity and Relation Extraction[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2019: 1409-1418.

[95] YAN Z, ZHANG C, FU J, et al. A Partition Filter Network for Joint Entity and Relation Extraction[C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Dominican Republic. Association for Computational Linguistics, 2021: 185-197.

[96] TONG M H, WANG S, XU B, et al. Learning from Miscellaneous Other-Class Words for few shot named entity recognition [C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, 2021: 6236-6247.

[97] WANG Y R, SHINDO H, MATSUMOTO Y J, et al. Nested Named Entity Recognition via Explicitly Excluding the Influence of the Best Path[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, 2021: 3547-3557.

[98] LI F, LIN Z C, ZHANG M S. A Span-Based Model for Joint Overlapped and Discontinuous Named Entity Recognition[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, 2021: 4814-4828.

[99] ZHU E W, LI J P. Boundary Smoothing for Named Entity Recognition[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Pa-

pers). Association for Computational Linguistics, 2022: 7096-7108.

[100]YE D, LIN Y, LI P, et al. A Simple but Effective Pluggable Entity Lookup Table for Pre-trained Language Models[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Association for Computational Linguistics, 2022: 523-529.

[101]YANG A, WANG Q, LIU J, et al. Enhancing Pre-Trained Language Representations with Rich Knowledge[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2019: 2346-2357.

[102]LU L, KONG F. Conversation-oriented entity relationship extraction with knowledge [C]// Proceedings of the 20th Chinese National Conference on Computational Linguistics. Chinese Information Processing Society of China, 2021.

[103]ZHANG R, HRISTOVSKI D, SCHUTTE D, et al. Drug repurposing for COVID-19 via knowledge graph completion[J]. Journal of Biomedical Informatics, 2021, 115: 103696.

[104]YEW K C, BING L D, PORIA S. RelationPrompt: Leveraging Prompts to Generate Synthetic Data for Zero-Shot Relation Triplet Extraction[C]// Findings of the Association for Computational Linguistics: ACL 2022. Association for Computational Linguistics, 2022: 45-57.

[105]LU Y, BARTOLO M, MOORE A, et al. Fantastically Ordered Prompts and Where to Find Them: Overcoming Few-Shot Prompt Order Sensitivity[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2021: 8086-8098.

[106]SI C L, ZHANG Z Z, CHEN Y F, et al. READIN A Chinese Multi-Task Benchmark with realistic and diverse noises[OL]. (2023-02) [2023-02-20]. https://arxiv.org/abs/2302.07324.

[107]WANG X, GUI M, JIANG Y, et al. ITA: Image-Text Alignments for Multi-Modal Named Entity Recognition[C]// Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies Seattle. Association for Computational Linguistics, 2022: 3176-3189.

[108]LI Y Q, LI W J, NIE L Q. MMCoQA: Conversational Question Answering over Text, Tables, and Images[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2022: 4220-4231.

[109]HAN J L, CHENG B, LU W. Exploring Task Difficulty for Few-Shot Relation Extraction[C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2021: 2605-2616.

**HOU Jing**, born in 1999, postgraduate. Her main research interests include information extraction and standard digitization.

**HAN Pengwu**, born in 1985, Ph. D, senior engineer. His main research interests include knowledge map, space Internet of things and computer control.

(责任编辑:喻藜)