

DOI: 10.19650/j.cnki.cjsi.J2514671

电力变压器缺陷记录嵌套实体识别方法*

缪希仁,曹喜龙,江 灏,陈 静,林蔚青
(福州大学电气工程与自动化学院 福州 350108)

摘要:随着近年来电网规模的不断扩大,海量非结构化的设备缺陷记录被采集与积累。这些数据中蕴涵着对设备状态评估与运维决策至关重要的实体信息。然而,此类数据中普遍存在的实体嵌套结构导致了实体边界模糊性与上下文语义复杂性的提升,对传统的命名实体识别方法构成了显著挑战。为实现对缺陷记录嵌套实体精准识别,以电力变压器为典型设备,提出了一种乱序预训练语言模型(PERT)、双向门控循环单元(BiGRU)与高效全局指针网络(EGP)相结合的电力变压器缺陷记录命名实体识别方法。该方法首先以 PERT 模型作为向量嵌入层进行动态语义编码,利用其乱序预训练特性深度捕捉文本的上下文依赖关系;随后引入 BiGRU 网络作为文本编码层,通过双向门控机制,更加全面地进行语义特征提取;最后采用 EGP 作为解码输出层,聚焦实体跨度与边界,实现对嵌套实体的精准抽取,避免了传统条件随机场(CRF)解码的标记冲突问题。实验结果表明,该实体识别方法能有效应对嵌套实体挑战,其综合 F1 分数达 96.01%,较传统嵌套识别方法双向变换器预训练模型-双仿射注意力(BERT-Biaffine)以及双向变换器预训练模型-机器阅读理解(BERT-MRC)分别高出 0.55% 和 0.70%。在全部 5 类实体标签识别中均取得了最高的 F1 分数。特别地,在嵌套现象最显著的缺陷设备与缺陷部位识别上,所提方法的 F1 分数分别达到 100% 与 94.74%,相较最优基准模型提升了 0.57% 与 0.13%。

关键词: 电力变压器缺陷记录;命名实体识别;乱序预训练语言模型;双向门控循环单元;高效全局指针网络;嵌套实体
中图分类号: TM721 TH86 **文献标识码:** A **国家标准学科分类代码:** 470.40

Nested entity recognition method for power transformer defect records

Miao Xiren, Cao Xilong, Jiang Hao, Chen Jing, Lin Weiqing

(College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China)

Abstract: With the continuous expansion of power grid scale in recent years, a vast amount of unstructured equipment defect records has been collected and accumulated. This data contains entity information crucial for equipment condition assessment and operational decision-making. However, the prevalent nested entity structures in such data lead to increased entity boundary ambiguity and contextual semantic complexity, posing significant challenges to traditional named entity recognition methods. To achieve accurate identification of nested entities in defect records, this paper takes power transformers as a typical example and proposes a named entity recognition method for power transformer defect records that combines pre-training bidirectional encoder representations from transformers with permuted language model (PERT), bidirectional gated recurrent unit (BiGRU), and efficient global pointer (EGP). This method first employs the PERT model as a vector embedding layer for dynamic semantic encoding, leveraging its permuted pre-training characteristics to deeply capture contextual dependencies. Subsequently, a BiGRU network is introduced as the text encoding layer to comprehensively extract semantic features through its bidirectional gating mechanism. Finally, EGP is adopted as the decoding output layer to focus on entity spans and boundaries, enabling precise extraction of nested entities while avoiding the label conflict issues associated with traditional CRF decoding. Experimental results show that this entity recognition method effectively addresses the challenges of nested entities, achieving a comprehensive F1 score of 96.01%, which is 0.55% and 0.70% higher than those of the bidirectional encoder representations from transformers-biaffine attention (BERT-Biaffine) and bidirectional encoder representations from transformers-machine reading comprehension (BERT-MRC), respectively. It attained the highest F1 scores across all five entity label categories. Specifically, for defect equipment and defect component recognition, where nesting is most prominent, the F1 scores of the proposed method reached 100% and 94.74%, representing improvements of 0.57% and 0.13% over the best baseline models.

Keywords: power transformer defect records; named entity recognition; pre-training BERT with permuted language model; bidirectional gated recurrent unit; efficient global pointer; nested entities

收稿日期:2025-11-11 Received Date: 2025-11-11

* 基金项目:福建省高校产学研合作项目(2023H6006)资助

0 引言

随着我国国民经济的快速发展和各行各业的用电需求的不断攀升,电力变压器的投运数量也在逐年增长。与此同时,电力变压器的缺陷次数和缺陷记录数量也呈现同步增长态势^[1]。电力变压器缺陷记录是运维检修人员对电力变压器进行缺陷诊断与处理过程中,以文本形式归档记录的设备缺陷特征、成因分析以及应对方案等关键信息,是电网公司核心数据资产的重要组成部分^[2-4]。然而,由于电力变压器缺陷记录呈现实体表述多样、非结构化和专业性等特点,导致人工查阅耗时较长且易因术语理解偏差导致信息提取遗漏,难以快速获取设备缺陷特征信息。为解决上述问题,科研人员提出引入命名实体识别(named entity recognition,NER)的方法,通过对电力变压器缺陷记录关键实体的精准识别,协助运维检修人员快速获取设备的缺陷位置与状态等关键信息。

NER作为自然语言处理(natural language processing,NLP)领域的基础任务之一,其目标是从非结构化文本中识别并分类出具有特定意义的专有名词,其识别精度直接决定下游任务的应用效能^[5-6]。然而,电力变压器缺陷记录中普遍存在的实体嵌套现象对仅适用于扁平实体识别的传统方法构成了挑战。传统模型因结构限制,难以处理同一文本对应多个类别标签的复杂情况,从而构成显著的技术瓶颈,因此亟需探索并设计出能够有效识别嵌套实体的新技术。

早期NER任务主要采用基于词典和规则的方法,通过人工设计规则模板与构建领域词典,利用模式匹配范式从文本中识别并分类实体^[7]。虽然该方法简单直观,但其依赖人工预定义模式的特点导致灵活性与扩展性不足,难以适应复杂多变环境下的实体识别需求。由此,基于统计机器学习技术应运而生。基于统计机器学习的典型算法^[8]包括隐马尔可夫模型(hidden Markov model,HMM)、最大熵模型(maximum entropy model, MEM)和条件随机场(conditional random field, CRF)等。上述典型机器学习算法虽提升了特征挖掘能力,但性能提升效果较为有限。

随着深度学习技术的不断发展,凭借其强大的特征学习与语义表征能力,有效增强了NER捕捉文本中复杂样式的能力。典型的深度学习算法^[9]包括卷积神经网络(convolutional neural network, CNN)、循环神经网络(recurrent neural network, RNN)、长短时记忆网络(long short-term memory, LSTM)和门控循环单元(gated recurrent unit, GRU)等。近年来,以双向变换器模型(bidirectional encoder representations from transformers,

BERT)为代表的基于Transformer架构的预训练语言模型逐渐在NER领域崭露头角。文献[10]提出了一种融合BERT与CRF的电力文本实体抽取模型,通过BERT依据上下文信息生成字符级向量并利用CRF层建模标签序列的依赖关系,取得了较好的效果。文献[11]提出了一种融合鲁棒性优化BERT(robustly optimized BERT pretraining approach, RoBERTa)、双向长短时记忆网络(bidirectional long short-term memory, BiLSTM)以及条件随机场的RoBERTa-BiLSTM-CRF模型,通过在RoBERTa与CRF之间加入BiLSTM层来进一步获取字符向量的上下文信息,并验证了其性能。然而,以上方法较少考虑识别过程中嵌套实体的问题。

嵌套实体在电力变压器缺陷记录中的普遍存在增大了NER任务的语义复杂度和识别难度。实体嵌套现象主要表现为一个实体的边界处于其他实体的边界之中,增大了实体边界的模糊性。为应对该挑战,目前较常采用的方法包括基于分层、基于序列生成、基于超图表示、基于大语言模型以及基于跨度这5种。首先,基于分层的方法通过嵌入多层次的层级结构,递归提取嵌套实体。文献[12]为实现对嵌套实体的识别,提出将BiLSTM-CRF层进行动态堆叠的方法,通过将实体信息由内到外传递,直到外部没有实体被识别出来为止。与之不同的是,基于序列生成的方法将实体识别任务转化为序列到序列的生成任务,利用编码器-解码器架构直接输出结构化结果。文献[13]提出使用以恢复被破坏文本为预训练目标的预训练模型(bidirectional and auto-regressive transformers, BART)作为基础模型,将实体线性化为实体指针索引序列,实现对多种类型实体的识别。此外,基于超图表示的方法通过将每个句子用复杂的图结构建模来实现嵌套实体的识别。文献[14]提出了一种分段超图表示方法,通过LSTM捕捉句子语境依赖信息,在拥有较低推理时间复杂度的情况下实现对嵌套实体的识别。而基于大语言模型的方法则是利用大语言模型庞大的预训练知识与强大的上下文理解能力,通过将序列标注任务转化为生成式任务实现对嵌套实体的识别。文献[15]将BERT与全局指针网络(global pointer, GP)相结合,通过对比BERT-GP模型与众多大语言模型对嵌套实体的识别效果,指出大语言模型在灵活性、小样本场景等方面的优势以及资源需求量大、推理速度慢和嵌套实体的位置感知较弱等方面的局限性。最后,基于跨度的方法主要通过实体跨度进行枚举以构建跨度表示并进行分类。文献[16]通过对生成的跨度语义表示矩阵进行语义差异信息提取与边界感知,提升了模型对嵌套实体边界感知和类型预测能力。

针对电力设备检修记录中存在嵌套实体和复杂语境的关键问题,本文以电力变压器缺陷记录为典型数据,提

出一种以乱序预训练语言模型 (pre-training BERT with permuted language model, PERT)、双向门控循环单元 (bidirectional gated recurrent unit, BiGRU) 与高效全局指针网络 (efficient global pointer, EGP) 相结合的电力变压器缺陷记录命名实体识别方法。该方法通过捕捉实体间的语义关系、聚焦实体边界位置信息,在保证模型训练速度的同时能够实现电力变压器缺陷记录中扁平与嵌套实体的精准识别。

1 电力变压器缺陷记录

1.1 电力变压器缺陷记录的特点

运维检修过程中记录的电力变压器缺陷记录具有 4 类特点:1) 实体多样:由于检修人员各自不同的记录习惯,对于一个设备实体会有一种不同的表述,如“主变压器”常被记录为“联变”或“主变”;2) 专业性强:电力变压器缺陷记录中存在大量诸如“冷却器系统”、“硅胶变色”等专业词汇。又如“开关拒动”中“开关”的种类需结合上下文语境进行判断;3) 非结构化:尽管检修人员通常遵循一定的记录模板,但在对缺陷进行详细描述时却呈现非结构化的自然语言特性,难以存在确定的形式;4) 嵌套实体:汉语文本中各实体间存在复杂的层次关系,使得实体间常存在嵌套现象。以“2 号散热片下方主变底箱螺栓发热”为例,该缺陷记录中缺陷部位实体“2 号散热片下方主变底箱螺栓”中包含了缺陷设备实体“主变”,即缺陷设备实体完全处于缺陷部位实体的内部。

1.2 电力变压器缺陷记录的实体类型

针对缺陷记录中多实体混杂所导致的分析效率低下的问题^[17],本文依据电力缺陷记录特点与领域专家经验,将缺陷实体划分为缺陷电压等级、缺陷变电站、缺陷设备、缺陷部位和缺陷现象 5 类实体,构建结构化的标签体系,使检修人员对电力变压器能够开展更为快速地评估与决策。以上划分的 5 类实体类型的示例如表 1 所示。

表 1 实体抽取类型及示例

Table 1 Entity extraction types and examples

实体类型	示例
缺陷电压等级	220 kV、110 kV、10 kV
缺陷变电站	某某变、某某变电站
缺陷设备	#1 主变、联变
缺陷部位	在线监测阀底部、#2 风机
缺陷现象	挂油珠、故障停转

2 PERT-BiGRU-EGP 模型

本文实现电力变压器缺陷记录嵌套命名实体识别的 PERT-BiGRU-EGP 模型结构如图 1 所示,该模型共计 3 层网络结构:PERT 向量嵌入层、BiGRU 文本编码层和 EGP 解码输出层。首先,通过 PERT 向量嵌入层精准地捕捉缺陷记录中字符的上下文语义信息;其次,将向量嵌入层的输出作为 BiGRU 文本编码层的输入,通过双向特征提取,进一步获取更为全面的上下文语义信息;最后,将富含上下文语义信息的向量输入 EGP 解码输出层,从全局角度出发,获取实体的起始与终止位置,实现嵌套实体的精准识别。3 层网络相互配合,共同完成实体识别任务。

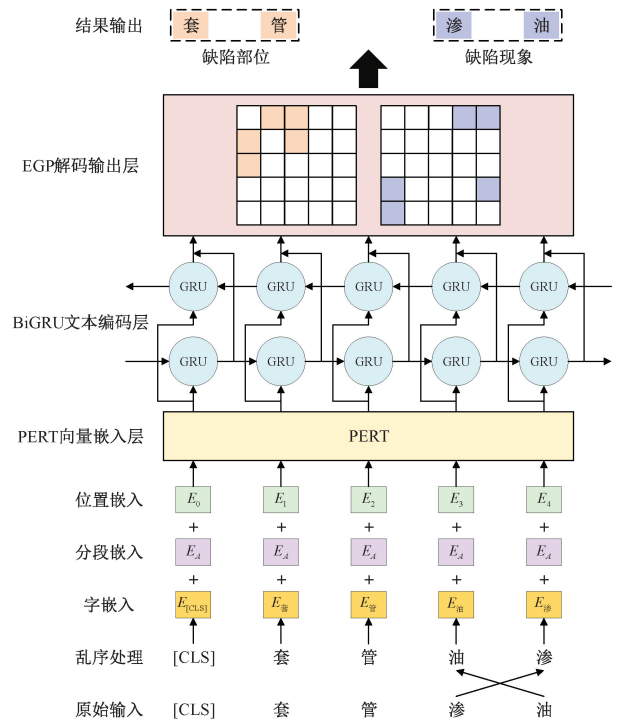


图 1 PERT-BiGRU-EGP 模型结构

Fig. 1 Structure of the PERT-BiGRU-EGP model

2.1 PERT 向量嵌入层

以乱序文本“研表究明,汉字的序顺并不一定会影响阅读”为例,基于人们能从部分字词顺序混乱的文本中获取真实语义的思想,PERT 模型应运而生^[18]。不同于 BERT 模型的核心预训练任务为掩码语言模型 (masked language model, MLM) 与下一句预测 (next sentence prediction, NSP), PERT 模型以乱序语言模型 (permuted language model, PerLM) 作为其预训练任务,模型从文本序列中选择 15% 的字词后,随机选择其中的 90% 进行乱

序处理,剩余 10% 的部分作为未乱序的负样本,目标是预测输入序列中顺序被随机打乱的字词在初始序列中的位置。PERT 模型通过引入全词掩码(whole word masking)和 N-gram 掩码(N-gram masking)的策略以提升其性能,在不添加掩码标记[MASK]的情况下进行文本语义信息的自监督学习,减轻了预训练任务与微调任务之间存在偏差的问题。此外,相较于 BERT 模型的 MLM 任务,PerLM 的预测空间取决于输入序列的长度而非整个词表空间,因此拥有更高的效率。

PERT 模型在将输入文本序列乱序处理后进入输入层,通过对字嵌入(token embedding)、分段嵌入(segment embedding)和位置嵌入(position embedding)进行加和以获取最终的输入表示向量。PERT 模型的结构如图 2 所示。其中,字嵌入将输入的字符表示为一个高维向量;分段嵌入用以获取输入中不同句子与段落的边界;位置嵌入用以添加字符的位置信息。此后,将输入表示向量输入 Transformer 编码层以获取字符的上下文相关表示向量。

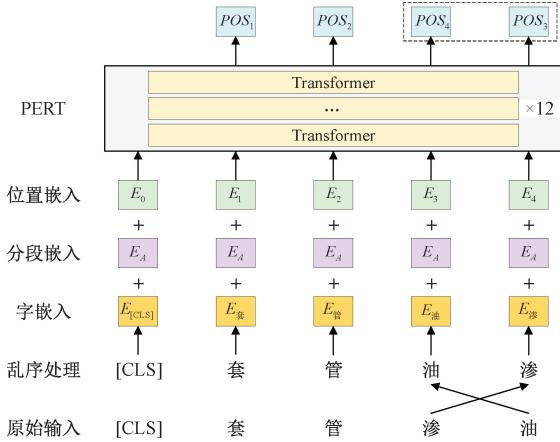


图 2 PERT 模型
Fig. 2 PERT model

Transformer 编码层由 12 个结构相同的 Transformer 编码器堆叠而成,后者结构如图 3 所示。

Transformer 编码器通过自注意力机制(self-attention)计算字符相关性权重来捕捉上下文信息。其计算公式如式(1)~(4)所示。

$$Q = XW^Q \tag{1}$$

$$K = XW^K \tag{2}$$

$$V = XW^V \tag{3}$$

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) V \tag{4}$$

式中: Q, K 和 V 分别表示查询矩阵、键矩阵和值矩阵; X 表示输入矩阵; W^Q, W^K 和 W^V 表示投影矩阵; $Attention(\cdot)$ 表示自注意力机制的输出矩阵; $Softmax(\cdot)$ 表示归一化函数,用以限制矩阵的数值范围; d_k 表示投影维度。

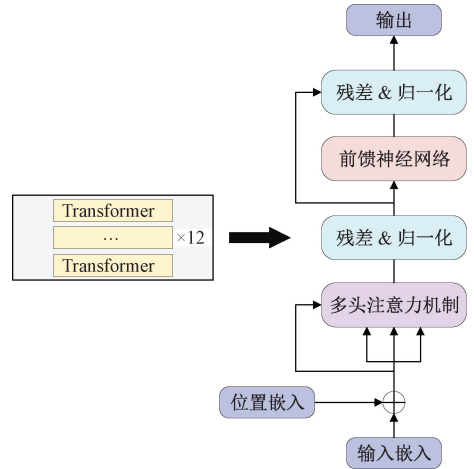


图 3 Transformer 编码器结构
Fig. 3 Structure of Transformer encoder

为增强模型对多样化依赖的捕捉能力,可拼接多个自注意力机制的并行计算结果以形成多头注意力机制(multi-head attention)^[19],提高模型捕捉字词、短语之间关系和上下文信息的能力以适应电力变压器检修文本的复杂结构。其计算公式如式(5)~(6)所示。

$$h_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \tag{5}$$

$$MultiHead(Q, K, V) = Concat(h_1, h_2, \dots, h_m)W^O \tag{6}$$

式中: $MultiHead(\cdot)$ 表示多头注意力机制的输出矩阵; $Concat(\cdot)$ 表示所有自注意力机制输出矩阵的拼接; W^O 表示输出投影矩阵; m 表示头的数量; h_i 表示第 i 个自注意力机制的输出矩阵。

2.2 BiGRU 文本编码层

GRU 与 LSTM 同为 RNN 的变体,都使用门结构来控制信息的传播,在一定程度上解决了 RNN 存在梯度消失、梯度爆炸和长期依赖的问题。相较于包含输入门、遗忘门和输出门 3 个门控单元的 LSTM,GRU 将门结构简化为更新门和重置门,其中,更新门是通过将 LSTM 的输入门与遗忘门的功能进行合并所得到的^[20-21]。可见,相较于 LSTM,GRU 的结构更加简单,这使其在拥有与 LSTM 相近甚至更优性能的同时,能够更加高效地进行训练^[22]。GRU 的结构如图 4 所示。

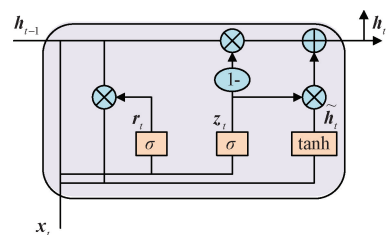


图 4 GRU 结构
Fig. 4 Structure of GRU

单向 GRU 的传播公式如式(7)~(10)所示。

$$\mathbf{r}_t = \sigma(\mathbf{W}_r \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_r) \quad (7)$$

$$\mathbf{z}_t = \sigma(\mathbf{W}_z \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_z) \quad (8)$$

$$\tilde{\mathbf{h}}_t = \tanh(\mathbf{W}_h \cdot [\mathbf{r}_t * \mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_h) \quad (9)$$

$$\mathbf{h}_t = (1 - \mathbf{z}_t) * \mathbf{h}_{t-1} + \mathbf{z}_t * \tilde{\mathbf{h}}_t \quad (10)$$

式中: \mathbf{r}_t 和 \mathbf{z}_t 分别表示重置门和更新门在 t 时刻的输出; \mathbf{x}_t 表示 t 时刻的输入字向量; \mathbf{h}_t 和 \mathbf{h}_{t-1} 分别表示 t 时刻和 $t-1$ 时刻的隐藏状态; $\tilde{\mathbf{h}}_t$ 表示 t 时刻的候选隐藏状态; \mathbf{W}_r 、 \mathbf{W}_z 和 \mathbf{W}_h 分别表示重置门、更新门及候选隐藏状态的权重矩阵; \mathbf{b}_r 、 \mathbf{b}_z 和 \mathbf{b}_h 分别表示重置门、更新门及候选隐藏状态的偏置向量; σ 表示 Sigmoid 激活函数; \cdot 表示点积; $[\]$ 表示向量间的拼接; \tanh 表示双曲正切激活函数; $*$ 表示矩阵的 Hadamard 积。

BiGRU 的结构如图 5 所示。相较于仅能利用历史上下文信息的单向 GRU, BiGRU 通过结合正向与反向 GRU 序列, 充分考虑历史与未来的上下文信息, 弥补了单相 GRU 无法利用全局上下文信息的局限性, 从而增强了序列特征的获取能力^[23]。

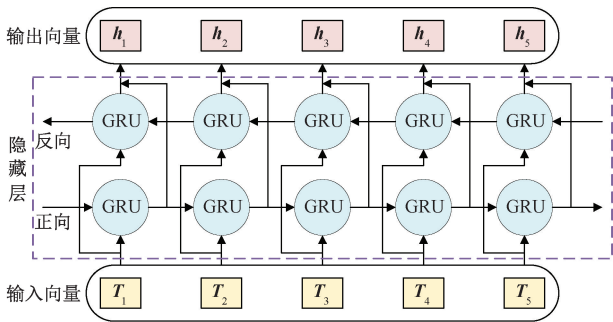


图 5 BiGRU 结构

Fig. 5 Structure of BiGRU

BiGRU 的传播公式如式(11)~(13)所示。

$$\vec{h}_t = GRU(\mathbf{x}_t, \vec{h}_{t-1}) \quad (11)$$

$$\overleftarrow{h}_t = GRU(\mathbf{x}_t, \overleftarrow{h}_{t+1}) \quad (12)$$

$$\mathbf{h}_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (13)$$

式中: \vec{h}_t 和 \overleftarrow{h}_t 分别表示 t 时刻正向与反向的隐藏状态; \vec{h}_{t-1} 表示 $t-1$ 时刻的正向隐藏状态; \overleftarrow{h}_{t+1} 表示 $t+1$ 时刻的反向隐藏状态; 通过拼接 \vec{h}_t 与 \overleftarrow{h}_t , 得到最终的双向隐藏状态 \mathbf{h}_t 作为输出。

2.3 EGP 解码输出层

传统 CRF 模型在解码时需要根据前一个状态来计算当前状态, 导致其无法并行预测所有实体标签, 且由于单个字符只能属于一个标签, 使其应对嵌套实体的能力不足。为此, 不同于只关注实体的起始与结束位置而并

未考虑全局信息的常规指针网络^[24], 全局指针网络通过直接并行计算“全局”中所有实体跨度的分数, 不加区分地将扁平、嵌套实体作为一个整体进行输出, 提高了模型对嵌套实体的识别效率和准确率。

全局指针网络通过构建一个边长与被识别缺陷记录长度 n 相同的, 大小为 $n \times n$ 的正方形矩阵来识别文本中所有可能实体的起始与终止位置, 并将获取的子序列输入实体分类器从而完成分类。由于利用矩阵的上三角部分即可完成实体信息的构造表示, 因此省略了矩阵的下三角部分。在如图 6 所示的识别示例中, 每个单元格各代表一个实体序列, 其中“1”代表子序列的起始与结束位置, 而“0”则代表不存在实体序列。在缺陷部位矩阵中, 由于其第 1 行第 2 列为 1, 故该节选缺陷记录的序列“套管”即为缺陷部位实体。

	套	管	渗	油
套	0	1	0	0
管		0	0	0
渗			0	0
油				0

	套	管	渗	油
套	0	0	0	0
管		0	0	0
渗			0	1
油				0

缺陷部位矩阵

缺陷现象矩阵

图 6 全局指针网络实体识别示例

Fig. 6 Example of NER using global pointer

在只存在一种实体类型的情况下, 缺陷记录中实体起始位置 i 与终止位置 j 所代表的实体位置表征向量 $\mathbf{u}_{i,\alpha}$ 和 $\mathbf{z}_{j,\alpha}$ 如式(14)~(15)所示。

$$\mathbf{u}_{i,\alpha} = \mathbf{W}_{u,\alpha} \mathbf{p}_i + \mathbf{b}_{u,\alpha} \quad (14)$$

$$\mathbf{z}_{j,\alpha} = \mathbf{W}_{z,\alpha} \mathbf{p}_j + \mathbf{b}_{z,\alpha} \quad (15)$$

式中: α 表示实体类型标签; \mathbf{p}_i 和 \mathbf{p}_j 分别表示实体起始、终止位置的字符向量表示; $\mathbf{W}_{u,\alpha}$ 和 $\mathbf{W}_{z,\alpha}$ 分别表示起始、终止索引前馈层的权重矩阵; $\mathbf{b}_{u,\alpha}$ 和 $\mathbf{b}_{z,\alpha}$ 分别表示起始、终止索引前馈层的偏置。

将 \mathbf{u}_i 和 \mathbf{z}_j 输入式(16) 即可得到实体跨度的打分函数 $S_\alpha(i, j)$, 即:

$$S_\alpha(i, j) = \mathbf{u}_{i,\alpha}^T \mathbf{z}_{j,\alpha} \quad (16)$$

相较于使用空格进行分隔的英语单词, 汉语字的边界更加难以确定。在数据量较少的情况下, 全局指针网络在进行识别时可能出现将两个实体的起始、终止位置随机匹配的情况。以缺陷记录“#2 主变本体体积灰严重”为例, 缺陷设备实体“#2 主变”与缺陷部位实体“本体”分别拥有各自的起始、终止位置, 但由于缺少相对位置的引导, 全局指针网络可能将缺陷设备实体的起始位置与缺陷部位实体的终止位置相结合, 错误地将“#2 主变本体”视作一个整体输出。为解决上述问题, 本文将旋转式位

置编码 (rotary position embedding, RoPE)^[25] 引入全局指针网络。该编码方法通过捕捉实体的相对位置信息来强化模型对实体相邻位置的关联性感知,使模型能够更加充分的利用实体边界信息,增强其对实体跨度的敏感性,减少此类随机匹配情况的出现次数。

RoPE 通过将实体起始、终止位置的特定矩阵 R_i 和 R_j 与对应位置的表征向量 $u_{i,\alpha}$ 和 $z_{j,\alpha}$ 相乘,即可得到引入相对位置 $(j-i)$ 的新实体跨度打分函数,如式(17)所示。

$$S_{\alpha}(i,j) = (R_i u_{i,\alpha})^T (R_j z_{j,\alpha}) = u_{i,\alpha}^T R_{j-i} z_{j,\alpha} \quad (17)$$

针对全局指针网络参数量较多、计算开销较大的问题,可以通过在全局指针网络中引入参数共享优化策略来解决,经过优化后的全局指针网络即为高效全局指针网络^[26-27]。该优化策略原理是将实体识别任务分为实体抽取与实体分类两个子任务,通过将实体抽取阶段的参数在不同类型实体之间进行共享,使得实体类别扩充时的参数量增加主要集中在分类任务中。由此高效全局指针网络在面对实体类型较多的情况时能够有效地降低模型的整体参数量和计算复杂度,其实体跨度打分函数如式(18)所示。

$$S_{\alpha}(i,j) = (W_{\alpha} p_i)^T (W_{\alpha} p_j) + w_{\alpha}^T [p_i; p_j] \quad (18)$$

为进一步减少参数,可以用跨度 $[u_i; z_j]$ 、 $[u_j; z_i]$ 代替 p_i 和 p_j , 所得到的最终跨度打分函数,如式(19)所示。

$$S_{\alpha}(i,j) = u_i^T z_j + w_{\alpha}^T [u_i; z_j; u_j; z_i] \quad (19)$$

式中: w_{α} 表示实体类型标签 α 的判别权重; $[u_i; z_j; u_j; z_i]$ 表示跨度表征。

此外,对于长度为 n 的缺陷记录,由于存在 $n(n+1)/2$ 个候选实体,故当缺陷记录长度较长时可能出现实际实体数远小于候选实体数的情况,即出现类别不平衡问题。如式(20)所示,为解决此问题,可以在网络中加入损失函数 L , 通过将多标签分类问题归结为目标类别与非目标类别之间得分的差异来平衡实体类型的权重。

$$L = \log\left(1 + \sum_{(i,j) \in P_{\alpha}} e^{-S_{\alpha}(i,j)}\right) + \log\left(1 + \sum_{(i,j) \in Q_{\alpha}} e^{S_{\alpha}(i,j)}\right) \quad (20)$$

式中: P_{α} 表示类型为 α 的实体跨度集合; Q_{α} 表示非实体或非 α 类型的实体跨度集合。

3 实验与分析

3.1 数据准备

本文选用某市国家电网公司近年来记录的电压等级在 10~220 kV 的 6 729 条电力变压器缺陷记录作为可用数据集。考虑到序列标注方法通常适用于单分类任务,难以应对单个字符同时归属多个标签的情况,因此本文

从电力变压器缺陷数据中存在复杂成分嵌套实体的现实出发,提出使用基于跨度标注的方法对关键实体进行标注,通过标注实体的起始与结束位置,结构化地存储电力变压器缺陷记录及其实体信息,并按照 8:1:1 的比例划分为训练集、验证集和测试集。数据集标注示例如图 7 所示。

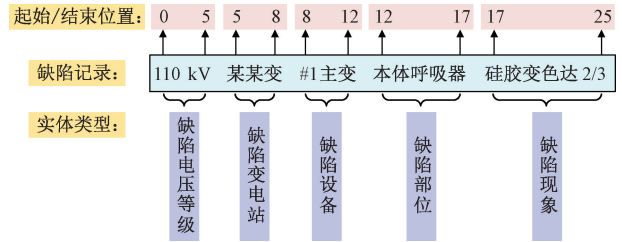


图 7 数据集标注示例

Fig. 7 Dataset annotation example

3.2 评价指标与环境配置

为了评估本方法在电力变压器缺陷记录实体识别任务中的有效性,本文选取精确率 (Precision, P)、召回率 (Recall, R) 以及 F1 分数 (F1-Score) 作为评估模型性能的核心指标。其中, P 是模型正确识别的实体数量与识别出的总实体数量之比,用以衡量模型预测结果的准确性; R 是模型正确识别的实体数量与实体总数量之比,用以衡量模型实体识别的完整性; $F1$ 分数是精确率和召回率的加权调和平均数,用以衡量模型在识别准确性与完整性之间的整体效能。三者的计算公式如式(21)~(23)所示。

$$P = \frac{TP}{TP + FP} \quad (21)$$

$$R = \frac{TP}{TP + FN} \quad (22)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (23)$$

式中: TP (true positive) 指模型正确预测正类样本为正类的数量; FP (false positive) 指模型错误预测负类样本为正类的数量; FN (false negative) 则是指模型错误预测正类样本为负类的数量。

本文所有实验所处环境的主要配置为: Windows 10 x64 操作系统、Intel (R) Xeon (R) w5-2465X 处理器、NVIDIA GeForce RTX 4090 显卡,并基于 Python 3.8 编程语言与 PyTorch 2.0.1 深度学习框架。本文所提出的 PERT-BiGRU-EGP 模型的关键参数配置如表 2 所示。

3.3 对比实验与分析

为验证本文所提模型在电力变压器缺陷记录嵌套实体识别任务上的有效性,在此选择与条件随机场 (CRF)

表 2 模型关键参数设置

Table 2 Model parameter configuration

参数名称	数值
优化器	Adam
学习率	5×10^{-5}
批次大小	16
最大序列长度	256
BiGRU 隐藏层维数	128
随机失活率	0.2

模型、双向长短时记忆网络-条件随机场 (BiLSTM-CRF) 模型、双向门控循环单元-条件随机场 (BiGRU-CRF) 模型、BERT 预训练模型-双向长短时记忆网络-条件随机场 (BERT-BiLSTM-CRF) 模型、PERT 预训练模型-条件随机场 (PERT-CRF) 模型、适用于嵌套实体识别的 BERT 预训练模型-双仿射注意力 (BERT-biaffine attention, BERT-Biaffine) 模型以及 BERT 预训练模型-机器阅读理解 (BERT-machine reading comprehension, BERT-MRC) 模型等 7 种模型进行对比实验,其结果如表 3 所示。

表 3 模型对比实验结果

Table 3 Comparative experimental results of models

实验结果	CRF	BiLSTM-CRF	BiGRU-CRF	BERT-BiLSTM-CRF	PERT-CRF	BERT-Biaffine	BERT-MRC	PERT-BiGRU-EGP(本文模型)	
训练时间/min	31.15	40.12	37.18	49.25	45.72	56.35	109.26	26.38	
<i>P</i> /%	缺陷电压等级	87.50	99.22	98.45	99.25	98.50	99.21	98.51	99.25
	缺陷变电站	93.80	99.26	98.53	98.90	98.89	99.26	98.90	100.00
	缺陷设备	95.47	98.03	98.57	97.47	98.00	98.60	99.15	100.00
	缺陷部位	83.38	86.47	87.65	91.76	91.10	93.77	94.15	94.31
	缺陷现象	74.93	85.83	85.56	86.35	85.88	88.24	87.78	88.52
	综合评价	86.57	92.65	92.76	93.84	93.67	95.10	95.08	95.78
<i>R</i> /%	缺陷电压等级	90.15	96.97	96.21	100.00	99.24	95.45	100.00	100.00
	缺陷变电站	95.54	100.00	100.00	100.00	99.63	100.00	100.00	100.00
	缺陷设备	96.01	99.43	98.01	98.86	97.72	100.00	99.72	100.00
	缺陷部位	84.89	88.82	90.03	94.26	92.75	95.47	92.45	95.17
	缺陷现象	75.99	87.29	87.01	87.57	84.18	88.98	89.27	89.27
	综合评价	87.89	93.88	93.67	95.34	93.74	95.82	95.55	96.24
<i>F1</i> /%	缺陷电压等级	88.81	98.08	97.32	99.62	98.87	97.30	99.25	99.62
	缺陷变电站	94.66	99.63	99.26	99.45	99.26	99.63	99.45	100.00
	缺陷设备	95.74	98.73	98.29	98.16	97.86	99.29	99.43	100.00
	缺陷部位	84.13	87.63	88.82	93.00	91.92	94.61	93.29	94.74
	缺陷现象	75.46	86.55	86.27	86.96	85.02	88.61	88.52	88.89
	综合评价	87.22	93.26	93.21	94.58	93.70	95.46	95.31	96.01

由表 3 可知,得益于对实体内部字符间的依赖关系的感知和实体边界信息的关注,本文所提出的 PERT-BiGRU-EGP 模型在存在嵌套实体的命名实体识别任务中表现出色,其综合评价指标 *P*、*R* 和 *F1* 均高于其他模型,分别达到了 95.78%、96.24% 和 96.01%,较其余模型取得的最优指标分别高出 0.68%、0.42% 和 0.55%。

其次,PERT-BiGRU-EGP 模型的综合 *F1* 分数相较于 CRF、BiLSTM-CRF 及 BiGRU-CRF 模型,分别高出了 8.79%、2.75% 与 2.80%。且通过在 BiLSTM-CRF 模型中加入预训练 BERT 模型,BERT-BiLSTM-CRF 取得了更好

的结果,其综合 *F1* 分数高出 BiLSTM-CRF 模型 1.32%,证明预训练模型的加入有助于提升模型对文本上下文语义关系的学习能力。

随后,通过对比表中 BiLSTM-CRF 和 BiGRU-CRF 模型的实验结果可知,BiGRU-CRF 模型的综合 *F1* 分数虽略低于 BiLSTM-CRF 模型 0.05%,但在训练时间上缩短了 7.33%。可见 BiGRU 在获取与 BiLSTM 相近效果的同时,能够一定程度上降低模型的训练时间与资源消耗。同样对比 3 种适用于嵌套实体识别的 BERT-Biaffine、BERT-MRC 和 PERT-BiGRU-EGP 模型效果可知,本文提

出的 PERT-BiGRU-EGP 模型不仅训练时间大幅降低,而且使模型的综合 $F1$ 分数分别提升了 0.55%、0.70%。各模型训练时间对比如图 8 所示。

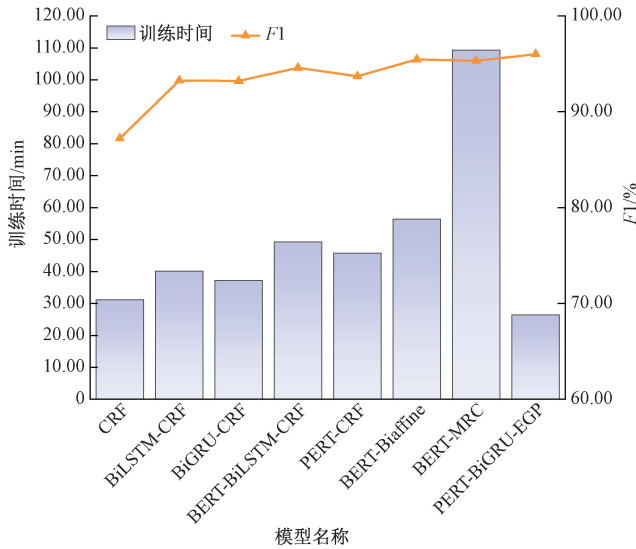


图 8 各模型训练时间对比

Fig. 8 Comparison of training time across models

此外,为验证本文所提方法在应对嵌套实体复杂识别场景的有效性,在此选取了综合 $F1$ 分数排名在前 4 位的 BERT-BiLSTM-CRF、BERT-Biaffine、BERT-MRC 和 PERT-BiGRU-EGP 模型分别对不同嵌套实体占比的数据集进行识别,其综合 $F1$ 分数对比如表 4 所示。

表 4 不同嵌套实体占比模型效果对比

Table 4 Comparative analysis of model performance under different proportions of nested entities (%)

模型	评价指标	不同嵌套实体占比				
		10	30	50	70	90
BERT-BiLSTM-CRF	$F1$	91.81	85.55	74.10	69.23	60.40
BERT-Biaffine	$F1$	95.28	94.43	92.58	88.15	86.09
BERT-MRC	$F1$	94.14	93.23	91.75	87.79	85.90
PERT-BiGRU-EGP	$F1$	95.79	95.16	93.28	90.03	87.61

需要注意的是,随着嵌套深度的不断加深,缺陷记录的文本复杂度也在不断增加,这将对模型识别造成一定的挑战,使所得指标出现不断下降的趋势。实验结果表明,适用于扁平实体识别的 BERT-BiLSTM-CRF 模型在全部 5 种数据集上的表现均不及具备嵌套实体识别能力的其余 3 种模型且存在较大的差距。而 PERT-BiGRU-EGP 模型则在各数据集中均取得了最佳的表现,以含有最高嵌套实体占比 90% 的数据集为例,PERT-BiGRU-EGP 模

型的识别效果分别领先 BERT-BiLSTM-CRF、BERT-Biaffine 和 BERT-MRC 模型 27.21%、1.52% 和 1.71%。

最后,通过观察表中不同实体类型的实验结果可以发现,模型在不同实体类型中表现存在较大差异。首先,对于实体标签“缺陷电压等级”与“缺陷变电站”,各模型均取得了较高的评价指标,除 CRF 模型外,其 P 、 R 和 $F1$ 都接近或等于 100%。这是由于缺陷记录中二者对应的实体拥有较为固定的表述且较少嵌套于其他标签的实体之中,使得模型能够较为轻松地学习其特征;其次,对于存在较多嵌套实体的“缺陷设备”与“缺陷部位”标签,本文所提出的 PERT-BiGRU-EGP 模型均取得了最优的 $F1$ 分数,高出适应嵌套实体识别的 BERT-Biaffine 和 BERT-MRC 模型中的最优值 0.57% 和 0.13%,证明了本文所提出模型在嵌套实体识别领域的优越性;最后,对于“缺陷现象”标签而言,虽然其实体嵌套现象并不明显,但因其描述多样性较高,影响了模型学习的效果,导致该标签的评价指标在各个模型的实验结果中均低于其他 4 类实体。对于该标签,本文所提出模型取得的综合评价指标分别为 88.52%、89.27% 和 88.89%,同样取得了最好的效果。为更加直观地展示本文所提出的 PERT-BiGRU-EGP 模型在 5 类实体类型中所取得的出色结果,另外选取了其余模型中综合 $F1$ 分数最高的嵌套实体识别模型 BERT-Biaffine 以及综合 $F1$ 分数最高的扁平实体识别模型 BERT-BiLSTM-CRF 进行对比,其各实体类型的 $F1$ 分数如图 9 所示。

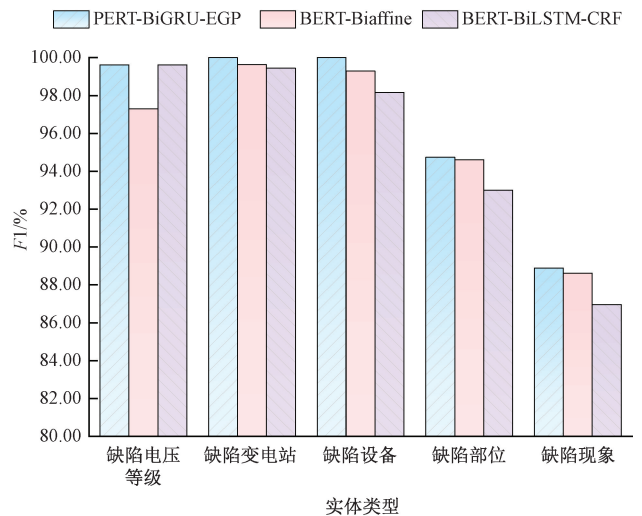


图 9 3 种模型的各实体类型 $F1$ 分数对比

Fig. 9 Comparison of $F1$ score for each entity type across three models

3.4 消融实验与分析

为分析本文所提出的 PERT-BiGRU-EGP 模型中各功能层的性能贡献,在相同环境下进行的消融实验结果

如表5所示。表中符号“√”代表实验中使用了此模型,而空白则代表并未使用。

表5 模型消融实验结果

Table 5 Ablation experimental results of models

PERT	BERT	BiGRU	EGP	评价指标		
				<i>P</i>	<i>R</i>	<i>F1</i>
		√	√	92.62	93.46	93.04
√			√	95.55	95.62	95.58
√		√		93.78	94.50	94.14
	√	√	√	95.08	95.41	95.24
√		√	√	95.78	96.24	96.01

由表5中结果可知,PERT-BiGRU-EGP模型的*F1*分数均处于领先地位,证明各功能层的加入对模型性能提升均有正面效果。具体来说,PERT-BiGRU-EGP模型的综合*F1*分数分别高出BiGRU-EGP、PERT-EGP、PERT-BiGRU和BERT-BiGRU-EGP模型2.97%、0.43%、1.87%和0.77%。其中,移除PERT的BiGRU-EGP模型*F1*分

数下降最多,证明PERT功能层的加入对模型捕获深层语义信息、应对复杂语境具有关键作用。此外,将PERT替换为BERT后模型效果有所降低,证明了PERT模型的优越性。

3.5 案例分析

以缺陷记录“110 kV 某某变本体在线油色谱检测装置靠#1主变取油样活门处漏油,约3 s一滴”为例,各标签与标准实体的对应关系包括:“缺陷电压等级:110 kV”、“缺陷变电站:某某变”、“缺陷设备:#1主变”、“缺陷部位:本体在线油色谱检测装置靠#1主变取油样活门处”以及“缺陷现象:漏油,约3 s一滴”。由于缺陷设备实体嵌套于缺陷部位之中,增大了文本解析的复杂度,对参与实体识别各模型对嵌套实体的语义辨别能力提出了一定的要求。为验证本文所提出模型具备在嵌套实体与复杂语境下的较强适应性,从参与对比实验的8种模型中选出具有代表性的CRF、BERT-BiLSTM-CRF、BERT-Biaffine和PERT-BiGRU-EGP(本文模型)模型分别对该缺陷记录进行命名实体识别,所得到的识别结果如表6所示。标注“(×)”的标签为识别错误项,未标注则为正确。

表6 各模型识别结果案例分析

Table 6 Case analysis of recognition results by various representative models

模型	缺陷电压等级/kV	缺陷变电站	缺陷设备	缺陷部位	缺陷现象
标准结果	110	某某变	#1主变	本体在线油色谱检测装置靠#1主变取油样活门处	漏油,约3 s一滴
CRF	110	某某变	#1主变	本体在线油色谱检测装置(×)	漏油(×)
BERT-BiLSTM-CRF	110	某某变	(×)	本体在线油色谱检测装置靠#1主变取油样活门处	漏油,约3 s一滴
BERT-Biaffine	110	某某变	#1主变	本体在线油色谱检测装置靠#1主变(×)	漏油,约3 s一滴
本文模型	110	某某变	#1主变	本体在线油色谱检测装置靠#1主变取油样活门处	漏油,约3 s一滴

由表6可知,得益于“缺陷电压等级”与“缺陷变电站”标签较低的实体多样性与复杂性,各模型均对其完成了准确识别,而对于更具挑战性的其余3类标签,各模型效果出现了较大差异。其中,对于“缺陷设备”与“缺陷部位”标签,由于嵌套实体的存在,适用于扁平实体识别的CRF、BERT-BiLSTM-CRF模型效果不及BERT-Biaffine和PERT-BiGRU-EGP模型。具体来说,CRF模型因表征能力有限,未能正确识别出缺陷部位实体,而BERT-BiLSTM-CRF模型通过引入语义理解与序列建模能力,实现了对该实体的完整识别,但遗漏了缺陷设备实体“#1主变”;BERT-Biaffine模型虽准确识别出缺陷设备实体,但因对实体边界检测不足而遗漏了复杂语境下缺陷部位中的“取油样活门处”;PERT-BiGRU-EGP模型通

过深度语义表示、序列上下文建模与全局实体边界评估,实现了对缺陷设备、缺陷部位实体的准确识别,表明该模型具备应对嵌套实体与复杂语境难题的能力。此外,对于“缺陷现象”标签,由于缺陷记录中有关“漏油”的标注一致性较高,各模型能够更加充分的学习到该实体类型的特征,使得除CRF模型未完整识别到该实体外,其余模型均正确识别出该缺陷现象。

4 结 论

本文针对电力变压器缺陷记录存在嵌套实体与复杂语境的特点,提出基于PERT-BiGRU-EGP模型的实体识别方法。首先,PERT的加入使得模型能够捕获缺陷记

录中的上下文深层语义信息;其次,BiGRU采用更加简单的门结构,通过结合正向与反向GRU序列,快速且全面地获取上下文语义信息;最后,考虑到传统CRF在嵌套实体处理方面的不足,本文提出使用EGP作为解码输出层,从全局角度出发,并行计算所有实体的头尾位置,无差别地输出扁平化和嵌套实体,实现对关键实体的快速精准识别。实验结果表明,该方法不仅拥有较快的训练速度,而且在5类标签的实体识别任务中表现优异,在全部5类标签中均取得了最优的F1分数,尤其对实体嵌套现象较为普遍的“缺陷设备”与“缺陷部位”标签,该方法展现出显著优势。需要指出的是,现阶段工作主要基于数据驱动范式,对如何将专业词典、设备实体间关系和缺陷演化模式等电力领域的结构化专业知识融入模型的探索尚显不足。后续研究可通过引入知识增强的预训练任务、知识图谱约束等方式来构建融合数据与知识的混合模型,进一步提升模型在电力变压器缺陷记录等领域专业文本上的识别精度与鲁棒性。

参考文献

- [1] 李元,李睿,林金山,等. 基于字词混用集成模型的电力变压器缺陷记录文本挖掘方法[J]. 电力工程技术, 2024, 43(6):153-162.
- LI Y, LI R, LIN J SH, et al. Character-word level ensemble integrated model for power transformer defect recording text mining method[J]. Electric Power Engineering Technology, 2024, 43(6):153-162.
- [2] 周杨泽,苏运,姚睿洋,等. 用电大数据共享关键技术与研究展望[J/OL]. 中国电机工程学报, 1-18[2025-09-05].
- ZHOU Y Z, SU Y, YAO R Y, et al. Key technologies and research prospects of smart meter data sharing [J/OL]. Proceedings of the CSEE, 1-18[2025-09-05].
- [3] 陈静,郑垂锭,李桂敏,等. 考虑行业关联度的工业用户用电异常识别研究[J]. 仪器仪表学报, 2023, 44(4):72-81.
- CHEN J, ZHENG CH D, LI G M, et al. Research on power consumption anomaly identification of industrial users considering industry relevance[J]. Chinese Journal of Scientific Instrument, 2023, 44(4):72-81.
- [4] 缪希仁,林蔚青,肖洒,等. 基于条件互信息与LSTNet的特高压变压器顶层油温预测方法[J]. 电网技术, 2022, 46(7):2601-2609.
- MIAO X R, LIN W Q, XIAO S, et al. Forecasting method for top oil temperature in ultra-high voltage transformers based on conditional mutual information and LSTNet[J]. Power System Technology, 2022, 46(7):2601-2609.
- [5] 丁建平,李卫军,刘雪洋,等. 命名实体识别研究综述[J]. 计算机工程与科学, 2024, 46(7):1296-1310.
- DING J P, LI W J, LIU X Y, et al. A review of named entity recognition research[J]. Computer Engineering & Science, 2024, 46(7):1296-1310.
- [6] 夏海南,贾宁,郭毅,等. 基于数据插值算法的潮流能发电装置发电性能分析优化方法研究[J]. 仪器仪表学报, 2025, 46(10):179-188.
- XIA H N, JIA N, GUO Y, et al. Research on optimization analysis method of power performance assessment for tidal energy converters based on data interpolation algorithms [J]. Chinese Journal of Scientific Instrument, 2025, 46(10):179-188.
- [7] 叶青,赖焯,程春雷,等. 融合词汇增强和跨度方法的中医药命名实体识别[J]. 计算机工程与应用, 2025, 61(17):222-231.
- YE Q, LAI X, CHENG CH L, et al. Named entity recognition for traditional Chinese medicine with lexical enhancement and span method[J]. Computer Engineering and Applications, 2025, 61(17):222-231.
- [8] 胡泽,李文君,杨宏宇. 基于字符表示学习与时序边界扩散的网络安全实体识别方法[J]. 电子与信息学报, 2025, 47(5):1554-1568.
- HU Z, LI W J, YANG H Y. A cybersecurity entity recognition approach based on character representation learning and temporal boundary diffusion [J]. Journal of Electronics & Information Technology, 2025, 47(5):1554-1568.
- [9] 钱丽萍,崔雨婷,廉露,等. 低资源场景下的命名实体识别研究综述[J]. 计算机工程与应用, 2025, 61(18):1-23.
- QIAN L P, CUI Y T, LIAN L, et al. Survey of named entity recognition for low-resource scenarios [J]. Computer Engineering and Applications, 2025, 61(18):1-23.
- [10] 赵磊,张越,蒙飞,等. 电力调度专业语言理解混合神

- 神经网络建模方法[J]. 电子器件, 2024, 47(6): 1640-1644.
- ZHAO L, ZHANG Y, MENG F, et al. Hybrid neural network modeling method for professional language understanding of electric power dispatching[J]. Chinese Journal of Electron Devices, 2024, 47(6): 1640-1644.
- [11] 田雪涵,董坤,赵剑锋,等. 基于增强优化预训练语言模型的电力数据实体识别方法[J]. 智慧电力, 2024, 52(6): 100-107.
- TIAN X H, DONG K, ZHAO J F, et al. Entity recognition method for power data based on enhanced optimization pre-trained language model [J]. Smart Power, 2024, 52(6): 100-107.
- [12] JU M ZH, MIWA M, ANANIADOU S. A neural layered model for nested named entity recognition[C]. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics, 2018: 1446-1459.
- [13] YAN H, GUI T, DAI J Q, et al. A unified generative framework for various NER subtasks[C]. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021: 5808-5822.
- [14] WANG B L, LU W. Neural segmental hypergraphs for overlapping mention recognition[C]. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018: 204-214.
- [15] 张逸勤,邓三鸿,王东波. 基于生成式大语言模型的非遗文本嵌套命名实体识别研究[J]. 现代情报, 2025, 45(10): 26-38.
- ZHANG Y Q, DENG S H, WANG D B. Research on nested named entity recognition of intangible cultural heritage texts based on generative language models[J]. Journal of Modern Information, 2025, 45(10): 26-38.
- [16] 蔡宇翔,骆姐,甘洋镭,等. 基于跨度边界感知的嵌套命名实体识别[J]. 软件学报, 2024, 35(11): 5149-5162.
- CAI Y X, LUO D, GAN Y L, et al. Nested named entity recognition based on span boundary perception [J]. Journal of Software, 2024, 35(11): 5149-5162.
- [17] 费正明,袁可为,黄弘扬,等. 基于 BERT-MRC 的电网现场作业文本关键实体识别方法[J]. 浙江电力, 2025, 44(7): 24-32.
- FEI ZH M, YUAN K W, HUANG H Y, et al. A BERT-MRC-based method for key entity recognition in power grid field operation texts[J]. Zhejiang Electric Power, 2025, 44(7): 24-32.
- [18] CUI Y M, YANG Z Q, LIU T. PERT: Pre-training BERT with permuted language model[J]. ArXiv preprint arXiv: 2203.06906, 2022.
- [19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]. Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 6000-6010.
- [20] 于春雨,高健玮,汤培泉,等. 基于 MPSR 和 IHBA-BiGRU-Attention 的短期电力负荷预测[J/OL]. 电子测量技术, 1-14[2025-12-31].
- YU CH Y, GAO J W, TANG P Q, et al. Short-term power load forecasting based on MPSR and IHBA-BiGRU-attention[J/OL]. Electronic Measurement Technology, 1-14[2025-12-31].
- [21] 王晓晨,许昕,潘宏侠,等. 基于 ResNet-RFA 和 BiGRU-SATT 的弧齿锥齿轮箱故障诊断[J/OL]. 电子测量技术, 1-10[2026-01-22].
- WANG X CH, XU X, PAN H X, et al. Fault diagnosis of spiral bevel gearbox based on ResNet-RFA and BiGRU-SATT[J/OL]. Electronic Measurement Technology, 1-10[2026-01-22].
- [22] 陈娜,孙艳秋,燕燕. 结合注意力机制的 BERT-BiGRU-CRF 中文电子病历命名实体识别[J]. 小型微型计算机系统, 2023, 44(8): 1680-1685.
- CHEN N, SUN Y Q, YAN Y, et al. Named entity recognition for Chinese electronic medical record based on BERT-BiGRU-CRF and attention mechanism[J]. Journal of Chinese Computer Systems, 2023, 44(8): 1680-1685.
- [23] 邹昕澎,李平华,支源,等. 随机振动与温度耦合环境下 MEMS 陀螺仪误差联合补偿技术研究[J]. 仪器仪表学报, 2025, 46(10): 294-306.
- ZOU X P, LI P H, ZHI Y, et al. Research on joint compensation technology for MEMS gyroscope errors in

coupled random vibration and temperature environments[J]. Chinese Journal of Scientific Instrument, 2025, 46(10):294-306.

- [24] LI H J, CHENG M ZH, YANG Z L, et al. Named entity recognition for Chinese based on global pointer and adversarial training[J]. Scientific reports, 2023, 13(1): 3242.
- [25] SU J L, AHMED M, LU Y, et al. RoFormer: Enhanced transformer with rotary position embedding[J]. Neurocomputing, 2024, 568: 127063.
- [26] 林蔚青, 郑垂锭, 陈静, 等. 融合 PERT 与高效全局指针网络的电力变压器缺陷文本实体识别方法[J]. 电网技术, 2025, 49(11):4876-4887.
- LIN W Q, ZHENG CH D, CHEN J, et al. Text entity recognition method for power transformer defects combining PERT and efficient global pointer[J]. Power System Technology, 2025, 49(11):4876-4887.
- [27] SU J L, MURTADHA A, PAN SH F, et al. Global pointer: Novel efficient span-based approach for named entity recognition[J]. ArXiv preprint arXiv: 2208.03054, 2022.

作者简介



Miao Xiren received his B.Sc. degree from Beihang University in 1986, and his M.Sc. and Ph.D. degrees both from Fuzhou University in 1989 and 2000, respectively. He is currently a professor at Fuzhou University. His main research interests include intelligent technology for electrical appliances and systems, as well as online monitoring and diagnosis of electrical equipment.



陈静 (通信作者), 分别于 2010 年、2013 年和 2016 年于厦门大学获得学士、硕士和博士学位, 现为福州大学副教授, 主要研究方向为电力大数据和智能电网。

Chen Jing (Corresponding author) received her B.Sc., M.Sc., and Ph.D. degrees all from Xiamen University in 2010, 2013, and 2016, respectively. She is currently an associate professor at Fuzhou University. Her main research interests include power big data and smart grid.

缪希仁, 1986 年于北京航空航天大学获得学士学位, 分别于 1989 年和 2000 年于福州大学获得硕士和博士学位, 现为福州大学教授, 主要研究方向为电器及其系统智能化技术、电气设备在线监测与诊断。

E-mail: mxr@fzu.edu.cn

陈静 (通信作者), 分别于 2010 年、2013 年和 2016 年于厦门大学获得学士、硕士和博士学位, 现为福州大学副教授, 主要研究方向为电力大数据和智能电网。

E-mail: chenjing@fzu.edu.cn