基于残差空洞卷积神经网络的网络安全实体识别方法

谢博[1,2]，申国伟[1,2]，郭春[1,2]，周燕[1,2]，于淼[3]

（1. 贵州大学计算机科学与技术学院，贵州 贵阳 550025；
2. 贵州省公共大数据重点实验室，贵州 贵阳 550025；
3. 中国科学院信息工程研究所，北京 100093）

摘　要：近年来，网络安全威胁日益增多，数据驱动的安全智能分析成为网络安全领域研究的热点。特别是以知识图谱为代表的人工智能技术可为多源异构威胁情报数据中的复杂网络攻击检测和未知网络攻击检测提供支撑。网络安全实体识别是威胁情报知识图谱构建的基础。开放网络文本数据中的安全实体构成非常复杂，导致传统的深度学习方法难以准确识别。在 BERT（pre-training of deep bidirectional transformers）预训练语言模型的基础上，提出一种基于残差空洞卷积神经网络和条件随机场的网络安全实体识别模型 BERT-RDCNN-CRF。通过 BERT 模型训练字符级特征向量表示，结合残差卷积与空洞神经网络模型有效提取安全实体的重要特征，最后通过 CRF 获得每一个字符的 BIO 标注。在所构建的大规模网络安全实体标注数据集上的实验表明，所提方法取得了比 LSTM-CRF 模型、BiLSTM-CRF 模型和传统的实体识别模型更好的效果。

关键词：网络安全；实体识别；残差连接；空洞卷积神经网络；BERT 预训练模型

中图分类号：TP391

文献标识码：A

doi: 10.11959/j.issn.2096−109x.2020009

Cyber security entity recognition method based on residual dilation convolution neural network

XIE Bo[1,2], SHEN Guowei[1,2], GUO Chun[1,2], ZHOU Yan[1,2], YU Miao[3]

1. College of Computer Science and Technology, Guizhou University, Guiyang 550025, China
2. Guizhou Provincial Key Laboratory of Public Big Data, Guiyang 550025, China
3. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China

Abstract: *In recent years, cybersecurity threats have increased, and data-driven security intelligence analysis has become a hot research topic in the field of cybersecurity.* In particular, the artificial intelligence technology repre-

收稿日期：2019−10−10；**修回日期**：2020−01−07

通信作者：申国伟，gwshen@gzu.edu.cn

基金项目：国家自然科学基金（61802081）；贵州省自然科学基金（20161052，20167428，20171051）；贵州省科技重大专项计划基金（20183001）

Foundation Items: The National Natural Science Foundation of China (61802081), The Natural Science Foundation of Guizhou Province, China (20161052, 20167428, 20171051), The Major Scientific and Technological Special Project of Guizhou Province, China (20183001)

论文引用格式：谢博，申国伟，郭春，等. 基于残差空洞卷积神经网络的网络安全实体识别方法[J]. 网络与信息安全学报, 2020, 6(5): 126-138.

XIE B, SHEN G W, GUO C, et al. Cyber security entity recognition method based on residual dilation convolution neural network[J]. Chinese Journal of Network and Information Security, 2020, 6(5): 126-138.

sented by the knowledge graph can provide support for complex cyberattack detection and unknown cyberattack detection in multi-source heterogeneous threat intelligence data. Cybersecurity entity recognition is the basis for the construction of threat intelligence knowledge graphs. The composition of security entities in open network text data is very complex, which makes traditional deep learning methods difficult to identify accurately. Based on the pre-training language model of BERT (pre-training of deep bidirectional transformers), a cybersecurity entity recognition model BERT-RDCNN-CRF based on residual dilation convolutional neural network and conditional random field was proposed. The BERT model was used to train the character-level feature vector representation. Combining the residual convolution and the dilation neural network model to effectively extract the important features of the security entity, and finally obtain the BIO annotation of each character through CRF. Experiments on the large-scale cybersecurity entity annotation dataset constructed show that the proposed method achieves better results than the LSTM-CRF model, the BiLSTM-CRF model and the traditional entity recognition model.

Key words: cybersecurity, entity recognition, residual connection, dilation convolution neural network, BERT pre-train model

1　引言

在网络空间安全态势日趋复杂的形势下，威胁情报驱动的网络安全防御成为业界关注的重点[1]。从海量碎片化的网络数据中挖掘威胁情报，采用知识图谱模型进行组织，支撑攻击路径预测、攻击溯源等，可实现海量数据驱动的威胁情报智能分析。

网络安全实体识别是威胁情报知识图谱构建任务中非常重要的一个基础任务[2]，其目标是从网络安全领域的文本数据中提取出安全实体的语义类，如攻击组织、企业、漏洞、软件等。

网络安全实体识别属于一种特定领域的命名实体识别。命名实体识别是自然语言处理中的一项重要研究内容，主要有基于规则的命名实体识别方法、基于机器学习的命名实体识别方法和基于深度学习的命名实体识别方法[3]。由于深度学习方法能够自适应地提取文本的特征信息，不依赖大量的特征工程和额外的语言学知识，因此，深度学习被广泛地应用在命名实体识别任务中[4]。

Georgescu 等[5]通过基于命名实体识别的解决方案来增强和检测物联网系统中可能存在的漏洞。王通等[2]使用深度置信网络对威胁情报知识图谱实体识别子任务中的安全实体进行有效识别。2003 年，Hammerton[6]利用长短记忆模型（LSTM）抽取句子的序列信息，并通过条件随机场（CRF，conditional random fields）对命名实体的标签进行分类。之后，很多命名实体识别方法是在 LSTM-CRF 网络架构下融入各种句子隐含的特征信息。Collobert 等[7]在 2011 年使用窗口化方法的神经网络和基于句子的卷积神经网络方法对命名实体识别进行了深入探索，在命名实体识别任务中取得了不错的效果。随后，Santos 等[8]使用字符级的特征向量作为卷积神经网络的输入，来增强 CNN-CRF 模型。Chiu 等[9]在 Hammerton 和 Collobert 等的工作基础上使用双向 LSTM 并融合卷积神经网络来获取词的字符特征，与 Collobert 的模型相比，利用双向 LSTM 模型打破固定窗口大小的限制。由于传统的卷积神经网络在提取大的上下文特征信息时会丢掉一些信息，2017 年，Strubell 等[10]提出利用空洞卷积神经网络进行命名实体识别，实验表明该网络能够弥补传统卷积神经网络的不足且进一步提高网络的训练速度。

He[11]、Liu[12]、Li[13]等通过实验表明了基于字符的命名实体识别方法一般比基于词的命名实体识别方法好。因此，秦娅等[14]针对传统的命名实体识别方法难以识别网络安全实体，提出一种基于特征模板的字符级 CNN-BiLSTM-CRF 网络安全实体识别模型。除了基于字的命名实体识别，还有基于词、融合字和词特征信息的命名实体识别方法。Xu 等[15]将字和词的特征信息联合训练进行融合，其中，Zhang 等[16]提出的 Lattice LSTM 中文命名实体识别网络结构取得了较好的实体识别效果，该方法将传统的 LSTM 单元改为网格 LSTM，在字模型的基础之上利用词典，从而得

到词与词序的特征信息，降低了分词带来的错误传播问题。

注意力机制（attention mechanism）[17]被广泛应用于自然语言处理任务。Bahdanau 等[18]将注意力机制和循环神经网络（RNN，recurrent neural network）结合起来解决机器翻译任务，使注意力机制成功融入自然语言处理领域。Yin 等[19]提出一种基于注意力机制的卷积神经网络，并成功地应用于句子对建模任务。Wang 等[20]在关系抽取任务中证明了注意力机制和卷积神经网络结合的有效性。

与传统领域的实体识别相比，网络安全领域的实体识别发展比较缓慢，且仍存在以下几方面的挑战。

1) 深度学习依赖大规模的标签数据，然而网络安全领域中缺乏大规模高质量的网络安全实体标注数据。

2) 网络安全实体识别比较复杂，包含大量的多种混合安全实体，如“SQL 注入”“80 端口”。

3) 全文非一致问题。即在一篇文章、段落或者长句子中相同的实体被赋予不同的类别标签。此外，存在实体缩写情况，即第一次提到实体时给出全称，在其后提到该实体时，仅仅给出该实体的缩写。

针对上述挑战，本文提出一种残差空洞卷积神经网络的安全实体识别方法。本文的主要贡献如下。

1) 构建并开放了一个网络安全实体识别语料库，标注了 6 类典型的网络安全实体。

2) 提出一种基于 BERT[21]语言模型的残差空洞卷积神经网络命名实体识别方法，与基于注意力机制的 BiLSTM-CRF 网络结构相比，本文的模型可以接收句子的平行化输入，降低了模型的训练时间。

3) 将残差连接与空洞卷积神经网络模型相结合，构建网络安全实体识别模型，无须使用诸如词性、依存句法等额外的特征信息。此外，以字符级特征向量作为模型的输入，降低了分词错误导致的误识别率。实验表明，本文提出的模型取得了比 BiLSTM-CRF 等模型好的安全实体识别效果。

2 任务定义

网络安全实体识别任务可以看作特定领域中的命名实体识别，其相当于自然语言处理中的序列标注问题。由于网络安全文本数据中的词条存在歧义和安全实体构成比较复杂，为了防止分词带来的错误传播问题，在字符层面进行序列标注是解决网络安全实体识别的关键问题。

本文以网络安全文本中的句子为基本单元，针对其中的任意一个句子 $s=[x]_1^N=[x_1,\cdots x_i\cdots, x_N]$，$x_i$ 是句子 s 中的第 i 个字符。在 BIO 标注方法[14]的指导下，识别句子 s 中的安全实体相当于给出标注序列 $L_s=[l]_1^N$。例如，句子 s 为“腾讯安全防御威胁情报中心检测到一款通过驱动人生系列软件升级通道传播的木马突然爆发”，其对应的序列标注 L_s 是“B-ORG I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG O O O O O O O B-SW I-SW I-SW I-SW O O O O O O O O O O O B-RT I-RT O O O O”。

3 基于残差空洞卷积神经网络的网络安全实体识别模型

针对网络安全实体识别任务，本文在 BERT 预训练模型的基础上，提出基于残差空洞卷积神经网络和 CRF 相结合的网络安全实体识别模型 BERT-RDCNN-CRF，如图 1 所示。

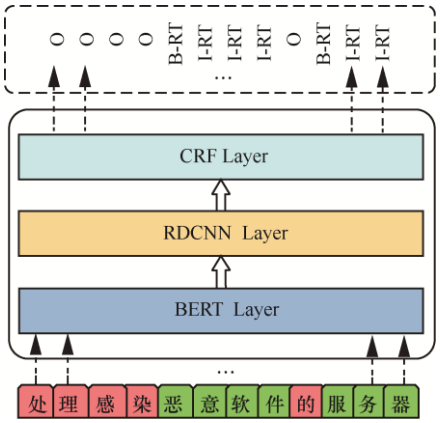

图 1 基于 BERT-RDCNN-CRF 的网络安全实体识别模型
Figure 1 Network security entity recognition model based on BERT-RDCNN-CRF

该模型在 BERT 预训练模型的基础上，采用 RDCNN 提取文本中的特征，为实体识别奠定基

础，最后通过 CRF 层处理给出标注序列。

3.1　BERT 预训练语言模型

BERT 模型能从大量的无标签数据中学习词条或者字的前向和后向这种双向表示，并且在学习词条或者字的上下文表示过程中通过微调来解决词或字的歧义问题。BERT 模型的框架如图 2 所示，其主要由输入层、双向的 Transformer 编码层和输出层组成。输入层接收输入句子的字嵌入（token embedding）、段嵌入（segment embedding）和位置嵌入（position embedding）拼接而成的一个特征矩阵。Transformer 编码层主要提取输入层特征矩阵中重要的特征信息。输出层通过一个前馈神经网络输出每一个字的嵌入表示。

本文以字为单位并将其作为 BERT 模型的输入，对于长度为 n 的句子 s，其字嵌入表示为

$$e_{1:n} = e_1 \oplus e_2 \oplus \cdots \oplus e_n \tag{1}$$

其中，$\oplus$ 为拼接操作符。

由于本文的输入为文本句子，所以对于长度为 n 的句子 s，引入段嵌入作为句子对的区分界限。段嵌入全部初始化为 0，如式(2)所示。

$$\hat{s} = [0, 0, \cdots, 0] \tag{2}$$

其中，$\hat{s} \in \mathbb{R}^n$。

在网络安全实体识别任务中，字的位置特征是识别的关键特征。因此，BERT 模型加入了位置嵌入，如式(3)~式(5)所示。

$$\mathrm{PE}_{(\mathrm{pos},2i)} = \sin \frac{\mathrm{pos}}{10\,000^{\frac{2i}{d_{\mathrm{model}}}}} \tag{3}$$

$$\mathrm{PE}_{(\mathrm{pos},2i+1)} = \cos \frac{\mathrm{pos}}{10\,000^{\frac{2i}{d_{\mathrm{model}}}}} \tag{4}$$

$$\mathrm{POS} = \mathrm{PE}_{(\mathrm{pos},2i)} \oplus \mathrm{PE}_{(\mathrm{pos},2i+1)} \tag{5}$$

其中，pos 为位置，i 为维度，d_{model} 为模型的输出维度，$\oplus$ 为拼接操作符。

最后，将字嵌入、段嵌入和位置嵌入拼接起来作为 BERT 模型的输入，如式(6)所示。

$$X = e_{1:n} \oplus \hat{s} \oplus \mathrm{POS} \tag{6}$$

双向 Transformer 编码层通过“多头”注意力机制（multi-head attention mechanism）扩展了模型专注于不同位置的能力，增大注意力单元的“表示子空间”。“多头”注意力机制的基础是自注意力机制，自注意力机制主要计算句子中的每个字对于这个句子中所有字的相互关系，即将与该字相关联的其他字的特征信息编码进该字的嵌入表示中。自注意力机制的计算如式(7)所示。

$$\mathrm{Attention}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \mathrm{soft\,max}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{7}$$

其中，$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 分别是自注意力机制中的查询向量（query vector）、键向量（key vector）和值向量（value vector），计算方法如下。

$$\boldsymbol{Q} = \boldsymbol{X}\boldsymbol{W}^{Q} \tag{8}$$

$$\boldsymbol{K} = \boldsymbol{X}\boldsymbol{W}^{K} \tag{9}$$

$$\boldsymbol{V} = \boldsymbol{X}\boldsymbol{W}^{V} \tag{10}$$

其中，$\boldsymbol{W}^{Q}$、$\boldsymbol{W}^{K}$、$\boldsymbol{W}^{V}$ 为权重矩阵，在模型开始

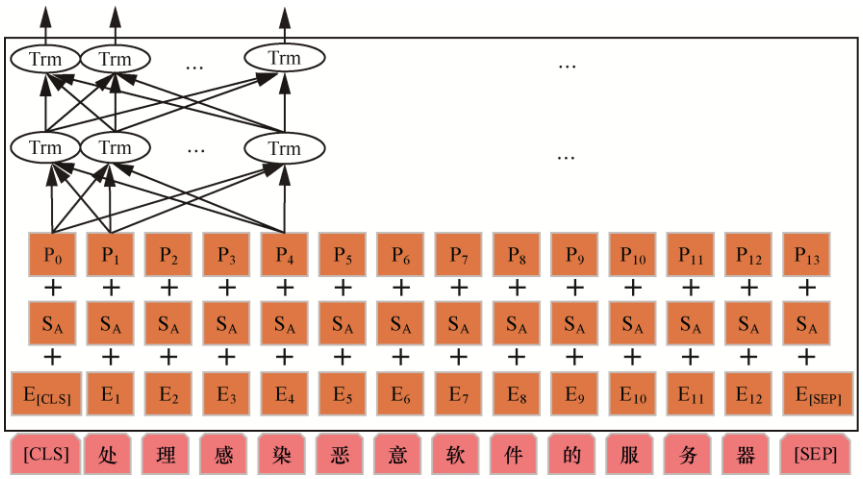

图 2　微调 BERT 模型在单句标注任务图示

Figure 2　Illustrations of fine-tuning BERT on single sentence tagging tasks

训练时随机初始化。

由此，实现“多头”注意力机制的计算如式(11)所示。

$$\begin{aligned}\text{MultiHead}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = (&\text{head}_1 \oplus \text{head}_2 \oplus \cdots \\ &\oplus \text{head}_i \oplus \cdots \oplus \text{head}_h)\boldsymbol{W}^o\end{aligned} \tag{11}$$

其中，$\oplus$ 为拼接操作符，$\boldsymbol{W}^o$ 为形状变换矩阵，是一个需要学习的参数，$\text{head}_i = \text{Attention}(\boldsymbol{Q}\boldsymbol{W}_i^Q, \boldsymbol{K}\boldsymbol{W}_i^K, \boldsymbol{V}\boldsymbol{W}_i^V)$。

最后，通过一个全连接层，输出每个字的嵌入表示，其计算如式(12)所示。

$$\text{FFN} = \max(0, \boldsymbol{X}\boldsymbol{W}_1 + b_1)\boldsymbol{W}_2 + b_2 \tag{12}$$

其中，$\boldsymbol{W}_1$、$\boldsymbol{W}_2$ 为权重矩阵，b_1、b_2 为偏置，FFN 为 BERT 模型的输出结果。

3.2 RDCNN -CRF 命名实体识别模型

本节在 BERT 模型的基础上，提出网络安全实体识别模型 RDCNN-CRF，如图 3 所示。RDCNN-CRF 主要由输入层、空洞卷积层和 CRF 层组成。输入层接收输入句子的特征矩阵；空洞卷积层利用卷积核对输入的基本单位进行卷积操作提取特征。RDCNN 的输入层将 BERT 模型的输出构建为一个特征矩阵传入模型中；CRF 层通过提取到的特征信息输出字的命名实体标签分类结果。

在空洞卷积神经网络的卷积层中，给定长度为 h 的卷积核，可以把句子分为 $x_{i-\delta(h-1):i+\delta(h-1)}$，然后对每一个分量进行卷积操作，通过式(13)得到卷积特征图。

$$\hat{C} = (\hat{c}_1, \hat{c}_2, \cdots \hat{c}_i, \cdots, \hat{c}_n) \tag{13}$$

其中，$\hat{c}_i$ 是对分量 $x_{i-\delta(h-1):i+\delta(h-1)}$ 进行卷积操作之后提取得到的信息，如式(14)所示。

$$\hat{c}_i = \text{relu}(\boldsymbol{W} \cdot x_{i-\delta(h-1):i+\delta(h-1)} + b) \tag{14}$$

其中，$\boldsymbol{W}$ 为卷积核权重，b 为偏置，$\delta > 1$ 为空洞卷积神经网络的膨胀系数。若 $\delta = 1$，此时的空洞卷积神经网络和传统的卷积神经网络等价。

由于对空洞卷积神经网络的卷积层进行简单的线性堆叠会使网络在训练过程中产生过拟合和退化问题，本文引用文献[22]提出的残差连接来防止退化问题，使用批正则化（BN，batch normalization）防止过拟合问题。残差连接的残差块计算如式(15)所示。

$$o = x + F(x) \tag{15}$$

其中，x 为输入，$F()$ 表示残差函数。

最后，通过 CRF 层得到字的实体标签分类结果。CRF 能够考虑相邻字的实体标签之间的关系，这符合字的实体标签关系之间并不独立的特点，且充分利用了字实体标签的上下文信息。而通过直接对残差空洞卷积神经网络的输出获取其对应字的实体标签的结果取决于数据的性质和质量。CRF 层的具体算法如下。

首先，定义一个状态转移矩阵 $\boldsymbol{A}$，这里的 $\boldsymbol{A}_{i,j}$

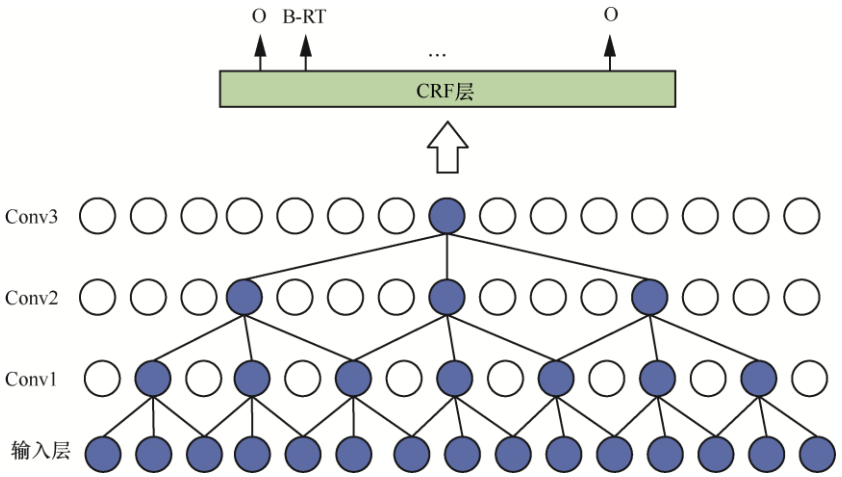

图 3 RDCNN-CRF 模型框架

Figure 3 Illustrations of RDCNN-CRF model framework

表示标签 i 转移到标签 j 的得分，该得分会随着模型的训练而更新。另外，定义分值矩阵 $f_\theta([x]_1^N)$ 为空洞卷积层的输出分值，其中，$[f_\theta]_{i,t}$ 是第 t 个字、第 i 个标签的 RDCNN 的输出分值，θ 是 RDCNN 的参数。针对句子 $s=[x]_1^N$，N 为句子的长度。定义 $\tilde{\theta}=\theta\cup\{\boldsymbol{A}_{i,j}\forall i,j\}$ 为整个 RDCNN-CRF 模型需要学习的参数。这样，一个句子在给定标签序列 $[h]_1^N$ 的总得分计算为

$$S\left([x]_1^N,[h]_1^N,\tilde{\theta}\right)=\sum_{t=1}^{N}\left(\boldsymbol{A}_{[h]_{t-1},[h]_t}+[f_\theta]_{[h]_t,t}\right) \quad (16)$$

然后，在句子 $[x]_1^N$ 上真实标签序列 $[l]_1^N$ 的条件概率可使用 softmax 函数计算得到。

$$p\left([l]_1^N\,|\,[x]_1^N,\tilde{\theta}\right)=\frac{\mathrm{e}^{S\left([x]_1^N,[l]_1^N,\tilde{\theta}\right)}}{\sum_{j}\mathrm{e}^{S\left([x]_1^N,[y]_1^N,\tilde{\theta}\right)}} \quad (17)$$

其中，$[y]_1^N$ 为可能的标签序列。最后，通过最大似然估计来训练模型的参数，计算为

$$\mathrm{lb}p\left([l]_1^N\,|\,[x]_1^N,\tilde{\theta}\right)=S\left([x]_1^N,[h]_1^N,\tilde{\theta}\right)-\mathrm{lb}\sum_{\forall[y]_1^N}\mathrm{e}^{S([x]_1^N,[y]_1^N,\tilde{\theta})} \quad (18)$$

模型训练结束后，本文采用维特比算法来找到最佳标签序列。

3.3　模型训练过程

本文基于 Tensorflow 深度学习框架实现所提出的网络安全实体识别方法，具体训练过程如下。

本文提出的安全实体模型在训练的过程中先对 BERT、RDCNN 和 CRF 模型进行初始化，然后 BERT 对字进行编码，随后 RDCNN 进行解码提取出句子局部特征信息，接着 CRF 模型计算出字的安全实体标签；最后将错误向前传播更新各个模型的参数。

算法 1　BERT-RDCNN-CRF

输入

训练数据集 $T=(X,Y)$，其中，$X,Y\in\mathbb{R}^{|V|\times n}$，$|V|$ 是数据集的大小，n 是句子最大长度；

输出

实体标签序列；

1) 初始化 BERT，RDCNN 和 CRF;

2) for i=0,1,2, do

3) for batch do

4) BERT 和 RDCNN 模型前向传播;

5) CRF 前向传播和后向传播，计算出序列的全局似然概率;

6) 对 RDCNN 和 BERT 模型进行后向传播;

7) 更新 BERT，RDCNN，CRF 模型的参数;

8) end for

9) end for

4　实验与结果分析

本节将所提出的方法在构建的数据集上进行实验。在本文的实验中，字向量均采用 Google 预训练好的 BERT 中文字向量。BERT 模型采用 Fine-tuning 策略，即模型参数采用 Google 预训练好的参数进行初始化，并可以在网络训练过程中自适应地调整 BERT 模型的参数。

4.1　实验数据

实验数据主要来自乌云漏洞数据库、Freebuf 网站、国家漏洞库等主流网络安全平台的公开数据。网络安全实体主要有 6 种类型，分别是人名（PER，person）、地名（LOC，location）、组织名（ORG，organization）、软件名（SW，software）、网络安全相关术语（RT，relevant term）和漏洞编号（VUL_ID，vulnerability ID）。网络安全实体数据均采用 BIO 命名实体标注策略。数据集的具体统计信息如表 1 所示。实验中将标注好的数据集划分为训练集、验证集和测试集，分别占总数据集规模的 70%、10%和 20%。详见 Github 官网。

表 1　数据集的统计信息
Table 1　Statistic of datasets

实体类别	训练集	验证集	测试集	总计
PER	8 737	1 256	2 570	12 563
LOC	16 454	2 372	4 890	23 716
ORG	12 755	1 813	3 566	18 134
SW	4 649	673	1 407	6 729
RT	56 190	8 045	16 212	80 447
VUL_ID	4 649	673	1 407	6 729
总计	103 434	14 832	30 052	148 318

本文采用精确率（P，Precision）、召回率（R，Recall）、F1 值（F1-measure）和准确率（Accuracy）作为评价标准。

4.2 对比实验

为了验证本文提出的网络安全实体识别方法 BERT-RDCNN-CRF 的有效性，对 12 种模型进行对比实验。其中，前 6 组实验的词向量和字符向量是基于 word2vec 语言模型训练的，后 6 组基于 BERT 预训练语言模型。实验代码可到 Github 官网下载。

1) CRF：文献[23]提出使用 CRF 对序列数据进行标注。

2) LSTM：文献[24]提出使用 LSTM 进行命名实体识别的模型。

3) LSTM-CRF：文献[25]提出结合 CRF 的 LSTM 命名实体识别模型。

4) BiLSTM-CRF：文献[26]提出考虑词上下文的双向 LSTM 结合 CRF 进行命名实体识别模型。

5) CNN-BiLSTM-CRF：文献[27]提出使用 CNN 学习词条的字符级特征信息，并将其和词拼接在一起作为 BiLSTM 的输入进行命名实体识别的模型。

6) FT-CNN-BiLSTM-CRF：文献[14]提出结合特征模板的网络安全实体识别模型。

7) BERT-CRF：在文献[21]中模型的基础上结合 CRF 进行命名实体识别的模型。

8) BERT-LSTM-CRF：在文献[25]的基础上将语言模型使用 BERT 代替。

9) BERT-BiLSTM-CRF：在文献[26]的基础上将语言模型使用 BERT 代替。

10) BERT-GRU-CRF：在 BERT 的基础上，使用普通 GRU 与 CRF 结合的命名实体识别模型。

11) BERT-BiGRU-CRF：在 BERT 的基础上，使用普通双向 GRU 与 CRF 结合的命名实体识别模型。

12) BERT-RDCNN-CRF：本文提出的识别方法弥补了 CNN 提取特征信息有限，并使用残差连接防止模型在训练的过程中出现过拟合的情况。

4.3 实验模型参数

在实验中，使用多种窗口卷积核对 BERT 的输出矩阵进行卷积操作。卷积核函数为 rectified linear units，激活函数为 Leaky ReLU。模型训练过程中采用 Zeiler[28]提出的 Adadelta 更新规则，其他参数见表 2。

表 2 实验参数设置

Table 2 Hyper parameters of experiment

参数名	参数描述	参数值
h	窗口大小	3
n	特征图数量	128
p	Dropout 率	0.5
b	批处理大小	64
δ	卷积空洞率	2
n_r	残差块数量	4

4.4 整体对比实验及分析

本文将 12 组实验在网络安全实体识别数据集上进行实验，分析网络安全实体识别。表 3 给出了 12 组实验在网络安全实体识别数据集上总体的实体识别准确率、精确率、召回率和 F1 值。

从表 3 结果可以看出，整体上本文提出的方法在网络安全实体数据集上取得了不错的实体标签分类效果。其中，BERT-CRF、BERT-LSTM-CRF、BERT- BiLSTM-CRF、BERT-GRU 和 BERT-BiGRU-CRF 模型在网络安全实体数据集中的命名实体识别准确率比传统基于特征的 CRF 命名实体识别模型、LSTM 模型、BiLSTM 模型、CNN-BiLSTM 模型和基于特征模板的 FT-CNN-BiLSTM 模型好。

对比文献[26]提出的 CNN-BiLSTM-CRF 模型和文献[13]提出的 FT-CNN-BiLSTM-CRF 模型可以看出，考虑了字特征信息的 BERT-RDCNN-CRF 模型在网络安全实体数据集上的实体识别准确率相比考虑了词特征信息和字特征信息的模型的准确率高。这是因为网络安全的实体是由字母、数字和中文构成的，在分词的过程中会产生大量的分词错误，这种错误会随着模型的训练往后传播，影响模型最后对实体标签的分类效果。

从表 3 的结果还可以看出，模型 BERT-CRF 的 F1 值相比不使用 BERT 模型的实体识

别模型没有提高，但是加了能够提取文本的句法和表层特征的序列模型后，其 F1 值有较大的提高。说明安全实体识别模型在含有丰富的语义特征基础上利用文本的句法和表层特征能提高实体识别性能。

表 3　不同模型的网络安全实体识别对比结果
Table 3　Comparison of cyber security entity recognition results of different models

模型	准确率	精确率	召回率	F1 值
CRF	0.915 0	0.842 6	0.733 4	0.784 2
LSTM	0.923 6	0.837 5	0.806 2	0.821 6
LSTM-CRF	0.929 5	0.861 7	0.820 7	0.840 7
BiLSTM-CRF	0.928 3	0.847 0	0.851 8	0.849 4
CNN-BiLSTM-CRF	0.931 0	0.864 7	0.840 7	0.852 5
FT-CNN-BiLSTM-CRF	0.933 1	0.884 5	0.836 8	0.860 0
BERT-CRF	0.959 4	0.824 7	0.809 5	0.817 1
BERT-LSTM-CRF	0.975 3	0.883 0	**0.919 6**	**0.901 0**
BERT-BiLSTM-CRF	0.973 4	0.846 9	0.903 1	0.874 1
BERT-GRU-CRF	0.976 4	0.884 3	0.913 0	0.898 4
BERT-BiGRU-CRF	0.974 1	0.826 7	0.902 5	0.862 9
BERT-RDCNN-CRF	**0.976 8**	**0.887 2**	0.910 7	0.898 8

为了进一步说明基于 BERT 预训练模型的 LSTM、BiLSTM、GRU 和 BiGRU 模型与本文提出的方法在网络安全实体识别任务中的实体识别效果，本文进一步做了对比实验。从表 3 的结果可以看出，在准确率和精确率方面，本文提出的方法比其他基于 BERT 模型的网络安全实体识别模型好，说明本文提出的方法在网络安全实体识别任务中的有效性，并且在使用 BERT 模型的前提下，单向的 LSTM 模型和 GRU 模型比双向的 LSTM 模型和 GRU 模型在网络安全实体识别任务中效果更好。然而，从召回率和 F1 值的结果来看，BERT- LSTM-CRF 均取得了最好的结果，分别是 91.96%和 90.10%。与本文提出的方法在召回率（91.07%）和 F1 值（89.88%）上相比，分别提升了 0.89%和 0.22%，说明本文提出的方法在这两个评价指标上与能够提取字的序列特征的模型相比差距不大。

4.5　6 类安全实体识别对比实验及分析

为了进一步比较基于 BERT 模型的安全实体识别模型在不同安全实体上的识别效果，本文计算出这 6 种安全实体的精确率、召回率和 *F*1 值，其中精确率如图 4 所示。

从图 4 可以看出，BERT-CRF 模型在 SW 和 VUL_ID 两类安全实体上的实体识别效果非常差。所有实体识别模型在安全实体 SW 上的精确度比较低，最高的精确率才 50.26%，其中 BERT-LSTM-CRF 模型、BERT-GRU-CRF 模型和 BERT-RDCNN-CRF 模型的精确度相近，说明这些模型不擅长识别安全实体 SW，这是因为一方面该实体在安全实体数据集中的数量较少，另一方面这类实体通常由数字、字母和汉字组成，构成非常复杂，特征不好提取。

6 种安全实体识别模型在安全实体 LOC、ORG 和 PER 上的实体识别效果相差无几且精确率较高，说明 6 种模型能够对 LOC、ORG 和 PER 这种简单的实体特征信息进行充分提取。对于安全实体 RT，BERT-CRF 模型的精确率最低，本文提出的方法比 BERT-LSTM 模型的精确率高 0.52%，说明无论是使用改进的卷积网络还是使用能够存储句子序列信息的 LSTM，在安全实体 RT 上都能取得较好的精确率。在安全实体 VUL_ID 的结果中，BERT-GRU-CRF 模型取得了最高的精确率，比 BERT-LSTM-CRF 模型提升了 0.28%，比本文提出的方法提升了 4.13%，说明能够存储句子序列信息的 LSTM 模型和 GRU 模型在安全实体 VUL_ID 的精确率上更有优势。

从图 4 中还可以看出，使用双向 BiLSTM 模型和双向 BiGRU 模型的精确率比使用单向的 LSTM 模型和 GRU 模型的精确率低，这是因为模型的复杂度增加会产生过拟合问题，损失函数的损失值在模型训练过程中很难下降。为了更进一步比较 6 种模型的安全实体识别性能，对不同模型在不同安全实体上的召回率进行对比，如图 5 所示。

从图 5 的结果可以看出，6 种模型在 6 类安全实体上的召回率和图 4 的结果差不多。值得注意的是，BERT-LSTM-CRF 模型和 BERT-GRU-CRF 模型比其他模型取得了更好的召回率，这说明在召回率这一评价指标上，这两种模型能够召回更多的安全实体。

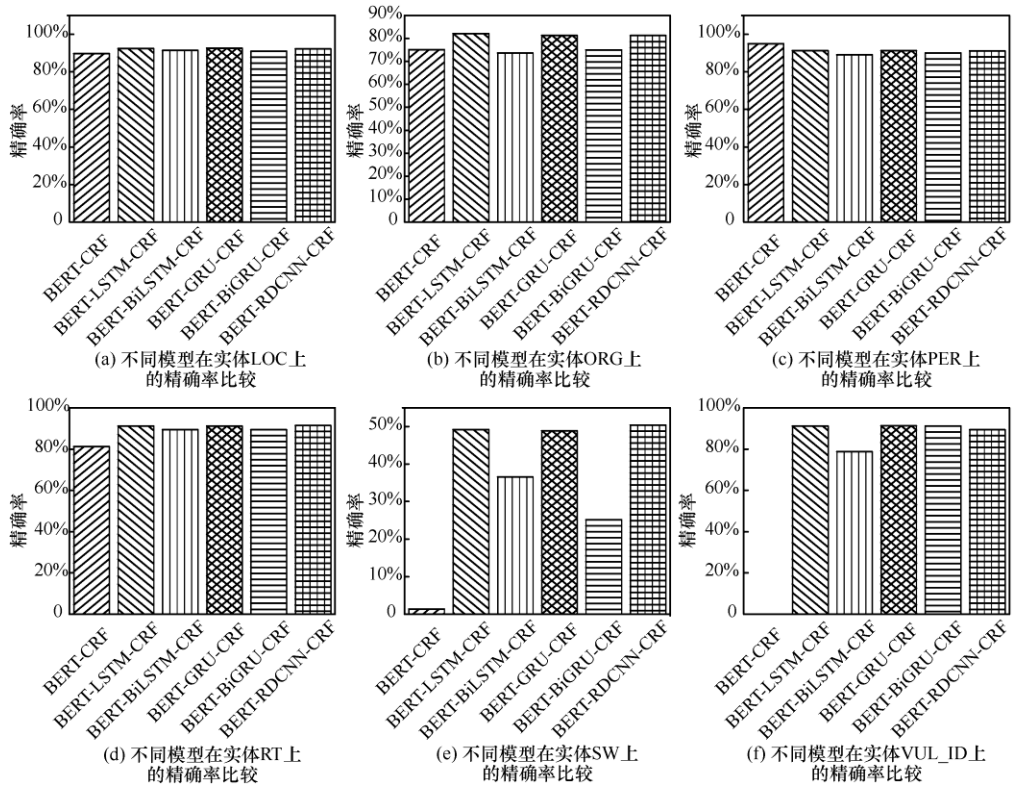

图 4　不同模型在不同安全实体上的精确率对比

Figure 4　Comparison of accuracy of different models on different security entities

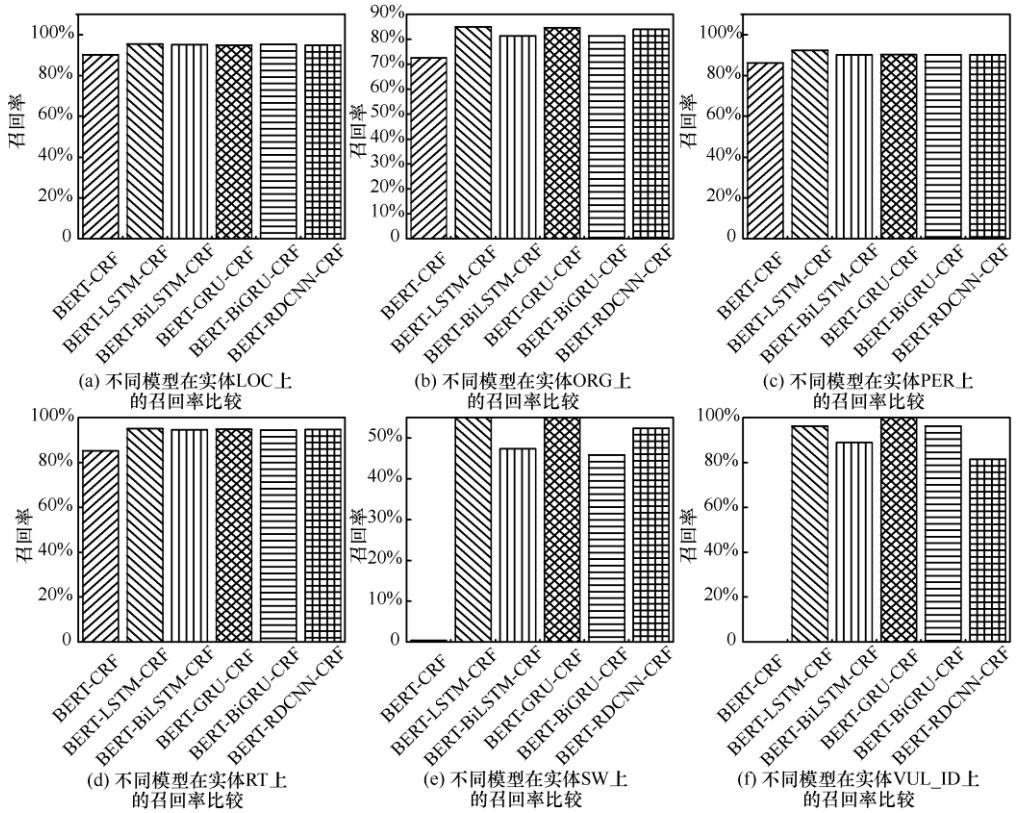

图 5　不同模型在不同安全实体上的召回率对比

Figure 5 Comparison of recall of different models on different security entities

为了平衡精确率和召回率，不同模型的 F1 值如图 6 所示。从图 6 中可以看出，BERT-CRF 模型、BERT-BiLSTM-CRF 模型和 BERT-BiGRU-CRF 模型对于复杂的安全实体 RT、SW 和 VUL_ID 的识别效果不是很好。对于安全实体 PER 的识别效果这 6 种模型相差无几，说明这 6 种模型都适合用来识别 PER 这样的安全实体。对于安全实体 LOC 的识别效果，BERT-RDCNN- CRF 模型、BERT-LSTM-CRF 模型、BERT-GRU- CRF 模型和 BERT-BiGRU-CRF 模型的 F1 值比 BERT-CRF 模型和 BERT- BiLSTM-CRF 的 F1 值有不同程度的提升。对于安全实体 ORG，BERT-RDCNN-CRF 模型和 BERT-LSTM-CRF 模型的 F1 值比其他模型的 F1 值有不同程度的提升。以上结果表明，本文提出的方法在各种安全实体识别任务中的鲁棒性和有效性。

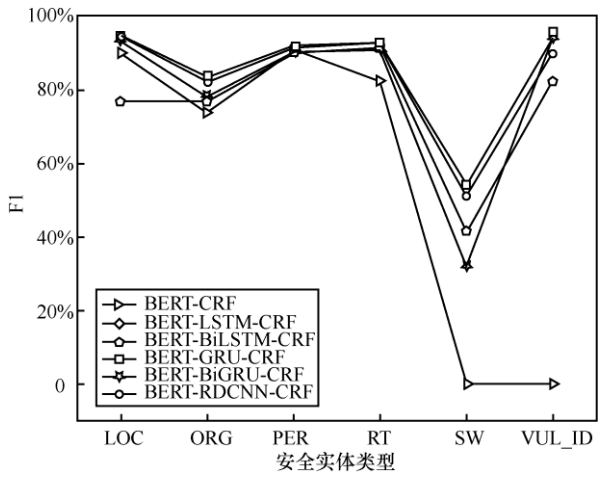

图 6　不同模型在不同安全实体上的 F1 值对比
Figure 6　Comparison of F1-measure of different models on different security entities

4.6　参数调整分析

本文提出的安全实体识别方法和其他传统的神经网络方法一样，通过最小化损失函数来进一步更新模型的参数。而损失值在模型训练过程中的变化情况表示该模型在训练过程中是否学习和是否稳定等。进一步分析模型在训练过程中损失值的变化情况如图 7 所示。

从图 7 可以看出，模型在整个训练过程中损失值是下降的，说明模型能够学习网络安全实体的相关特征信息。此外，由于本文提出的方法参数太多并且损失值整体上下降，进一步表明本文模型在训练过程中具有鲁棒性。从图 7 中还可以看出，损失值的下降曲线并不是平滑的，这主要与最小化损失函数的优化算法和学习率的设置有关，从整体上看依然能够说明本文所提方法的鲁棒性。

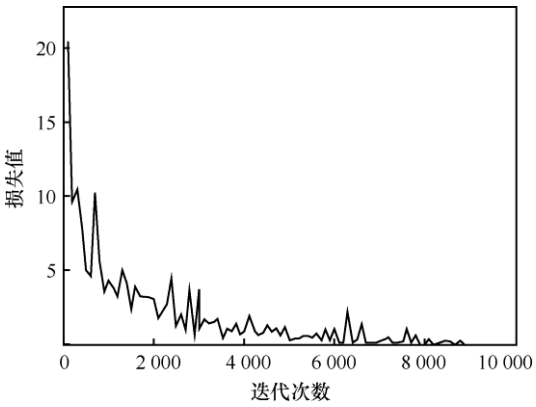

图 7　损失值曲线
Figure 7　Loss value curves during training on network security entity datasets

4.7　实例分析

为了进一步分析本文提出方法的实用性，从网络安全实体识别数据集中抽取一些典型句子的实体标签分类结果进行分析。经典句型分析如表 4 所示。

从表 4 中可以看出，对于句子 1 这种结构简单的实体识别，本文提出的方法能正确识别出实体的标签。在句子 2 中，本文提出的方法除了准确识别出安全实体 PER 和 RT，还识别出了安全数据测试集中没有标注的 ORG 实体类型。另外，在句子 3 和句子 7 中，本文提出的方法均能准确识别出安全数据测试集中没有标注的 ORG 和 SW 实体类型，说明本文提出的方法具有识别与训练集中相似实体的新实体的能力。对于句子 8 和句子 9 这种 VUL_ID 类型的英文字母和数字构成的实体，不管该实体是单独出现在句子中还是成对出现在句子中，本文提出的方法均能准确识别出该类型的实体。此外，诸如“暗云Ⅲ”这种汉字加罗马字符的安全实体，本文提出的方法均能准确将其识别出来。“IP 地址”这类英文字母加汉字构成的实体也能被准确识别出来，如句子 5 所示。

表 4　经典句型例子分析
Table 4　Analysis of typical sentences

序号	例句	BERT-RDCNN-Attention-CRF	真实标签
1	Wardle 在某篇博文中表示有了 RansomWhere，RansomWhere 在被发现和阻止前，理想情况下最多只能加密几个文件	RT：加密，文件 SW：RansomWhere PER：Wardle	RT：加密，文件 SW：RansomWhere PER：Wardle
2	早在 5 年多以前，SemiAccurate 专家 Charlie，Demerjian 在研究硬件后门时就知晓了该漏洞	ORG：SemiAccurate PER：Charlie，Demerjian RT：硬件，后面，漏洞	PER：Charlie，Demerjian RT：硬件，后面，漏洞
3	这种方案是 FIDO 联盟提出来的，苹果的芯片和 Android 手机芯片基本遵循这套方案	ORG：FIDO 联盟，苹果，Android RT：芯片	ORG：苹果 SW：Android RT：芯片
4	需要通过逆向工程来找到 Sign nature 的位置	RT：逆向工程	RT：逆向工程
5	将加密后的 IP 地址去除首尾的分隔符经过 base64 解密并于 036 异或之后得到真实的 IP 地址在本样本中	RT：IP 地址	RT：IP 地址
6	暗云Ⅲ木马专杀工具	SW：暗云Ⅲ RT：木马	SW：暗云Ⅲ RT：木马
7	当然，选择 OpenResty 也可以，如果选择 OpenResty 就不需要单独安装 lua 相关的组件	SW：OpenResty	无
8	编号 CVE-2017-0882 的漏洞可导致拥有向其他用户发送 issue 或 merge 请求权限的攻击者获取到该用户的信息	VUL_ID：CVE-2017-0882 RT：漏洞，用户，请求，权限，攻击者	VUL_ID：CVE-2017-0882 RT：漏洞，用户，请求，权限，攻击者
9	除了 CVE-2017-0199 之外，漏洞受欢迎程度排名第二位和第三位的分别是 CVE-2012-0158 和 CVE-2015-1641	VUL_ID：CVE-2017-0199，CVE-2012-0158，CVE-2015-1641	VUL_ID：CVE-2017-0199，CVE-2012 - 0158，CVE-2015-1641

然而，在句子 3 中，前一个实体“芯片”被准确识别出，而后一个实体“芯片”却没有被准确识别出来，Android 却被误识别为 ORG 实体类型，这可能是因为 Android 前面有两个 ORG 实体类型，算法在提取上下文信息的时候误认为 android 也是 ORG 类型的实体。

除了表 4 展示的这些典型例子，在实验过程中，对于较长的 ORG 这种实体类型算法，往往不能准确将其识别。换句话说，对于名称较长的实体，不管是构成简单的还是构成复杂的模型往往不能将其准确识别出来。所以，本文提出的方法更加适用于安全实体中名称不是很长和实体构成相对规律的安全实体。

5　结束语

针对开放网络文本数据中的安全实体构成非常复杂的问题，本文提出了一种基于残差空洞卷积神经网络的网络安全实体识别方法。使用 BERT 模型对字进行向量化表示，进一步结合空洞卷积神经网络和 CRF 准确地识别网络安全实体。在网络安全实体识别数据集上的实验结果表明，本文提出的基于残差空洞卷积神经网络的方法在准确率和精确率方面优于许多已有的实体识别方法。

通过分析经典句型可以看出，本文提出的方法在某些安全实体类型的识别上是存在不足的，仍然不能准确识别出一些安全实体。在下一步的工作中，考虑从网上爬取大量的网络安全数据，使用表示能力较强的语言模型训练出网络安全领域的字嵌入。并且，针对网络安全数据集中安全实体的数量不平衡问题，对相应的安全实体数量进行补充或者在模型训练的过程中采用解决数据

不平衡的训练技巧。此外，进一步提高较长的安全实体的识别准确率。

参考文献：

[1] SHU X, ARAUJO F, SCHALES D L, et al. Threat intelligence computing[C]//ACM SIGSAC Conference on Computer and Communications Security. 2018: 1883-1898.

[2] 王通, 艾中良, 张先国. 基于深度学习的威胁情报知识图谱构建技术[J]. 计算机与现代化, 2018(12): 21-26.
WANG T, AI Z L, ZHANG X G. Knowledge graph construction of threat intelligence based on deep learning[J]. Computer and Modernization, 2018(12): 21-26.

[3] 刘浏, 王东波. 命名实体识别研究综述[J]. 情报学报, 2018(3): 329-340.
LIU L. WANG D B. A review on named entity recognition[J]. Journal of the China Society for Scientific and Technical Information, 2018(3): 329-340.

[4] 张晓斌, 陈福才, 黄瑞阳. 基于 CNN 和双向 LSTM 融合的实体关系抽取[J]. 网络与信息安全学报, 2018, 4(9): 44-51.
ZHANG X B. CHEN F C. HUANG R Y. Relation extraction based on CNN and Bi-LSTM[J]. Chinese Journal of Network and Information Security, 2018, 4(9): 44-51.

[5] GEORGESCU T M, IANCU B, ZURINI M. Named-entity- recognition-based automated system for diagnosing cybersecurity situations in IoT networks[J]. Sensors, 2019, 19(15): 3380.

[6] HAMMERTON J. Named entity recognition with long short-term memory[C]//The 7th Conference on Natural Language Learning at Hltnaacl. 2003: 172-175.

[7] COLLOBERT R, WESTON J, BOTTOU L, et al. Natural language processing (almost) from scratch[J]. Journal of Machine Learning Research, 2011, 12(8): 2493-2537.

[8] SANTOS C N, GUIMARAES V. Boosting named entity recognition with neural character embeddings[J]. arXiv preprint arXiv:1505.05008, 2015.

[9] CHIU J P C, NICHOLS E. Named entity recognition with bidirectional LSTM-CNNs[J]. Transactions of the Association for Computational Linguistics, 2016, 4: 357-370.

[10] STRUBELL E, VERGA P, BELANGER D, et al. Fast and accurate entity recognition with iterated dilated convolutions[J]. arXiv preprint arXiv:1702.02098, 2017.

[11] HE J, WANG H. Chinese named entity recognition and word segmentation based on character[C]//The Sixth SIGHAN Workshop on Chinese Language Processing. 2008.

[12] LIU Z, ZHU C, ZHAO T. Chinese named entity recognition with a sequence labeling approach: based on characters, or based on words[C]//International Conference on Intelligent Computing. 2010: 634-640.

[13] LI H, HAGIWARA M, LI Q, et al. Comparison of the impact of word segmentation on name tagging for Chinese and Japanese[C]//LREC. 2014: 2532-2536.

[14] 秦娅, 申国伟, 赵文波, 等. 基于深度神经网络的网络安全实体识别方法[J]. 南京大学学报(自然科学), 2019, 55(1): 29-40.
QIN Y, SHEN G W, ZHAO W B, et al. Research on the method of network security entity recognition based on deep neural network[J]. Journal of Nanjing University (Natural Sciences), 2019, 55(1): 29-40.

[15] XU Y, WANG Y, LIU T, et al. Joint segmentation and named entity recognition using dual decomposition in Chinese discharge summaries[J]. Journal of the American Medical Informatics Association, 2013, 21(e1): e84-e92.

[16] ZHANG Y, YANG J. Chinese NER using lattice LSTM[J]. arXiv preprint arXiv:1805.02023, 2018.

[17] MNIH V, HEESS N, GRAVES A. Recurrent models of visual attention[C]//Advances in Neural Information Processing Systems 27 (NIPS 2014). 2014: 2204-2212.

[18] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[J]. arXiv preprint arXiv:1409.0473, 2014.

[19] YIN W, SCHÜ TZE H, XIANG B, et al. AbCNN: attention-based convolutional neural network for modeling sentence pairs[J]. Transactions of the Association for Computational Linguistics, 2016, 4: 259-272.

[20] WANG L, CAO Z, DE-MELO G, et al. Relation classification via multi-level attention CNNs[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. 2016: 1298-1307.

[21] DEVLIN J, CHANG M W, LEE K, et al. Bert: pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv:1810.04805, 2018.

[22] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.

[23] LAFFERTY J, MCCALLUM0-A, PEREIRA F C N. Conditional random fields: probabilistic models for segmenting and labeling sequence data[C]//Proceedings of the Eighteenth International Conference on Machine Learning. 2001: 282-289.

[24] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

[25] PENG N, DREDZE M. Named entity recognition for chinese social media with jointly trained embeddings[C]//2015 Conference on Empirical Methods in Natural Language Processing. 2015: 548-554.

[26] LAMPLE G, BALLESTEROS M, SUBRAMANIAN S, et al. Neural architectures for named entity recognition[J]. arXiv preprint

arXiv:1603.01360, 2016.

[27] MA X, HOVY E. End-to-end sequence labeling via bi-directional LSTM-CNNS-CRF[J]. arXiv preprint arXiv:1603.01354, 2016.

[28] ZEILER M D. ADADELTA: an adaptive learning rate method[J]. arXiv preprint arXiv:1212.5701, 2012.

[作者简介]

谢博（1996− ），男，云南昭通人，贵州大学硕士生，主要研究方向为网络安全、知识图谱、数据挖掘。

申国伟（1986− ），男，湖南邵东人，博士，贵州大学副教授，主要研究方向为大数据、网络与信息安全、数据挖掘。

郭春（1986− ），男，贵州贵阳人，博士，贵州大学副教授，主要研究方向为网络安全。

周燕（1980− ），女，贵州贵阳人，贵州大学讲师，主要研究方向为密码学与网络安全。

于淼（1987− ），男，黑龙江牡丹江人，博士，中国科学院信息工程研究所高级工程师，主要研究方向为网络与信息安全、数据挖掘。