

## 面向危险化学品事故的事理图谱抽取方法研究

陈观林<sup>1,2</sup>,孔振华<sup>1,2</sup>,汪俊霖<sup>1\*</sup>,翁文勇<sup>1</sup>

(1.浙大城市学院 计算机与计算科学学院,浙江 杭州 310015;

2.浙江大学 计算机学院,浙江 杭州 310027)

**摘要:**为防范危险化学品的储运事故衍生出化学事故的发生,优化事故风险识别流程以减少对人工的依赖。首先从中文语料角度,基于融入了诸多语义信息,通过多粒度空洞卷积网络,利用联合预测层来推断词与词之间的关系,抽取危化品事故实体;然后,针对已经抽取的实体,融入位置编码,对每个主语以及对应的多个关系,识别出宾语,达到危化品事件关系识别。结果表明,实体抽取模型的F1值在自建数据集和公开数据集分别达到92.91%和96.59%,关系识别模型的F1值达到了76.04%;与现有方法相比,两个模型性能均有提升,尤其是关系识别模型优势明显,取得了6.28%的性能领先。抽取的危化品实体和关系,进一步建立了危化品事理图谱及预警系统。

**关键词:**多粒度空洞卷积网络;联合预测层;实体抽取;位置编码;关系识别

**中图分类号:**TP391 **文献标志码:**A **文章编号:**0253-2395(2025)05-0871-09

## Research on the Construction Method of Eventic Graph of Hazardous Chemicals Accidents

CHEN Guanlin<sup>1,2</sup>, KONG Zhenhua<sup>1,2</sup>, WANG Junlin<sup>1\*</sup>, WENG Wenyong<sup>1</sup>

(1. School of Computer and Computing Science, Hangzhou City University, Hangzhou 310015, China;

2. College of Computer Science, Zhejiang University, Hangzhou 310027, China)

**Abstract:** In order to prevent the occurrence of chemical accidents derived from storage and transportation accidents of hazardous chemicals, the accident risk identification process is optimized to reduce the reliance on manual work. Firstly, from the perspective of Chinese corpus, based on the incorporation of many semantic information, the hazardous chemicals accident entities are extracted by Multi-Granularity Dilated Convolution Networks using Co-Predictor Layer to infer the relationship between words; then, for the entities that have been extracted, positional coding is incorporated to recognize the object of each subject as well as the corresponding multiple relations to achieve the Hazardous Chemicals Incident Relationship Recognition. The results show that the F1 value of the Namedity Extraction model reaches 92.91% and 96.59% in the self-constructed and public datasets, respectively, and the F1 value of the Relation Recognition model reaches 76.04%; the performance of the two models improves compared with the existing methods, and the Relation Recognition model, in particular, has a clear advantage and achieves a performance lead of 6.28%. Named Entity Recognition of hazardous chemicals and Relation Extraction further establishes the hazardous chemicals Eventic Graph and early warning system.

**Key words:** multi-granularity dilated convolution networks; co-predictor layer; named entity recognition; positional encoding; relation extraction

收稿日期:2024-05-30;接受日期:2025-03-08

基金项目:浙江省重点研发计划项目(2020C03091)

作者简介:陈观林(1978-),男,浙江台州人,博士,教授,研究方向为人工智能、大数据分析。E-mail:chenguanlin@hzcu.edu.cn

\* 通信作者:汪俊霖(WANG Junlin),E-mail:2230101017@stu.hzcu.edu.cn

引文格式:陈观林,孔振华,汪俊霖,等.面向危险化学品事故的事理图谱抽取方法研究[J].山西大学学报(自然科学版),2025,48(5):871-879. DOI:10.13451/j.sxu.ns.2025013.

## 0 引言

一直以来,危险化学品在各行各业扮演着必不可少的角色,在工业、医疗、农业以及日常用品等领域,都提供显著的价值和收益。尽管如此,国内的危险化学品事故频发,如2015年8月,天津市滨海新区某一化学品仓库发生了重大火灾爆炸案,导致165人遇难,798人受伤<sup>[1]</sup>。危化品事件的发生凸显出在事故预警方面存在的不足,尤其是针对这类化学品的安全管理机制尚有改进空间。因此,如何有效融合危险化学品的专业知识与现代科技手段,构建一套高效的预警系统,成为当前亟待攻克的难题。

国内外学者对于危化品的预警开展了大量研究。陈文坚<sup>[2]</sup>通过对危化品的物质特性和生产工艺角度,预测事故后果,提出针对性措施。Yuan等<sup>[3]</sup>深入探究事故的动态演化路径,挖掘每个事件环节,构建了一套火灾风险源识别模型。王好一等<sup>[4]</sup>通过蝴蝶结模型分析事故前因后果,后经贝叶斯网络模型,建立动态风险评估系统。Je等<sup>[5]</sup>通过光电离检测器传感器和网络控制程序,实现有机化合物的多通道连续监测。贾进章等<sup>[6]</sup>则提出一种基于网络层次分析,并运用灰色聚类的火灾风险评价方法。马全宇等<sup>[7]</sup>基于深度学习模型,设计了一套海关危化品数智能监管系统,识别并分析危险化学品。目前对于危化品的预警大部分依靠高成本的人工标注,通过对危化品事件演化关系从而寻求解决方法的策略较少,同时互联网中大量积累的危化品事件数据未被有效利用,由此本文采用事理图谱对危化品事件演变展开深入分析,构建危化品预警系统。

2012年,谷歌团队提出了知识图谱,引起全球学术界以及工业界的广泛关注,国内外学者随即开展研究,包括对通用知识图谱的构建,和对某些特殊领域知识图谱的构建<sup>[8-11]</sup>。然而知识图谱存在一定局限性,主要体现在知识图谱一般只描绘静态的确定性事实,研究名词实体、实体属性和实体之间的关系,缺乏对事物逻辑演化和发展过程的描绘能力。事理图谱是一种事理逻辑知识库,它以事件为节点,以事件之间的因果、顺承关系为边,构成一个刻画事件的发展规律和演化逻辑的有向有环图<sup>[12]</sup>。

事理图谱是知识图谱动态的延伸,重点解决Why、How问题,通过事件的多维度信息并利用高级语义关系分析技术,有效挖掘了事件序列中的内在联系和发展脉络<sup>[13-14]</sup>。目前国内外对于事理图谱有一定研究。宁慧涵等<sup>[15]</sup>提出一种顾及时空关系的事故灾难事理图谱构建方法。Liu等<sup>[16]</sup>通过事件关联识别关键风险源,构建了金融事理图谱。李诗轩等<sup>[17]</sup>针对突发公共卫生事件,使用规则模板和基于转换器的双向编码器表示的主题建模(Bidirectional Encoder Representations from Transformers Topic Modeling, BERTopic)模型,来构建支原体肺炎网络舆情的事理图谱。事理图谱可以利用事件信息和语义关系来挖掘出对应的事件,可以充分用于对危化品事故的预警。但是目前对于危化品事故的事理图谱构建工作较少,并且危化品事故的数据集有限,对于危险化学品实例图谱的构建以及研究都缺乏相关的数据支持。

为了优化事故风险识别流程,减少对人工的依赖,需要构建基于危化品事故的事理图谱,本文针对国家发布的权威性文件以及高质量、大规模的危化品语料库进行收集和清洗。而在抽取阶段中,发现现有的抽取方法存在三个问题:1)抽取危化品潜在事故事件实体时,缺少了中文显式语言学知识;2)在危化品事件实体中,缺少对实体边界的识别;3)在识别危化品事件关系时,缺少对主语和宾语的抽取能力。因此本文提出了一种基于危险化学品事故的事理图谱抽取方法,该方法融合了中文语义对实体显式识别任务进行词词分类处理,并添加位置编码实现关系识别。在抽取后本文将大量复杂的事件信息和关系进行整合并可视化,最后设计了一套基于Web的危化品潜在风险预警系统。具体而言,本文的贡献如下:1)提出了危化品潜在危险事件实体抽取模型SEMNER(Semantic W2NER);2)提出了危化品潜在危险事件关系抽取模型POS-Rel(Position Case-Rel);3)构建了危化品潜在风险事理图谱及预警系统。

## 1 危化品事故的抽取方法研究

事件的实体和关系识别都是构建事理图谱

的核心技术要素。本文分别设计了事件实体抽取模型 SEM-NER, 以及事件关系识别模型 POS-Rel。

### 1.1 事件实体抽取

为了捕获更多的分词信息和词信信息, 来增强模型对语义的感知, 本文从中文语料处理角度出发, 根据语言学信息编码器, 基于词-词关系分类的实体识别 (Word-to-Word Relation Entity Recognition, W2NER)<sup>[18]</sup> 模型, 设计了 SEM-NER 模型, 由预训练模型、卷积层、联合预测层组成, 模型结构如图 1 所示。

#### 1.1.1 预训练模型

BERT<sup>[19]</sup> (Bidirectional Encoder Representations from Transformers) 是一种基于 Transformer 的预训练语言模型, 其高效利用海量未标注文本数据, 借助双重预训练任务策略, 深化了对语言语义的理解层次。虽然 BERT 对上下文的编码很有帮助, 但是缺少中文语义信息。本文采用 LERT<sup>[20]</sup> (Linguistically-motivated bidirectional Encoder Representation from Transformer) 模型来进行预训练, 其基于 BERT, 并结合语言技术平台 (Language Technology Platform, LTP) 工具。由哈尔滨工业大学开发, 提供了一系列中文自然语言处理工具, 包括语义角色标注, 句法分析, 分词和词性标注, 依赖分析和命名实体识别等。本文使用该工具获取危化品事件句子的分词信息和词性标注信息。同时使用双向长短期记忆网络 (Long Short-Term Memory, LSTM) 对 LERT 的输出, 来捕捉序列中的长期依赖关系, 得到每个单词或字的隐层向量。将

LSTM 的输出传入条件层归一化 (Conditional Layer Normalization, CLN) 中, 条件层归一化有助于将数值范围控制在一个相对小的区间内, 提高学习速度, 同时避免了梯度消失, 提高了网络的通用性, 降低了模型的敏感度。

#### 1.1.2 卷积层

受到 BERT 模型的启发, 卷积首先对单词、位置和句子三个维度进行编码, 依次表示为  $V \in \mathbb{R}^{N \times N \times d_{E_d}}$ 、 $E^d \in \mathbb{R}^{N \times N \times d_{E_d}}$  和  $E^t \in \mathbb{R}^{N \times N \times d_{E_t}}$ 。将三种嵌入拼接, 并用全连接网络进行特征混合, 同时降低维度, 得到词对表示为  $C \in \mathbb{R}^{N \times N \times d_c}$  个计算过程如式 (1) 所示。

$$C = \text{MLP}_1([V; E^d; E^t]), \quad (1)$$

$$Q^l = \sigma(\text{DConv}_l(C)). \quad (2)$$

接下来, 采用多个不同扩张率的二维膨胀卷积 (Dilated Convolutions, DConv), 用于词对之间的交互。计算过程如式 (2)。其中  $Q^l \in \mathbb{R}^{N \times N \times d_c}$  扩张率为  $l$  的扩张卷积输出,  $\sigma$  为 GELU 激活函数。最终得到的词对表示为  $Q = [Q^1, Q^2, Q^3] \in \mathbb{R}^{N \times N \times 3d_c}$ 。

#### 1.1.3 联合预测层

卷积操作得到了词对表示  $Q$ , 随后采用全连接层, 来预测词对之间的关系。在此, 先前的研究表明, 在关系分类方面, 结合全连接预测器和双仿射预测器效果更佳, 因此本文采取选择两种预测分别再组合的策略。

编码器的输出  $H = \{h_1, h_2, \dots, h_N\} \in \mathbb{R}^{N \times d_h}$  仿射预测器的输入。给定词表示  $H$ , 在全连接网络之后, 分别计算出主体  $x_i$  示  $s_i$  及客体  $x_j$  的表示  $o_j$ 。然后采用双放射分类器计算词对

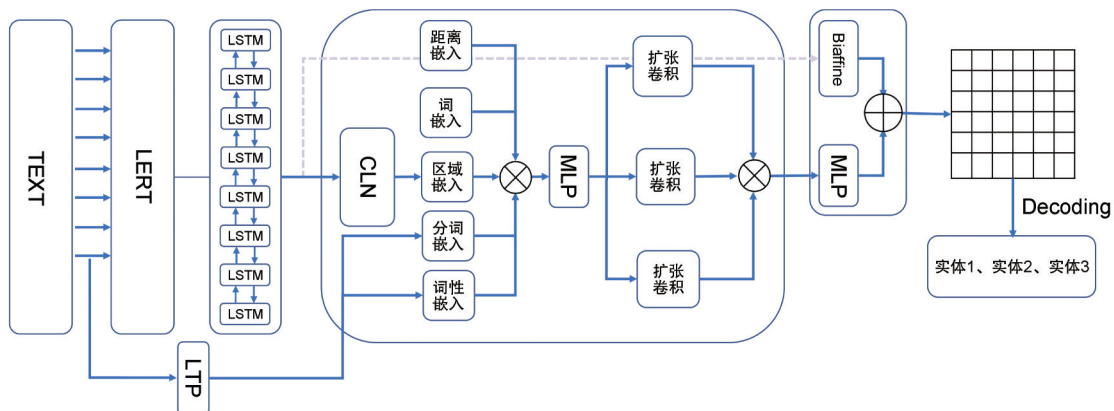


图 1 SEM-NER 模型

Fig. 1 SEM-NER model

$(x_i, x_j)$  的分数关系, 如式(3)一式(5)所示:

$$s_i = \text{MLP}_2(h_i), \quad (3)$$

$$o_j = \text{MLP}_3(h_j), \quad (4)$$

$$y'_{ij} = s_i^\top U o_j + W[s_i; o_j] + b, \quad (5)$$

其中  $U$ 、 $W$  和  $b$  均为可学习的参数,  $y'_{ij} \in \mathbb{R}^{|R|}$  为预定义关系  $R$  的分数。

用词对  $(x_i, x_j)$  的表示  $Q_{ij}$  全连接网络预测器的输入, 得到的关系分数为, 如式(6):

$$y''_{ij} = \text{MLP}(Q_{ij}), \quad (6)$$

其中  $y''_{ij} \in \mathbb{R}^{|R|}$  为预定义关系  $R$  的分数。

最终词对  $(x_i, x_j)$  的关系概率向量  $y''_{ij}$  由两个预测器 Biaffine 和 MLP 的计算结果得出, 如式(7):

$$y_{ij} = \text{softmax}(y'_{ij} + y''_{ij}). \quad (7)$$

对于每个句子  $X = \{x_1, x_2, \dots, x_N\}$ , 其目标是 minimized 负对数似然损失, 如式(8):

$$L = -\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \sum_{r=1}^{|R|} \hat{y}_{ij}^r \log y_{ij}^r, \quad (8)$$

其中  $N$  是词的数量,  $\hat{y}_{ij}^r$  是词对  $(x_i, x_j)$  的正确标注标签,  $y_{ij}^r$  为预测概率向量,  $r$  则表示预定义的关系集合  $R$  中第  $r$  个关系。

### 1.2 事件关系识别

为了获取危化品事故中的关系, 本文设计了一种实体位置嵌入信息的关系联合抽取模型 POS-Rel, 如图 2 所示, 该模型基于 CasRel<sup>[21]</sup> 模型进行优化, 不同于 CasRel, 本文已从句子中抽取实体, 对已获取的实体, 加入位置编码, 增强 CasRel 中编码器对主语和谓语的抽取关系。

#### 1.2.1 编码器

本文基于 BERT 改进了 CasRel 中编码部分。首先已经抽取的实体嵌入  $W_s$  经过 BERT 模型, 得到隐状态输出  $h_\alpha$ , 具体表示如下:

$$h_0 = S W_s + W_p, \quad (9)$$

$$h_\alpha = \text{trans}(h_{\alpha-1}), \alpha \in [1, N], \quad (10)$$

其中  $W_s$  表示实体嵌入,  $S$  表示文本的词组成的 one-hot 编码,  $\text{trans}(x)$  代表 Transformer 模块,  $h_\alpha$  代表的是第  $\alpha$  个隐状态输出。

接着需要融合实体位置信息, 本文提供了三种位置信息的编码方式, 如下所示。

1) 将实体开头位置标注 1, 该实体剩余部分位置标注 2, 不属于任意实体部分位置

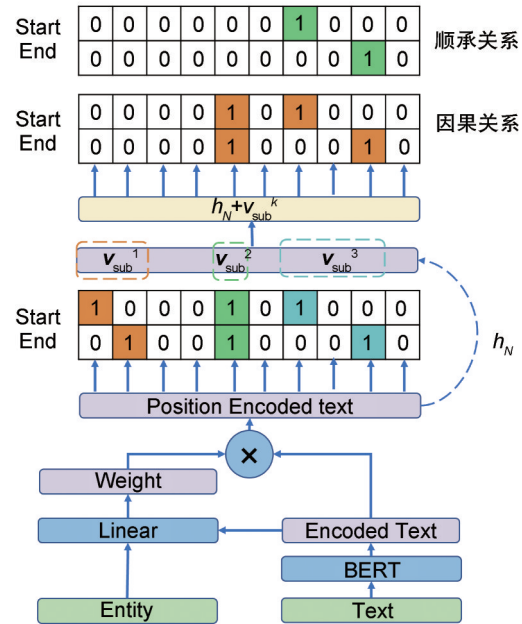


图2 POS-Rel模型

Fig. 2 POS-Rel model

标注 0。

2) 将实体所有位置标注 1, 不属于任意实体位置标注 0。

3) 将实体开头位置标注 1, 实体结束位置标注 2, 不属于任意实体位置标注 0。

对位置长度为  $N$  的张量  $P$ , 以及 BERT 输出长度同样为  $N$  的句子张量  $H$  进行拼接, 得到长度为  $2N$  的新张量  $Z$ , 通过线性 (Linear) 层, 得到长度为  $N$  的张量  $Y$ , 接着对其进行归一化处理, 最后与 BERT 的输出  $H$  逐个相乘, 得到长度为  $N$  的编码。计算过程如式(11)一式(12):

$$Y = UZ + B, \quad (11)$$

$$x_i = H_i \text{softmax}(Y_i), \quad (12)$$

其中  $U$  和  $B$  均为可学习的参数,  $Z$  为 BERT 与位置张量拼接后的新张量,  $Y$  为  $Z$  通过线性层变换后的张量,  $H_i$  为 BERT 的第  $i$  个输出,  $x_i$  为编码器的第  $i$  个输出。

#### 1.2.2 解码器

将上述编码器的输出作为解码器的输入。解码器分为主体标注器和特定关系的客体标注器, 其过程需要经过两步, 先是选取一个主语, 对其进行识别, 并对每个候选主体, 逐一排查每种关系, 查询是否有对应的客体。

在主体标注器阶段, 采用了二元标签分配, 进行 0 或者 1 的标签方式, 分别检测主体的起

始与末尾位置。若此 token 是主体的起始或者末尾索引,则给当前位置分类为 1,反之为 0。具体公式如(13)一式(14):

$$p_i^{\text{start}} = \sigma(\mathbf{W}_{\text{start}} \mathbf{x}_i + \mathbf{b}_{\text{start}}), \quad (13)$$

$$p_i^{\text{end}} = \sigma(\mathbf{W}_{\text{end}} \mathbf{x}_i + \mathbf{b}_{\text{end}}), \quad (14)$$

其中  $\mathbf{x}_i$  是第  $i$  个 token 的编码,  $p_i^{\text{start}}$  为起始索引的概率,  $p_i^{\text{end}}$  为末尾索引的概率。 $\sigma$  代表 sigmoid 激活函数。

对于给定的 token 编码结果,我们的训练目标为最大化似然函数,如式(15):

$$p_\theta(s|x) = \prod_{t \in \{\text{start}, \text{end}\}} \prod_{i=1}^L (p_i^t)^{I\{y_i^t=1\}} (1-p_i^t)^{I\{y_i^t=0\}}, \quad (15)$$

其中当且仅当  $z = \text{ture}$  时,  $I(z) = 1$ , 否则为 0。 $L$  为整个句子长度,  $y_i^{\text{start}}$  表示第  $i$  个 token 对应开始位置的二元标签(0 或者 1)。对于  $p_\theta(s|x)$ , 给定句子  $x$ , 尽可能找出句子主体。

而在特定关系的客体标注器中,需要同时识别关系和客体,每种关系都对应了各自的客体识别器,如式(16—17):

$$p_i^{\text{start}_o} = \sigma(\mathbf{W}_{\text{start}}^r (\mathbf{x}_i + \mathbf{v}_{\text{sub}}^k) + \mathbf{b}_{\text{start}}^o), \quad (16)$$

$$p_i^{\text{end}_o} = \sigma(\mathbf{W}_{\text{end}}^r (\mathbf{x}_i + \mathbf{v}_{\text{sub}}^k) + \mathbf{b}_{\text{end}}^o), \quad (17)$$

其中  $\mathbf{v}_{\text{sub}}^k$  表示第  $k$  个主体的编码表示向量。 $\mathbf{x}_i$  与  $\mathbf{v}_{\text{sub}}^k$  的维度需要保持一致。

训练目标与主体标注器一致,如式(18):

$$p_\theta(o|s, x) = \prod_{t \in \{\text{start}, \text{end}\}} \prod_{i=1}^L (p_i^t)^{I\{y_i^t=1\}} (1-p_i^t)^{I\{y_i^t=0\}}. \quad (18)$$

## 2 实验与分析

### 2.1 实验数据

本文采用自建数据集以及开放的数据集 Resume-zh。构建的危险化学品中文数据集,描述了危化品潜在危险事件。具体方法为:首先从《危险化学品目录(2015版)》和公开网站的危化品信息上采集危化品的信息,根据物质安全数据单(Material Safety Data Sheet, MSDS)信息清洗出消防信息、危险性描述和操作方法等 16 项信息,并剔除和本文研究无关的信息,再根据危险化学品事故本体框架进行人工标注。得到的标注风险信息为 995 条,据此本文训练了一个标注模型,通过此模型对剩余数据

进行处理。数据集由 6 620 条危化品潜在危险事件组成,实验中将数据集按照 7:1.5:1.5 的比例划分为训练集、验证集和测试集,其中训练集为 4 634 条,验证集为 993 条,测试集也为 993 条。

### 2.2 评价指标

实验中采用精确率(Precision,  $P$ )、召回率(Recall,  $R$ )和  $F1$  值来作为衡量事件实体和关系识别的评价指标,计算公式如下:

$$P = \frac{TP}{FP + TP}, \quad (19)$$

$$R = \frac{TP}{FN + TP}, \quad (20)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (21)$$

其中  $TP$  (True Positive) 代表真正例,  $FP$  (False Positive) 代表假正例,  $FN$  (False Negative) 代表假反例。

### 2.3 对比模型

#### 2.3.1 事件抽取

1) BERT+BiLSTM+CRF: 由 BERT 作为预训练模型, BiLSTM 用 LSTM 来获取上下文的信息, 条件随机场(Conditional Random Field, CRF)用来处理相邻标签的依赖关系。

2) BERT+BiGRU+CRF: 包含 BERT 和 CRF, BiGRU 用双向门控循环单元(Gated Recurrent Unit, GRU), 结合过去和未来的信息。

3) BERT+IDCNN+CRF: 包含 BERT 和 CRF, 以及迭代空洞卷积网络(Iterated Dilated Convolutional Neural Network, IDCNN), 其中 IDCNN 比卷积神经网络(Convolutional Neural Network, CNN)更能提取上下文信息。

4) LERT: 基于 BERT, 融入语言学特征。

5) POS (Part-of-speech Tagging): 词性标注方法。

6) CWS (Chinese Word Segment): 中文分词方法。

7) W2NER: 将命名实体识别(Named Entity Recognition, NER)问题视为词词分类问题。

#### 2.3.2 关系识别

1) Graphrel<sup>[22]</sup>: 基于图卷积网络, 联合学习命名实体和关联。

2) CopyR<sup>[23]</sup>: Seq2seq 学习框架, 采用 Copy

机制,将关系识别转化为三次生成过程。

3) CasRel: 基于参数共享的联合实体关系识别模型,解决实体重叠问题。

## 2.4 参数设置

实验参数设置如表1和表2所示。

表1 POS-Rel模型主要参数默认值及其描述

Table 1 The default values of the main parameters of the POS-Rel model and their descriptions

参数名称	默认值	描述
bert_lr	0.000 01	预训练模型学习率
other_lr	0.003	其他部分学习率
batch_size	8	批处理大小
max_epoch	200	最大训练轮数
max_len	300	序列最大长度
bert_dim	768	预训练模型的嵌入维度

表2 SEM-NER模型主要参数设定值及其描述

Table 2 The setting values of the main parameters of SEM-NER model and their descriptions

参数名	参数设定值	描述
dist_emb_size	20	距离信息词嵌入长度
type_emb_size	20	实体类型词嵌入长度
epochs	30	训练轮次
batch_size	12	批大小
learning_rate	0.001	学习率

## 2.5 实验结果

### 2.5.1 事件抽取

本文提出的SEM-NER模型与7个对比模型在危化品潜在危险事件数据集和Resume-zh数据集上的事件抽取效果如表3所示。在Resume-zh数据集上,F1达到了96.59%。可以看出SEM-NER模型比未基于词对关系抽取效果更好。同时,SEM-NER也有效地提高了W2NER模型的效果,证实融入语义信息的有效性,尤其是通过LERT提供的内容以及本文新提出的分词和词性矩阵,这些方法均能显著提升命名实体识别的性能。与原W2NER模型效果对比如图3。

W2NER:	遇水、潮湿空气或酸分解释出剧毒和自然的磷化氢气体。
	CON CON CON RES
SEM-W2NER:	遇水、潮湿空气或酸分解释出剧毒和自然的磷化氢气体。
	CON CON CON RES

图3 W2NER和SEM-NER的对比示例

Fig. 3 Example of comparison between W2NER and SEM-NER

### 2.5.2 关系识别

表4是本文的POS-Rel三种标记方法与其他三个模型的不同实验结果,其中第一种标记方法都优于对比模型,并且方案一对比其他两种方案更能提取事件的关系,精确率达到了75.86%,召回率达到了76.23%,F1达到了76.04%,相比于最初的CasRel模型,这三个指标都有不小的提升,方案一更是提升了6.28%。

表3 危化品在自建数据集和Resume-zh下不同模型的NER性能

Table 3 NER performance of different models of hazardous chemicals in self-built data set and Resume-zh %

模型	自建数据集			Resume-zh		
	精确率	召回率	F1	精确率	召回率	F1
BERT+BiL-STM+CRF	90.66	89.00	89.82	96.19	91.85	93.97
BERT+BiG-RU+CRF	85.38	85.11	85.24	92.84	93.69	93.26
BERTH-DCNN+CRF	82.45	80.18	81.30	80.35	84.86	82.54
-AII(W2NER)	88.94	90.74	89.83	95.65	95.71	95.68
-POS	91.34	90.49	90.91	95.68	96.38	96.03
-CWS	<b>94.04</b>	87.92	90.88	95.84	96.01	95.92
-LERT	92.39	89.68	91.01	95.45	<b>96.63</b>	96.04
SEM-NER	93.33	<b>92.50</b>	<b>92.91</b>	<b>96.74</b>	96.44	<b>96.59</b>

表4 不同模型事件关系识别实验结果

Table 4 Entification of experimental results of different model event relationships %

模型	精确率	召回率	F1
Graphrel	48.73	47.52	48.12
CopyR	55.78	54.95	55.36
CasRel	68.75	70.79	69.76
POS-Rel <sub>tag1</sub>	<b>75.86</b>	<b>76.23</b>	<b>76.04</b>
POS-Rel <sub>tag2</sub>	68.54	72.27	70.36
POS-Rel <sub>tag3</sub>	74.74	73.26	73.99

## 3 危险化学品事故的事理图谱构建

本文方法构建的危险化学品事故事理图谱,可以十分直观地展示危险化学品事故的发展过程,可视化整个事件过程,并且搭建了一个基于事理图谱的预警系统。

### 3.1 事理图谱构建

根据本文提出的危险化学品事故的事理图谱构建方法,首先从《危险化学品目录(2015版)》和公开网站的危化品信息上采集危化品



## 危险化学品预警系统

仓库内所包含的危化品:

仓库状态:

解析结果:

分析潜在危险

图5 预警系统运行图

Fig. 5 Operational diagram of the early warning system

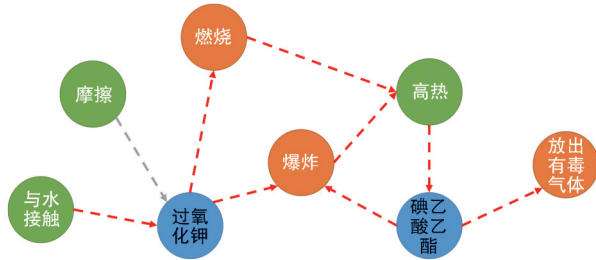


图6 预警链路示意图

Fig. 6 Schematic diagram of the early warning link

中还可以结合危化品事故发生的图像、视频等信息,深入研究危化品事故的多模态事理图谱构建方法;深入挖掘潜在的危化品事故,进行事理图谱的补充,完善整个危化品事故的事理图谱,提高预警水平。

### 参考文献:

- [1] 国务院天津港"8·12"瑞海公司危险品仓库特别重大火灾爆炸事故调查组. 天津港"8·12"瑞海公司危险品仓库特别重大火灾爆炸事故调查报告[R/OL]. (2016-02-05) [2025-09-15]. [http://www.gov.cn/foot/2016-02/05/content\\_5039788.htm](http://www.gov.cn/foot/2016-02/05/content_5039788.htm).
- [2] 陈文坚. 化学危险品储运项目环境风险评估实例[J]. 福建环境, 2003(5): 27-30.  
CHEN W J. A Case of the Environmental Risk Assessment of the Storage and Transportation of Chemical Hazardous Articles[J]. *Fujian Environ*, 2003(5): 27-30.
- [3] YUAN C F, WANG H, CHEN Y. The Hierarchical Risk Source Identification Method Connected with Event Causal Chain in the Emergency Process of Fire Accident of Petroleum Storage and Transportation[J]. *Appl Mech Mater*, 2014, **501**: 2411-2414. DOI:10.4028/www.scien-
- [4] 王好一, 赵东风, 孟亦飞, 等. 基于前导事件和贝叶斯网络的储罐溢油动态风险评估[J]. 安全与环境学报, 2018, **18**(2): 446-450. DOI: 10.13637/j.issn.1009-6094.2018.02.007.  
WANG H Y, ZHAO D F, MENG Y F, *et al.* Dynamic Risk Analysis Model for the Oil-container Overflow Scenario Based on the Bayesian Network and Precursor Events[J]. *J Saf Environ*, 2018, **18**(2): 446-450. DOI: 10.13637/j.issn.1009-6094.2018.02.007.
- [5] JE C H, STONE R, OBERG S G. Development and Application of a Multi-channel Monitoring System for near Real-time VOC Measurement in a Hazardous Waste Management Facility[J]. *Sci Total Environ*, 2007, **382**(2/3): 364-374. DOI:10.1016/j.scitotenv.2007.04.017.
- [6] 贾进章, 陈怡诺. 基于网络层次分析-灰色聚类法的高层建筑火灾风险分析[J]. 安全与环境学报, 2020, **20**(4): 1228-1235. DOI: 10.13637/j.issn.1009-6094.2019.0751.  
JIA J Z, CHEN Y N. Fire Risk Assessment of the High-rise Buildings Based on the Network Hierarchical Analysis-gray Clustering Method[J]. *J Saf Environ*, 2020, **20**(4): 1228-1235. DOI: 10.13637/j.issn.1009-6094.2019.0751.
- [7] 马全宇, 张茂盛, 丁宗超, 等. 基于深度学习的海关危险化学品数智监管新模式及应用[J]. 中国口岸科学技术, 2024, **6**(4): 41-48. DOI: 10.3969/j.issn.1002-4689.2024.04.008.  
MA Q Y, ZHANG M S, DING Z C, *et al.* The Building and Application of a New Model of Digital Intelligence Supervision and Control of Customs Hazardous Chemicals Based on Deep Learning[J]. *China Port Sci Technol*, 2024, **6**(4): 41-48. DOI: 10.3969/j.issn.1002-4689.2024.04.008.
- [8] SONG Y X, SUN P L, LIU H Y, *et al.* Scene-driven Multimodal Knowledge Graph Construction for Embodied AI [J]. *IEEE Trans Knowl Data Eng*, 2024, **36**(11): 6962-

6976. DOI:10.1109/TKDE.2024.3399746.
- [9] FAN Z L, CHEN C C. CuPe-KG: Cultural Perspective-Based Knowledge Graph Construction of Tourism Resources via Pretrained Language Models[J]. *Inf Process Manag*, 2024, **61**(3): 103646. DOI: 10.1016/j.ipm.2024.103646.
- [10] LIU Z K, LU Y Q. A Task-centric Knowledge Graph Construction Method Based on Multi-modal Representation Learning for Industrial Maintenance Automation [J]. *Eng Rep*, 2024, **6**(12): e12952. DOI: 10.1002/eng2.12952.
- [11] GAO J L, PENG P, LU F, *et al.* Mining Tourist Preferences and Decision Support via Tourism-oriented Knowledge Graph[J]. *Inf Process Manag*, 2024, **61**(1): 103523. DOI:10.1016/j.ipm.2023.103523.
- [12] LIU T. From Knowledge Graph to Event Evolutionary Graph[R]. Shanghai: CCF YOCSEF, 2017.
- [13] SUN X K, MENG Y, WANG W L. Identifying Traffic Events from Weibo with Knowledge Graph and Target Detection[J]. *Data Analysis and Knowledge Discovery*, **4**(12): 136-147. DOI: 10.11925/infotech. 2096-3467.2020.0596.
- [14] 李纲, 王施运, 毛进, 等. 面向态势感知的国家安全事件图谱构建研究[J]. *情报学报*, 2021, **40**(11): 1164-1175. DOI: 10.3772/j.issn.1000-0135.2021.11.004.  
LI G, WANG S Y, MAO J, *et al.* Construction of National Security Event Map and Its Application for Situation Awareness[J]. *J China Soc Sci Tech Inf*, 2021, **40**(11): 1164-1175. DOI: 10.3772/j.issn.1000-0135.2021.11.004.
- [15] 宁慧涵, 眭海刚, 王金地, 等. 顾及时空关系的事灾难事理图谱构建方法研究[J]. *武汉大学学报(信息科学版)*, 2024, **49**(5): 831-843. DOI: 10.13203/j.whugis20230291.  
NING H H, SUI H G, WANG J D, *et al.* *Geomat Inf Sci Wuhan Univ*, 2024, **49**(5): 831-843. DOI: 10.13203/j.whugis20230291.
- [16] LIU Z H, ZHANG Z J, ZENG X. Risk Identification and Management Through Knowledge Association: A Financial Event Evolution Knowledge Graph Approach [J]. *Expert Syst Appl*, 2024, **252**: 123999. DOI:10.1016/j.eswa.2024.123999.
- [17] 李诗轩, 沈愿, 王璐, 等. 面向支原体肺炎网络舆情的  
事理图谱构建与演化分析[J/OL]. *情报科学*, 2024: 1-22. (2024-09-12). <https://kns.cnki.net/kcms/detail/22.1264.G2.20240624.1801.002.html>.
- LI S X, SHEN Y, WANG L, *et al.* Construction and Evolution Analysis of the Event Map for *Mycoplasma Pneumonia* Network Public Opinion[J/OL]. *Inf Sci*, 2024: 1-22. (2024-09-12). <https://kns.cnki.net/kcms/detail/22.1264.G2.20240624.1801.002.html>.
- [18] LI J Y, FEI H, LIU J, *et al.* Unified Named Entity Recognition as Word-word Relation Classification[J]. *Proc AAAI Conf Artif Intell*, 2022, **36**(10): 10965-10973. DOI:10.1609/aaai.v36i10.21344.
- [19] DEVLIN J, CHANG M W, LEE K, *et al.* 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). Stroudsburg, PA: ACL, 2019: 4171-4186. DOI: 10.18653/v1/N19-1423.
- [20] CUI Y M, CHE W X, WANG S J, *et al.* Lert: A Linguistically-motivated Pre-trained Language Model [EB/OL]. (2022-11-11) [2025-09-08]. <https://doi.org/10.48550/arXiv.2211.05344>.
- [21] WEI Z P, SU J L, 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. Stroudsburg, PA, USA: ACL, 2020: 1476-1488. DOI:10.18653/v1/2020.acl-main.136.
- [22] 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. Stroudsburg, PA, USA: ACL, 2019: 1409-1418. DOI:10.18653/v1/p19-1136.
- [23] ZENG X R, ZENG D J, HE S Z, *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). Stroudsburg, PA: USA: ACL, 2018: 506-514. DOI:10.18653/v1/p18-1047.