



开放科学  
(资源服务)  
标识码  
(OSID)

# 基于神经网络的嵌套命名实体关系抽取模型

陈浩 王兴芬

北京信息科技大学 北京 100192

**摘要:** [目的/意义] 网络安全形势日益严峻,从威胁情报中抽取网络安全实体及其关系,构建结构化威胁情报信息,对于网络安全防护来说尤为重要。在过去的工作中嵌套实体的关系抽取一直是难点,嵌套实体和关系重叠不能被有效识别导致准确率低,为威胁情报信息抽取研究带来巨大的挑战。[方法/过程] 基于上述情况,针对嵌套实体关系抽取过程中存在的问题,本文基于图注意力网络提出了一种新型知识抽取模型 SRG。采用 Bert-Bi-LSTM 作为共享编码层,与边界检测模块所得的多跨度实体共同进行跨度表示后,利用图注意力网络提取特征进行关系抽取,可有效改善实体边界信息与类别信息的检测效果。[结果/结论] 在公共数据集上进行了实验验证,验证结果表明,其在解决实体嵌套与关系重叠的问题上有显著的效果。

**关键词:** 深度学习; 嵌套实体识别; 实体关系抽取; 威胁情报;

**中图分类号:** G35; TP391

## Nested Named Entity Recognition Based on Neural Network

CHEN Hao WANG Xingfen

Beijing Information Science & Technology University, Beijing 100192, China

**Abstract:** [Objective/Significance] The network security situation is becoming much more serious. It is particularly important for network security protection to extract network security entities and their relationships from threat intelligence and build structured threat intelligence information. In the past work, the relationship extraction of nested entities are always difficult. The overlapping of nested entities and relationships cannot be effectively identified, resulting in low accuracy, which brings great challenges to the

**基金项目** 国家重点研发计划课题“大宗商品电子商务市场的交易风险智能分析与预警技术”(2019YFB1405003);北京市教委科技一般项目“数字化转型背景下面向企业协同治理的群体智能决策机制与方法研究”(KM202111232017)。

**作者简介** 陈浩(1993-),硕士研究生,研究方向为网络空间安全,威胁情报,文本挖掘,知识图谱,E-mail: chenhaom1993@hotmail.com;王兴芬(1968-),博士,教授,研究方向为电子商务信用评价与风险预警、大数据分析 with 智能决策。

**引用格式** 陈浩,王兴芬.基于神经网络的嵌套命名实体关系抽取模型[J].情报工程,2021,7(6):67-75.

research of threat intelligence information extraction. [Methods/Process]According to such situation and aiming at the problems existing in the process of nested entity relationship extraction, this paper proposes a new knowledge extraction model SRG based on graph attention network. Bert-Bi-LSTM is used as the shared coding layer to represent the span together with the multi span entities obtained by the boundary detection module, and the graph attention network is used to extract features for relationship extraction, which can effectively improve the detection effect of entity boundary information and category information. The experimental verification is carried out on the public data set. [Results/Conclusions]The verification results show that, it has a significant effect on solving the problem of entity nesting and relationship overlap.

**Keywords:** Deep learning; nested entity recognition; entity relationship extraction; threat intelligence

## 引言

近年来,在网络攻击数量迅猛增长、复杂度日渐增加情况下,为了更有效的促进安全信息共享,保证信息系统的安全,Gartner于2013年提出了威胁情报的概念:“威胁情报是基于证据的知识,包括场景、机制、指标、含义和可操作的建议。这些知识是关于现存的、或者即将出现的、针对资产的威胁或危险,可为主体响应相关威胁或危险提供决策信息”。威胁情报所包含的网络安全威胁信息,对于由企业或组织实施主动网络防御有重要作用。所以,从开源的互联网文章或报告中有效抽取威胁情报信息,并将其转换为标准化、结构化的形式,对网络安全研究有着非常重要的意义与实际应用价值。

对威胁情报进行信息抽取主要包括命名实体识别和关系抽取两个步骤,实体识别是从大量的原始信息中找出特定的某种实体信息,如人名、地点等文本信息中的实体具有丰富的语义,同时也直接影响对整个上下文的理解。关系抽取是从非结构化文本中抽取实体和关系的文本处理技术,它是自然语言理解的基础,在智能问答、信息检索等领域有重要应用。因此如何能够从大量的信息库中找出具有参考价值的实体数据和关系非常关键。常采用的方法有

管道模型和联合模型。管道模型通常需要两个子任务单独工作顺序执行,也就是可以先进行实体识别,再做关系抽取工作,这种方案流程灵活便捷,且子任务之间可以独立工作单独实现各自的工作内容,但这种模型也会存在问题,如子任务之间传递识别信息互相影响可能会产生较大数据误差等。相比之下,联合模型能够有效的改善上述管道模型中存在的问题。在命名实体识别过程中,传统的实体识别的相关模型无法解决实体嵌套的情况,因为在嵌套实体识别中,同一个字可能会有两个及以上不同的序列标签。例如:“北京信息科技大学”中的“北”同时属于B-Location,也属于B-Organization,而“京”也拥有I-Location与I-Organization两个标签,所以实体嵌套问题是研究难点。关系重叠问题在实体关系抽取中会影响抽取的性能,所以解决各种关系重叠类型的实体关系抽取、提高抽取性能是对威胁情报进行信息抽取研究的重点。基于此,本文基于边界检测与跨度表示进行实体识别,利用图注意力网络模型对实体关系联合抽取进行处理,不仅能够快速准确的检索到实体边界与类别信息,还能够将非主要的实体信息与主要的关系信息相互结合,有效的缓解关系重叠对嵌套实体关系抽取模型的影响。

综上, 本文利用联合模型提出一种准确率更高的基于图注意力网络的实体关系抽取模型 SRG 应用于威胁情报的信息抽取, 创新点如下: 命名实体识别方面, 针对实体识别过程中存在实体嵌套导致准确率下降问题, 采用边界预测的方法得到更为准确的不同跨度实体, 再与隐藏层实体共同进行跨度表示, 得到多尺度跨度的实体特征, 可有效识别嵌套实体, 提升实体边界与类型识别准确率; 关系抽取方面, 为了提升模型在抽取时的准确性和泛化性, 利用图注意力神经网络对实体间关系进行特征提取, 有效捕捉实体内部的依赖关系, 减少边界错误数量, 提升模型泛化能力以及抽取的准确率。

## 1 相关工作

### 1.1 嵌套实体识别

目前基于深度学习的实体识别已经能够在最初大量的语料数据中主动学习目标的表示, 还可以在非线性映射函数里对较为复杂的特征进行学习和掌握。大多数研究将命名实体识别任务视为传统的序列标注问题<sup>[1]</sup>, 并提出了大量基于 Bi-LSTM-CRF 架构的命名实体识别模型<sup>[2]</sup>, 该模型主要是从输入词里捕捉对词语在语料中的表达, 并利用上下文编码器获取新的语义单词的表达, 再在标签解码器中, 形成对应的输出标签。

Finkel 等<sup>[3]</sup>指出大量的命名实体都是嵌套的。对于嵌套命名实体识别的早期研究工作是基于规则的<sup>[4]</sup>。但是这种基于人工规则的缺陷也是十分明显的, 针对不同数据集,

都需要手工的构建相关规则来进行处理。现在对于嵌套命名实体识别的处理方式更多的是基于跨度的方法<sup>[5]</sup>。在基于跨度的方法中, 嵌套实体属于不同的子序列从而易被检测。Xu 等<sup>[6]</sup>尝试通过将每个子序列编码表示为固定大小来直接对句子的所有子序列进行分类。但是由于在边界检测的过程中缺乏有效的监督信息, 与基于序列标签的方法相比, 基于跨度的方法在确定实体的边界时性能较差。

为了缓解上述问题, 本文采用了跨度分类框架, 通过将边界检测学习过程纳入到模型中, 帮助模型学习实体的边界信息, 可以有效减少候选实体的数量。通过生成高质量的实体, 减少整体模型的实体数量, 降低模型的时间复杂度。

### 1.2 实体关系抽取

关系抽取简单来说指的是从一段指定文本中抽出其中的三元组。即按照语言学的习惯, 形式化的将一个句子表示为(主语, 谓语, 宾语), 也就是(subject, predicate, object)。因此, 三元组有时候也被称为“SPO 三元组”。因为语言表达存在复杂性和多样性, 特别是在汉语语言关系的表达中, 常常存在形式简单却语义复杂的表述, 因此, 想要从这样的语言环境中抽取出正确的关系是研究难点, 且非常依赖于系统对于语篇的分析和理解程度。

关系抽取在自然语言处理领域中挖掘文本中事实知识的任务<sup>[7]</sup>。对于有标注的实体识别, 该任务就可以简化为一个简单的分类任务, 即关系分类。Zeng 等<sup>[8]</sup>利用 CNN 或 RNN 来解决关系分类任务。然而, 这些方法忽略了从句子中提

取实体关系时，不能真正的确定实体与关系的对应事实。当句子中没有任何标注实体时，研究人员提出了几种联合提取实体和关系的方法。现有对多关系抽取任务的研究可以分为四种范式：（1）基于流水线的方法，如 Zelenko 等<sup>[9]</sup>先识别实体，再进行关系分类；（2）基于表格填充的方法，如 Miwa 等<sup>[10]</sup>使用共享参数表示实体和关系，但单独提取实体和关系；（3）基于标签的方法，如 Zheng<sup>[11]</sup>和 Dai 等<sup>[12]</sup>，将此任务视为序列标注问题，需要设计复杂的标注模式来进行实体关系抽取；（4）基于 Seq2seq 的方法，如 Zeng<sup>[13]</sup>、Zhang<sup>[14]</sup>、Liu 等<sup>[15]</sup>应用 seq2seq 模型直接生成关系三元组。

## 2 SRG 模型

### 2.1 模型框架

本文提出一种基于图注意力网络的实体关系抽取联合模型 SRG（Span Representation

GAT），用来提升威胁情报在信息抽取过程中嵌套实体识别及关系重叠提取的准确率。该方法主要将实体关系联合抽取划分为三个子任务，即边界预测、实体跨度表示和实体关系抽取。SRG 由四个模块构成：编码（Bert-Bi-LSTM）模块、边界预测（Boundary prediction）模块、跨度表示（Span representation）模块和基于图注意力网络（Graph Attention Networks）的嵌套实体关系抽取模块。首先通过 Bert-Bi-LSTM 编码器对语料库进行特征提取，再联合执行边界预测和跨度表示模块进行多尺度实体跨度预测，利用联合模型对不同跨度实体边界与类别信息进行识别。以实体为节点，实体关系为边构造图结构作为图注意力网络输入，通过 Attention 机制的引入，使得图模型在进行训练的过程中，能够同时获取边信息与相邻节点特征进行，丰富了实体跨度与关系间的交互，提升模型抽取效果。

本文中提出的联合实体关系抽取 SRG 模型总体框架设计如图 1 所示。

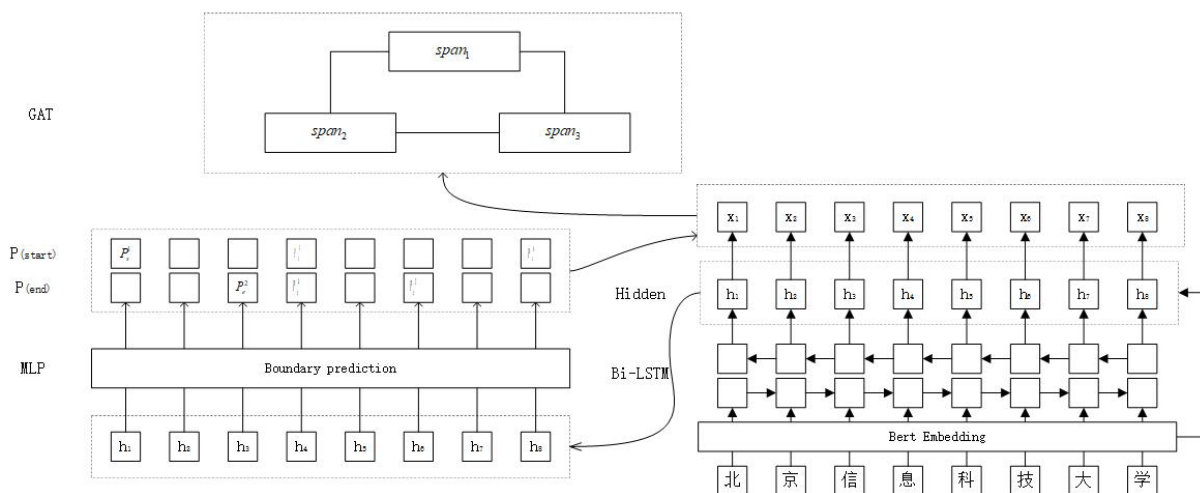


图 1 SRG 模型框架图



## 2.2 编码器模块

本文中编码器采用双向长短时记忆网络。这种编码器的好处在于能够在同一时间段内多方获取目标单词在语境中的上下文含义，这种编码器能够把词嵌入层检索到的词向量放入到整体的语境中，从而得到层次更为多元的数据。

例如，给定一个句子  $S=[w_1, w_2, w_u, \dots, w_n]$ ，句子中的  $w_u$  为第  $u$  个单词， $n$  是句子长度。每一个字输入嵌入  $x_u$  均由字符级嵌入、词级

嵌入这两部分组成。其中字符级嵌入向量通过 Bert 模型得到，词级嵌入则通过将字符级嵌入表达输入到 Bi-LSTM 模型通过计算得到的最终隐藏状态生成。 $h_u$  为通过将  $w_u$  输入到 Bi-LSTM 模型获得的隐藏状态向量，它是由现在所在位置的前向和后向长短时记忆网络两个隐层状态向量拼接而成，从两个方向深入挖掘输入序列之间存在的依赖关系。最终将两个层次的嵌入表达拼接在一起共同构成输入语料的最终编码向量  $x_u=[w_u;h_u]$ 。

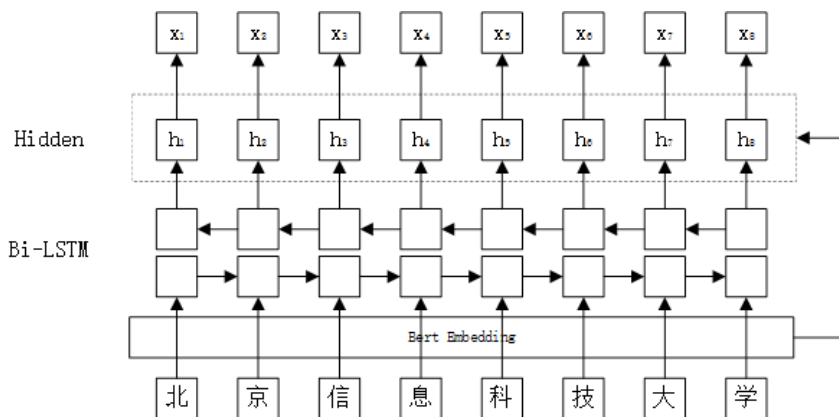


图2 编码器模块框架图

## 2.3 边界检测模块

有别于传统采用 BIO 和 BILOU 序列标注法进行的命名实体识别，在 BIO 中 B 代表 Begin 表示一个命名实体的开头字符，I 代表 Inside 表示一个命名实体中的字符，O 代表 Outside 表示不属于命名实体中的字符；而 BILOU 则更进一步，其中 L 代表 Last 表示一个命名实体的最后一个字符，U 代表一个单自命名实体。边界检测的目的在于对一个实体的首尾字进行预判，并选用两个标记分类器来明确首尾边界的位置信息。这样对于命名实体识别不再停留在

token-level 的分类问题上，避免了每一个字符级仅能对应一种标签类型，不能够有效解决嵌套实体识别的问题。

如公式 (1) 和 (2) 所示， $P_s^u$  和  $P_e^u$  分别表示将单字隐藏状态  $h_u$  输入到多层感知 MLP 分类器中，然后使用 softmax 层计算该单字  $w_u$  是为一个命名实体首尾单字的概率。

$$P_s^u = \text{softmax}(MLP_{start}(h_u)) \quad (1)$$

$$P_e^u = \text{softmax}(MLP_{end}(h_u)) \quad (2)$$

同时定义首尾单字的两个交叉熵之和作为该边界预测任务的损失函数。其中  $y_s^u$  和  $y_e^u$  分别代表了该单字是否是命名实体的首尾单字。

$$L_b = -\left(\sum_{u=1}^n \left[ y_s^u \log P_s^u + (1 - y_s^u) \log(1 - P_s^u) \right] + \sum_{e=1}^n \left[ y_e^u \log P_e^u + (1 - y_e^u) \log(1 - P_e^u) \right] \right) \quad (3)$$

## 2.4 跨度表示模块

通过边界检测明确了单字为命名实体的首尾概率后,对首尾单字进行有效的组合,得到跨度下的实体的向量表达。跨度表示旨在获得与跨度对应的语义标签,并根据跨度边界对单词的表示进行概括。跨度表示的方法通常为:平均池化、最大池化和基于自注意力机制的池化。这里我们将对应实体的跨度表示定义为  $e_k = [x_s; \hat{x}_{ij}; x_e]$ , 其中当给定一个跨度窗口为  $s=[i, j]$  时, 对应  $\hat{x}_{ij}$  即为该跨度窗口下的编码向量的平均池化  $\hat{x}_{ij} = \frac{1}{j-i+1} \sum_{k=i}^j x_k$ ,  $x_s$  为首位字符的编码向量,  $x_e$  为末尾字符的编码向量。得到实体的跨度表示后, 将  $e_k$  送入多层感知器 MLP 中。然后使用 softmax 层计算该实体所对应的实体类型。

## 2.5 图注意力网络

本文全面的考虑到实体与关系之间的相互作用, 采用基于图注意力网络扩展的一种对带有关系特征边的图进行编码的关系抽取方法。本部分就结合数学方法来对图注意力网络的基本工作原理过程进行概述。已知图注意力网络可以直接作用于图结构  $G=(V, E)$  的神经网络, 如公式(4)所示, 其中  $V$  和  $E$  分别是节点和边的集合。对于顶点  $e_j$ , 逐个计算它的邻居们和它自己之间的相似系数。

$$e_j = a([Wh_i || Wh_j]), j \in N_{(i)} \quad (4)$$

公式(4)中,  $e_i$  为目标节点,  $N_{(i)}$  是节点  $e_i$

的邻域;  $W$  是可学习的共享权重参数, 相当于对顶点的特征进行了一个线性映射, 进行了数据增维;  $[\cdot || \cdot]$  是对于顶点  $e_{ij}$  的变换后的特征进行了拼接。

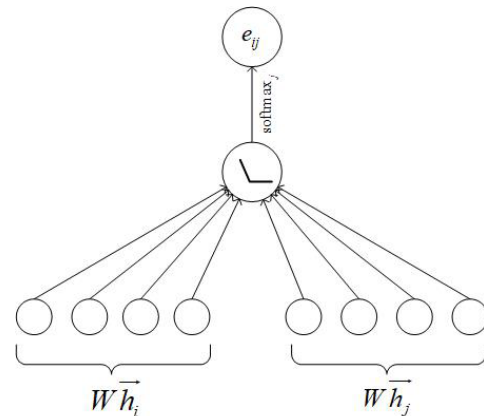


图3 注意力机制的计算过程

注意力机制计算过程如图3所示。具体计算过程如下所示:

$$h'_i = \sigma \left( \sum_{j \in N_{(i)}} \alpha_{ij} Wh_j \right) \quad (5)$$

上述公式(5)中  $h'_i$  就是图注意力网络输出的对每一个顶点  $e_i$  融合了相关邻域信息的新特征向量, 为了进一步增加在实体关系对匹配时候的健壮性, 引入多头注意力机制进一步增强注意力的优势, 具体公式(6)如下所示:

$$h'_i(K) = \parallel_{k=1}^K \sigma \left( \sum_{j \in N_{(i)}} \alpha_{ij}^k Wh_j^k \right) \quad (6)$$

本文采用关系图注意力网络优势在于其可以将关系特征集成到节点中, 这样相对于传统方法来说, 更有利于捕获长距离依赖, 并且在计算过程中, 图注意力网络是逐结点进行运算的, 每一次运算都需要循环遍历图上的所有顶点来完成, 因此可以更好的将顶点之间的特征的相关性融入到模型中。

### 3 实验结果与分析

#### 3.1 数据集

本文采用 2005 年自动内容提取 (Auto-matic Content Extraction, ACE) 公开数据集, 开展实验, 数据集综合统计情况如表 1 所示。

其中, 公共数据集 ACE05 选用通过 LDC 公开发布的标准数据进行测试, 这套数据由于保有非常多的嵌套实体, 其实体识别的难度较

大也较繁杂。因此, 多数的关系抽取工作也更常选用其进行分析, 具有较高的代表性。此外, 还可以用于事件抽取方面的操作, 能够支持对实体识别、关系、事件、值和事件表达式等多方面的研究, 普适性也极强。但在本文中, 主要是通过选用 ACE05 数据集进行实体识别操作, 其含有地理命名实体 (GPE)、组织 (ORG)、人 (PER)、地点 (LOC)、车辆 (VEH)、设施 (FAC) 和武器 (WEA) 这七类实体, 适合本文的识别研究的现实需要和理论专业要求。

表 1 ACE2005 语料库统计

粒度	统计项	AEC2005		
		训练集	验证集	测试集
文章	总数	370	43	51
语句	总数	7336	958	1047
	含嵌套命名实体	2686 (37)	341 (36)	330 (32)
实体	总数	24687	3217	3027
	含嵌套命名实体	9937 (40)	1192 (37)	1184 (39)
	长度>6	1343 (5)	148 (5)	160 (6)
	平均长度	19.24	18.93	17.19
	最大嵌套层数	6	3	4

#### 3.2 参数设计

模型的基本参数进行设置如下: 选择 256 维 Bert 作为固定字嵌入; Bi-LSTM 的隐藏状态为 200 维; 随机失活设置为默认值 0.5; Adam 优化器的学习率设置为 0.001。

#### 3.3 结果分析

为进一步保证 SRG 模型结果的有效性, 本文利用基线模型进行辅助论证, 具体模型如下所述: 利用 Multi-trunQA<sup>[16]</sup> 为实体关系的抽取找到了一种多轮对话的全新视角, 可以基于上

下文语境的整体环境对实体关系进行检索和抽取。利用 Span-level<sup>[17]</sup> 模型可以将全部跨度进行建模操作以达到联合实体关系检索的效果。利用 GraphRel 模型<sup>[18]</sup> 能够对端到端的关系建立加权关系图, 此外, 通过图可以分析出实体与关系的影响与逻辑。

##### 3.3.1 实验一: 嵌套实体关系抽取

结合表 1 和表 3 中的数据可以看出, 虽然数据集中存在着许多嵌套实体, 其中中重叠实体就有 37.5% 的占比, 但本文中设计的 SRG 模型的在该数据集上进行实体识别仍取得了较高

的准确率, 达到 87.1% 的性能, 在 Span-level 方面有 1.1% 的进步。由此也可以论证, 本文涉及的相关模型的实体嵌套解决能力非常高效; 在关系抽取工作的解决方面, 尽管数据集内含有的关系重叠占比很低, 但该模型的 F1 值依然能够得到 62.1% 的良好成绩。Multi-trunQA 能够对远距离实体和关系的相关数据进行编码, 通过多轮问答形式解决实体关系的抽取工作, 相比于其他模型来说, 这种模型因为多次的编码而增加了模型的操作难度, 通过对比上表数据可以发现, 这种情况更适合基于跨度的表示的模型。

表 3 数据集 ACE05 上的结果

System	Entity	Relation
Multi-trunQA	0.848	0.602
Span-level	0.860	0.628
SRG	0.871	0.621

Span-level 模型为准确得到跨度表示, 需要先在输入端操作跨度排列步骤, 然后对关系作 softmax 分类, 这种操作极易产生实体冗余。相比之下, 本文设计的模型通过对关系加以分类管理, 能够极大的降低分析错误的产生。此外, 尽管 DYGIE 模型的动态相关跨度表示比较全面, 却在实体与关系的分析上较为不足, 相比之下, 本文设计的模型由于具备边界检测能够极大的强化跨度识别的精准, 此外, 实体关系图还能将跨度与关系间的交互作用得以强化, 能够快速弥补前模型在实体嵌套方面的不足。

### 3.3.2 实验二: 消融实验

为了详尽的说明各个组件的功能, 本文还基于 ACE05 和 NYT 两组数据集完成了消融实

验研究, 实验主要从边界预测和图神经网络等方面进行分析, 结果如表 5 所示。

表 5 消融实验

Model	ACE05		NYT	
	Entity	Relation	Entity	Relation
All	0.871	0.621	0.843	0.701
-Boundary pre	0.832	0.587	0.811	0.638
-GAT	0.809	0.605	0.798	0.652

对 ACE05 和 NYT 两组数据集来看, 去除边界预测模块导致的降低有 3.9% 和 3.2%, 还对模型在关系抽取工作产生作用。此外, 普通模型在区分重叠时会受到假阳性的影响, 而降低了模型的准确度, 但边界预测能够快速的对重叠进行有效区分, 能够很好的保障跨度表示的准确性。通过上述分析可以发现, 各组件都对实体识别与关系抽取的有效进行具有重要意义, 应注意不同子任务间的协调配合。

## 4 总结与展望

针对嵌套命名实体识别的问题本文所提出了一种基于图注意力神经网络的嵌套命名实体识别模型来对威胁情报进行信息抽取, 这种模型结构主要采用 Bert-Bi-LSTM 网络作为共享编码层, 对系统输入的句子进行编码, 然后提取其特征。为有效解决实体嵌套问题, 模型中还加入了边界检测模块, 该模块的作用是为模型提供高质量的实体跨度。将边界检测后实体与编码器输出隐藏层共同进行跨度表示后输入图注意力网络构建实体间关系图完成信息抽取。在公共数据集上进行了实验验证, 验证结果表



明, 其在解决实体嵌套的问题上可以起到十分关键的作用。

虽然本文提出的基于图注意力神经网络的嵌套命名实体识别模型, 相较于传统管道模型解决了一些现存缺陷, 也在各项实验中取得了不错的性能, 能够通过参数共享的方式使得实体关系抽取子任务之间产生了依赖关系, 但在获取高质量候选实体时, 仍然会产生冗余信息, 后续研究将对模型不足之处继续完善。

## 参 考 文 献

- [1] 杨沛安, 武杨, 苏莉娅, 等. 网络空间威胁情报共享技术综述 [J]. 计算机科学, 2018, 45(6):9-18+26.
- [2] Huang Z, Xu W, Yu K. Bidirectional LSTM-CRF models for sequence tagging[J]. arXiv preprint arXiv:1508.01991, 2015.
- [3] Finkel J R, Manning C D. Nested named entity recognition[C]. Proceedings of the 2009 conference on empirical methods in natural language processing. 2009:141-150.
- [4] 李建华. 网络空间威胁情报感知、共享与分析技术综述 [J]. 网络与信息安全学报, 2016(2):16-29.
- [5] 杨泽明, 李强, 刘俊荣, 等. 面向攻击溯源的威胁情报共享利用研究 [J]. 信息安全研究, 2015(1):31-36.
- [6] Xu M, Jiang H, Watcharawittayakul S. A local detection approach for named entity recognition and mention detection[C]. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017:1237-1247.
- [7] Bach N, Badaskar S. A review of relation extraction[J]. Literature review for Language and Statistics II, 2007(2):1-15.
- [8] Zeng D, Liu K, Lai S, et al. Relation classification via convolutional deep neural network[C]. Proceedings of COLING 2014, the 25<sup>th</sup> International Conference on Computational Linguistics: Technical Papers. 2014:2335-2344.
- [9] Zelenko D, Aone C, Richardella A. Kernel methods for relation extraction[J]. Journal of machine learning research, 2003, 3(2):1083-1106.
- [10] Miwa M, Bansal M. End-to-end relation extraction using lstms on sequences and tree structures[J]. arXiv preprint arXiv:1601.00770, 2016.
- [11] Zheng S, Wang F, Bao H, et al. Joint extraction of entities and relations based on a novel tagging scheme[J]. arXiv preprint arXiv:1706.05075, 2017.
- [12] Dai D, Xiao X, Lyu Y, et al. Joint extraction of entities and overlapping relations using position-attentive sequence labeling[C]. Proceedings of the AAAI conference on artificial intelligence. 2019, 33(1):6300-6308.
- [13] Zeng X G, Zeng D J, He S Z, et al. Extracting relational facts by an end-to-end neural model with copy mechanism. Proceedings of the 56<sup>th</sup> Annual Meeting of the Association for Computational Linguistics. 2018:506-514.
- [14] Zhang M, Wang Q, Fu G. End-to-end neural opinion extraction with a transition-based model[J]. Information Systems, 2019(80):56-63.
- [15] Zhang R H, Liu Q, Fan A X, et al. Minimize exposure bias of Seq2Seq models in joint entity and relation extraction[J]. arXiv preprint arXiv:2009.07503, 2020.
- [16] Luan Y, He L, Ostendorf M, et al. Multi-Task Identification of Entities, Relations, and Coreference for Scientific Knowledge Graph Construction[J]. arXiv preprint arXiv: 1808.09602-2018.
- [17] Dixit K, Al-Onaizan Y. Span-level model for relation extraction[C]. Proceedings of the 57<sup>th</sup> Annual Meeting of the Association for Computational Linguistics. 2019:5308-5314.
- [18] 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 57<sup>th</sup> Annual Meeting of the Association for Computational Linguistics. 2019:1409-1418.