·人工智能及识别技术· 文章编号：1000-3428(2019)11-0213-05 文献标志码：A 中图分类号：TP391

基于局部实体特征的事件触发词抽取

柳亦婷，李培峰

（苏州大学 计算机科学与技术学院，江苏 苏州 215000）

摘 要： 基于神经网络的触发词抽取模型利用实体信息判别触发词，但大量无关实体会影响触发词抽取效果。提出一种借助局部实体特征的事件触发词抽取方法，该方法先初步过滤无关实体，并将保留实体分为核心与非核心2类分别进行建模。利用卷积神经网络（CNN）抽取局部特征的特性，从众多实体中定位有助于触发词识别的局部重要实体，采用注意力机制提高其权重，同时利用有效非核心实体的语义排除干扰实体，从而借助重要实体的特征信息判别触发词。在特定和通用领域事件语料库上的实验结果均表明，该方法能够减少无关实体对触发词抽取的干扰，其触发词抽取性能的 *F*1 值比基准系统最高可提升 0.017。

关键词： 事件抽取；触发词抽取；CNN 模型；局部实体特征；核心实体；非核心实体

中文引用格式： 柳亦婷，李培峰. 基于局部实体特征的事件触发词抽取[J]. 计算机工程，2019，45(11)：213-217，224.

英文引用格式： LIU Yiting，LI Peifeng. Event trigger extraction based on local entity characteristics[J]. Computer Engineering，2019，45(11)：213-217，224.

Event Trigger Extraction Based on Local Entity Characteristics

LIU Yiting，LI Peifeng

（School of Computer Science and Technology，Soochow University，Suzhou，Jiangsu 215000，China）

【Abstract】 The current neural network-based event trigger extraction model distinguishes triggers by entity information，but lots of unrelated entities have negative interference on trigger extraction. Therefore，we propose an event trigger extraction method based on local entity characteristics. First，we filter unrelated entities，separate the reserved entities into core and non-core types，and build the reserved entity model. Then，with the utilization of CNN network，which is superior in the extraction of local features，the model manages to locate the entities that are important for trigger recognition from all the entities. Subsequently，we adopt the Attention mechanism to improve the weights of important entities，so as to identify triggers by their characteristic information. Experimental results on domain-specific and domain-general event corpus show that the proposed method can reduce the interference of unrelated entities on trigger extraction，and its *F*1-score of trigger extraction performance，if at the highest level，can exceed the baseline by 0.017.

【Key words】 event extraction；trigger extraction；CNN model；local entity characteristics；core entities；non-core entities

DOI： 10.19678/j.issn.1000-3428.0051052

0 概述

事件抽取是指从无结构化的文本中抽取结构化的事件信息，如事件触发词及其类型、事件论元及其角色。本文以例句“Next week，Hurtado will discuss details of this collaboration with officials in Peru.”阐释事件抽取任务。上述例句包含“Next week”“Hurtado”“officials in Peru”3 个实体实例，其类型分别是 TIME（时间）、PER（人）、PER（人）。从该例句中抽取出一个事件实例，其类别是“会谈”，触发词是“discuss”。“Next week”“Hurtado”“officials in Peru”是该事件的 3 个实体论元，分别担任的角色是 Time（时间）、Actor0（施事者）、Actor0（施事者）。

事件抽取包含触发词抽取和论元抽取 2 个子任务。其中，触发词抽取任务就是抽取文档中的事件触发词并预测触发词所属的事件类别。在触发词抽取中，实体信息是重要依据。当前，部分基于神经网络的触发词抽取模型在利用实体信息时，没有排除无关实体对触发词抽取的干扰。例如事件抽取：“NASA is launching a series of small suborbital rockets，called sounding rockets，to study the Earth and space，according to CNN.”从该例句中抽取到一个“发射”类事件，触发词为“launching”。“NASA”与“CNN”同为 ORG（组织）类型实体。其中，“NASA”

基金项目： 国家自然科学基金（61772354，61773276，61472265）。

作者简介： 柳亦婷（1991—），女，硕士研究生，主研方向为事件抽取；李培峰，教授、博士。

收稿日期： 2018-04-02 **修回日期：** 2018-09-17 **E-mail：** lytliu0719@163.com

计算机工程 www.ecice06.com

在该事件中担任 Actor0(施事者)论元,“CNN”在该事件中不担任论元,是无关实体。显然,对于触发词“launching”的识别而言,“NASA”是有用的实体,“CNN”反而可能造成干扰。如何排除无关实体的干扰是触发词抽取过程中需要考虑的问题。

受触发词与论元联合抽取模型的启发,本文提出一个事件触发词抽取方法,首先筛选出触发词周边的有效实体,再利用卷积神经网络(Convolutional Neural Networks,CNN)模型挖掘出有效实体中最重要的局部实体信息,从而借助局部实体信息实现触发词识别。此外,本文构建面向特定领域的时政事件(Political Event,PE)语料库,对比本文方法在 PE 语料库和 ACE 语料库上的实验结果。

1 相关工作

在句子级别的事件抽取上,传统方法采用规则方法或基于分类器的特征工程方法。文献[1]引入最大熵分类器,将事件抽取看作是分类问题。文献[2]提出管道模型,将事件抽取任务分为若干串行的子任务。文献[3]在基于结构化感知机的基础上,利用多种局部特征和全局特征,提出触发词与论元抽取的联合模型,以减少串行抽取模式造成的级联错误。

近年来,部分事件抽取方法采用了神经网络方法。文献[4]将神经网络模型应用于事件抽取,输入候选触发词的上下文信息、周边的实体信息,利用 CNN 模型自动挖掘隐含特征,从而避免了特征工程中词性标注、句法分析等预处理步骤错误对分类结果的影响。文献[5]提出 DMCNN 网络,通过在 CNN 网络的基础上构建一个动态多池化层来保留更多的有效信息。文献[6]利用循环神经网络(Recurrent Neural Network,RNN)实现触发词与论元的联合抽取模型,并探究了多种不同的记忆特征。文献[7]提出 HNN 模型,即把 CNN 模型与 Bi-LSTM 模型进行融合,同时挖掘触发词上下文的局部特征和全局特征。文献[8]在测试集论元未知和论元已知的条件下,分别利用多种有监督和无监督 Attention 机制挖掘实体信息,研究其对触发词抽取产生的不同作用。

在文档级和篇章级的事件抽取上,文献[9]使用隐马尔科夫模型(Hidden Markov Model,HMM)和条件随机场(Conditional Random Field,CRF)等序列模型获取全局信息。文献[10]提出了篇章树结构和局部依存关系的框架从而引入全局信息。文献[11]提出跨文档的事件抽取方法,从相近主题的文档中抽取事件信息。文献[12]利用上下文中的相关事件信息和共指事件信息填补时间类成员的缺失。文献[13]提出单文档中的跨事件推理方法,将文档中的其他相关事件信息传播到当前事件。文献[14]提出跨实体抽取方法,利用特定实体总是出现在特定类别事件中的特性,对实体进行聚类,根据实体的类别判断触发词的类别。

2 模型构建

本文将论元分为核心论元和非核心论元,将填充相应论元的实体类型分为核心实体和非核心实体,根据核心与非核心实体各自的特点,分别构建模型挖掘特征,从而提升触发词的识别效率。

2.1 核心实体和非核心实体的筛选

参考文献[15]的定义,本文将论元角色分类为核心角色和非核心角色。其中,核心角色是事件的参与者角色,回答了事件要素中最核心的 2W(Who,Whom)问题。非核心角色是辅助角色,主要包括职位、武器、金额等角色。而时间和地点角色不能为触发词的识别提供较多有益信息,因此,本文不对这两者划分。

在此基础上,本文将能够填充核心角色的实体类型定义为核心类型,能够填充非核心角色的实体类型定义为非核心类型。属于核心类型的实体实例称为核心实体,属于非核心类型的实体实例称为非核心实体。

对于候选触发词而言,在上下文中的核心与非核心实体中可能存在干扰实体。例如 NUM 类型的实体可能是“经费”和“合同”类事件的 Amount(金额)论元,但在“会谈”类事件中则是干扰实体。本文在预处理阶段对部分明显的干扰实体进行初步筛除,筛选方法为:1)遍历训练集,构建“触发词-事件类别”映射表;2)对于测试集中的候选触发词,根据“触发词-事件类别”映射表查找该候选触发词对应的所有可能事件类别;3)根据事件体系定义查找这些事件类别论元的核心、非核心角色集和对应实体的核心、非核心类型集;4)根据核心、非核心类型集从候选触发词上下文中筛选出对应的核心、非核心实体集。

用以上筛选方法筛选出的核心实体和非核心实体分别称为有效核心实体和有效非核心实体。

例 1 Boeing today announced that the UK Ministry of Defense has signed an agreement for the acquisition of one C-17 Globemaster III.

在例 1 中共有 4 个实体“Boeing”“today”“the UK Ministry of Defense”“C-17 Globemaster III”,类型分别为 ORG、TIME、ORG 以及 WEA。遍历训练集可知,候选触发词“acquisition”可能触发“购入”类事件。该事件类别对应的核心与非核心角色集分别为{Actor0,Actor1}、{Weapon},对应的核心与非核心类型集为{PER,ORG,GPE}、{WEA}。则有效核心实体集合 p = {“Boeing”,“the UK Ministry of Defense”},有效非核心实体集合 q = {“C-17 Globemaster III”}。

2.2 模型输入

本文对预处理阶段筛选出的有效实体,分别用

www.ecaeo.com

CNN 网络构建基本模型、核心实体子模型和非核心实体子模型,将这 3 种模型在 Pooling 层的输出进行拼接,最后通过 Softmax 层分类。以例 1 中的候选触发词“acquisition”为例,系统模型如图 1 所示。

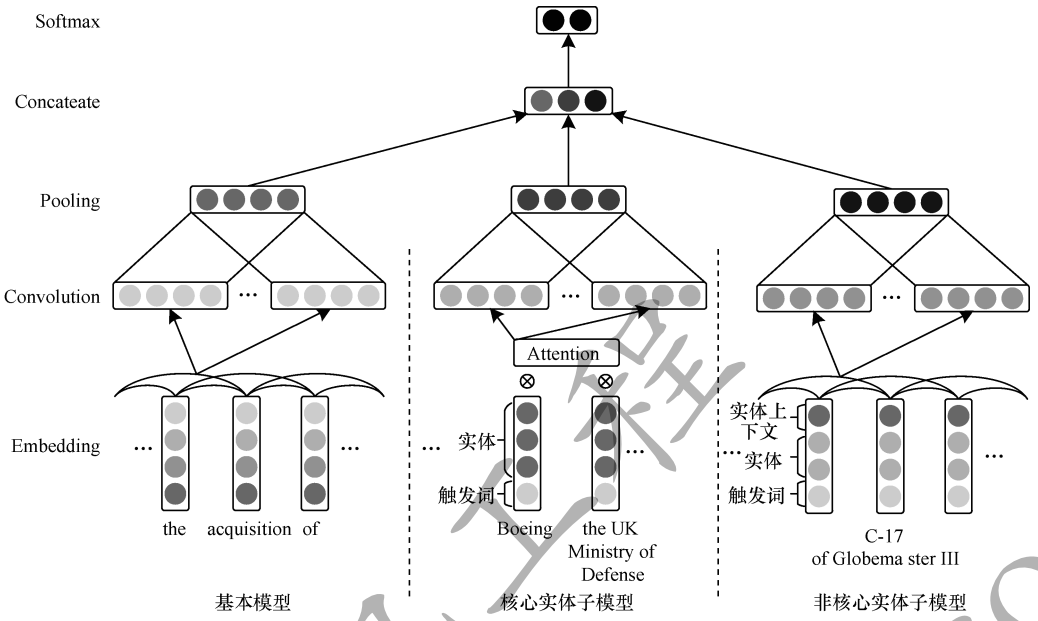

图 1　事件触发词抽取系统模型

基本模型、核心实体子模型和非核心实体子模型,介绍如下:

1)基本模型

为了挖掘触发词周围的上下文信息,本文采用文献[4]提出的 CNN 模型作为基本模型。

基本模型取候选触发词前后大小为 $2n+1$ 的词窗口作为输入。若候选触发词为 x_0,则窗口为[x_{-n}, x_{-n+1},…,x_0,…,x_{n-1},x_n]。窗口中每个词 x_i($-n \leqslant i \leqslant n$)的输入向量为:

$$\boldsymbol{V}_{1i} = \boldsymbol{V}_{w_1i} \oplus \boldsymbol{V}_{p_1i} \oplus \boldsymbol{V}_{e_1i}$$

其中,$\boldsymbol{V}_{1i}$为词 x_i在基本模型中的输入向量,$\boldsymbol{V}_{w_1i}$为词 x_i的词向量,$\boldsymbol{V}_{p_1i}$为词 x_i的位置向量,$\boldsymbol{V}_{e_1i}$为词 x_i的实体类型向量,⊕表示拼接操作。

2)核心实体子模型

候选触发词周边可能存在多个有效核心实体,而这些有效核心实体中可能只有一个或两个是当前候选触发词的核心论元(如果当前候选触发词的确是真实触发词)。为了能够从这些实体中挖掘出最重要的、最可能填充核心论元槽的实体,本文构建核心实体子模型。该模型以触发词周边的若干个有效核心实体作为输入矩阵,通过 Attention 层为每个有效核心实体分配重要度权重,再对 Attention 层的输出进行卷积。而 CNN 模型作为一种挖掘局部特征的模型,其最大池化层可以在若干有效核心实体中捕捉到最重要实体周围的局部特征,而忽略其他不重要的实体。这个最重要的局部实体与候选触发词的组合特征能够更高效地对候选触发词进行分类。

核心实体子模型在候选触发词前后的实体中选取了一个大小为 $2k$ 的窗口作为输入,窗口中的实体为[e_{-k}, e_{-k+1}, …, e_{-1}, e_1, …, e_{k-1}, e_k]。其中,e_i($-k \leqslant i \leqslant -1$)为候选触发词向左第 $|i|$ 个实体,e_i($1 \leqslant i \leqslant k$)为候选触发词向右第 i 个实体。窗口中每个实体 e_i的输入向量为:

$$\boldsymbol{V}_{2i} = \boldsymbol{V}_{tw_2i} \oplus \boldsymbol{V}_{e_2i} \oplus \boldsymbol{V}_{pr_2i} \oplus \boldsymbol{V}_{d_2i} \oplus \boldsymbol{V}_{i_2i} \oplus \boldsymbol{V}_{ew_2i}$$

其中,$\boldsymbol{V}_{2i}$为实体 e_i在核心实体子模型中的输入向量,$\boldsymbol{V}_{tw_2i}$为候选触发词的词向量,$\boldsymbol{V}_{e_2i}$为 e_i的实体类型向量,$\boldsymbol{V}_{pr_2i}$为 e_i与触发词的位置关系向量,$\boldsymbol{V}_{d_2i}$为 e_i与触发词的距离向量,$\boldsymbol{V}_{i_2i}$为 e_i的序号向量,$\boldsymbol{V}_{ew_2i}$为 e_i的词向量。

3)非核心实体子模型

非核心实体与核心实体的不同之处在于,非核心实体数目较少,其中的干扰实体也较少。统计语料可知,在所有距离触发词最近的非核心实体中,有 82.14% 可作为该触发词的论元;在距离触发词最近的核心实体中,仅有 39.15% 可作为该触发词的论元。这说明对于触发词来说,距离它最近的非核心实体非常重要。为了获得非核心实体的上下文信息,挖掘该实体的语义,进一步帮助判断该候选触发词的语义,本文构建非核心实体子模型。该模型选择一个距离触发词最近的有效非核心实体,以该实体的上下文信息作为输入,通过 CNN 模型挖掘其语义,从而实现对触发词的判别。

非核心实体子模型为每个非核心角色都选取了一个离候选触发词最近的相应类型非核心实体来填充该角色槽,并且为每个角色槽中的实体构建一个 CNN 网络。若当前论元槽中的实体为 E_i($1 \leqslant i \leqslant 4$),

计算机工程 www.ecice06.com

将 E_i 看作整体，选取 E_i 前后大小为 $2j+1$ 的词窗口，则窗口为 $[x_{-j},x_{-j+1},\cdots,x_0,\cdots,x_{j-1},x_j]$，其中 x_0 是实体 E_i 的第一个词，窗口中每个词 $x_k(-j\leqslant k\leqslant j)$ 在模型中的输入向量为：

$$\boldsymbol{V}_{3i}=\boldsymbol{V}_{\text{tw_3}i}\oplus\boldsymbol{V}_{\text{e_3}i}\oplus\boldsymbol{V}_{\text{pr_3}i}\oplus\boldsymbol{V}_{\text{d_3}i}\oplus\boldsymbol{V}_{\text{w_3}i}\oplus\boldsymbol{V}_{\text{p_3}i}$$

其中，$\boldsymbol{V}_{3i}$ 为词 x_k 在非核心实体子模型中的输入向量，$\boldsymbol{V}_{\text{tw_3}i}$ 为候选触发词的词向量，$\boldsymbol{V}_{\text{e_3}i}$ 为 E_i 的实体类型向量，$\boldsymbol{V}_{\text{pr_3}i}$ 为 E_i 与触发词的位置关系向量，$\boldsymbol{V}_{\text{d_3}i}$ 为 E_i 与触发词的距离向量，$\boldsymbol{V}_{\text{w_3}i}$ 为词 x_k 的词向量。

4）向量表

上述输入向量可以从以下向量表中查询得到：

（1）词向量表，以谷歌预训练的 Word2Vector[16] 词向量表进行初始化。$\boldsymbol{V}_{\text{w_1}}$、$\boldsymbol{V}_{\text{tw_2}}$、$\boldsymbol{V}_{\text{ew_2}}$、$\boldsymbol{V}_{\text{tw_3}}$、$\boldsymbol{V}_{\text{w_3}}$ 从该向量表中查询获得。

（2）实体类型向量表，随机初始化。$\boldsymbol{V}_{\text{e_1}}$、$\boldsymbol{V}_{\text{e_2}}$、$\boldsymbol{V}_{\text{e_3}}$ 从该向量表中查询获得。

（3）位置向量表，用以表示上下文词在窗口中的相对位置，随机初始化。$\boldsymbol{V}_{\text{p_1}}$、$\boldsymbol{V}_{\text{p_3}}$ 从该向量表中查询获得；

（4）位置关系向量表，用以表示触发词与实体的位置关系，共有 3 个值："1"表示实体在候选触发词左侧，"2"表示实体在候选触发词右侧，"3"表示实体与候选触发词相交，随机初始化。$\boldsymbol{V}_{\text{pr_2}}$、$\boldsymbol{V}_{\text{pr_3}}$ 从该向量表中查询获得。

（5）距离向量表，用以表示触发词与实体的距离，随机初始化。$\boldsymbol{V}_{\text{d_2}}$、$\boldsymbol{V}_{\text{d_3}}$ 从该向量表中查询获得。

（6）实体序号向量表，用以表示核心实体子模型中，实体在输入窗口中的序号，随机初始化。$\boldsymbol{V}_{\text{i_2}}$ 从该向量表中获得。

2.3 模型输出

在非核心实体子模型中，若存在 n 个非核心角色，则选取 n 个非核心实体，为这些非核心实体构建的 n 个 CNN 子网络在 Pooling 层的输出向量分别为 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$，则非核心实体子模型在 Pooling 层的输出向量为：

$$\boldsymbol{O}_3=\boldsymbol{X}_1\oplus\boldsymbol{X}_2\oplus\cdots\oplus\boldsymbol{X}_n$$

对基本模型、核心实体子模型、非核心实体子模型在 Pooing 层的输出向量进行拼接，通过 Softmax 变换获得模型的最终输出，即模型整体的输出向量为：

$$\boldsymbol{O}=Softmax(\boldsymbol{O}_1\oplus\boldsymbol{O}_2\oplus\boldsymbol{O}_3)$$

其中，$\boldsymbol{O}_1$、$\boldsymbol{O}_2$、$\boldsymbol{O}_3$ 分别为基本模型、核心实体子模型、非核心实体子模型在 Pooling 层的输出向量。

3 实验与结果分析

3.1 语料和实验设置

为了验证本文方法在特定领域语料上的实验结果，本文构建 PE 语料库。PE 与对比语料库 ACE 的统计数据如表 1 所示。

表 1 PE 与 ACE 语料库对比

语料库名称	事件类型数	文档数	事件数
PE	36	531	5 624
ACE	33	599	5 349

在 PE 语料库中，本文定义了核心论元角色 2 种，非核心角色 4 种。在 ACE 语料库中，定义了核心角色 11 种，非核心角色 5 种。

本文使用文献[4]的基本 CNN 模型作为基准系统，实验使用五折交叉验证，在训练集语料中选取 10% 作为开发集。

在实验中，基本模型窗口大小设置为 31，卷积核大小为{2,3,4,5}，卷积核数目为 150；核心实体子模型的实体窗口大小为 4，卷积核大小为{2,3}，卷积核数目为 100，实体类型向量维度为 50，位置关系向量维度为 10，距离向量维度为 70，实体序号向量维度为 20；非核心实体子模型的实体上下文窗口大小为 11，卷积核大小为{2,3}，卷积核数目为 100，实体类型向量维度为 50，位置关系向量维度为 10，距离向量维度为 70。Mini-batch size 为 50，优化算法使用 Adadelta。

3.2 实验结果

在 PE 语料库上的抽取结果如表 2 所示。其中，基本模型 + 核心子模型指 Pooling 层仅拼接了基本模型的输出和核心实体子模型输出的结果，基本模型 + 核心子模型 + 非核心子模型指 Pooling 层拼接了基本模型的输出、核心实体子模型输出和非核心实体子模型输出的结果；P、R、$F1$ 分别是模型的分类精度、召回率和 $F1$ 值。

表 2 在 PE 语料库上的抽取结果

模型	P	R	$F1$
基准系统	0.713	0.642	0.675
基本模型 + 核心子模型	0.732	0.635	0.681
基本模型 + 核心子模型 + 非核心子模型（未筛选）	0.736	0.638	0.684
基本模型 + 核心子模型 + 非核心子模型（已筛选）	0.729	0.658	0.692

从表 2 可以看出，在未预先筛选实体情况下，加入核心子模型后相比基本模型，其精度和 $F1$ 值分别提高升了 0.019 和 0.006。这是由于核心实体子模型剔除了"launch + The authorities"这样显然不是"触发词-实体"组合，排除了不能作为事件论元的无关实体对触发词识别造成的干扰。在基本模型 + 核心子模型基础上加入非核心实体子模型，其精度、召回率和 $F1$ 值分别提升了 0.004、0.003、0.003。这表明非核心实体子模型对确定触发词语义有帮助，但是由于非核心实体较少且没有经过筛选，部分不含有某种非核心论元槽的事件类型也被预填充了论

计算机工程 www.ecice06.com

元,因而整体提升幅度较小。

在对实体进行筛选后,相比未筛选情况,其 *F*1 值提升了 0.008,召回率有较明显的提升,但精度反而下降。非核心实体的存在相当于给触发词做了一个明显的标记,因而可以使得触发词的召回率上升。另外,由于部分事件的上下文中原本非核心论元和非核心实体较少,经过实体筛选后留下了更稀少的有效实体,对于部分触发词来说,句子中甚至不含有非核心实体,因而难以和那些也不含有非核心实体的候选触发词负例相区分。

在 ACE 语料上的实验使用了文献[15]对核心论元和非核心论元的划分,抽取结果如表 3 所示。与特定领域不同的是,通用领域中不同事件的核心论元普遍对应不同的实体类型,因此,本文对通用领域的核心论元也做了有效实体的筛选。

表 3　在 ACE 语料库上的抽取结果

模型	*P*	*R*	*F*1
基准系统	0.720	0.660	0.689
基本模型 + 核心子模型(未筛选)	0.767	0.642	0.699
基本模型 + 核心子模型(已筛选)	0.776	0.645	0.704
基本模型 + 核心子模型 + 非核心子模型(已筛选)	0.765	0.668	0.713

从表 3 可以看出,加入未筛选的核心实体子模型后,召回率降低,*F*1 值和精度分别提升了 0.010 和 0.047。进行实体筛选后,精度又提升了 0.009。这说明核心实体子模型能够排除部分不符合语义的“触发词-实体”组合。对有效实体进行筛选能够进一步筛除更多无效实体。加入已筛选的非核心实体子模型后,*F*1 值和召回率都有了较为明显的提升,其原因与特定领域上的提升类似。

表 2 和表 3 实验结果表明,本文方法在特定领域和通用领域事件语料上都能有效提升识别性能。但是,本文的核心实体子模型仍然不能区分 Actor0 论元和 Actor1 论元。此外,在同时出现多个重要论元且它们相隔较远的情况下,本文的核心实体子模型只能定位其中一个重要论元。例如“NASA signed a contract at Washington with its partner SpaceX……”,对于触发词“contract”来说,“NASA”与“SpaceX”是它的核心实体,“Washington”虽然是它的 Place 论元,但并不是它的核心实体,对于触发词的识别并没有帮助。由于本文 CNN 模型的卷积核是连续的,因此在最大池化层,或是挖掘出“NASA”周围的局部实体特征,或是挖掘出“SpaceX”周围的局部实体特征,不能在跳过“Washington”的情况下同时挖掘“NASA”与“SpaceX”的局部实体特征。这就意味着:当事件中的 Actor0 论元槽可填充多个实体时,一般只能大致定位出其中一个实体的局部特征;或者当事件中同时出现 Actor0 论元和 Actor1 论元时,只能大致定位出其中一个论元的某一个实体的特征。这样显然造成了信息的浪费。如何在模型中同时挖掘出多个重要实体的局部特征是下一步工作的重点。

4　结束语

针对现有基于 CNN 模型的触发词抽取方法不能筛选出重要实体的问题,本文对实体进行筛选,分别对核心实体和非核心实体建模,利用注意力机制和 CNN 抽取重要实体的局部特征,并根据重要实体与触发词的组合信息进行分类。基于 PE 和 ACE 语料库的实验结果表明,相比文献[4]的基本 CNN 模型,本文方法可以减少无关实体的干扰,显著提升识别性能。下一步将完善实体筛选方法,利用词法、句法信息对本文筛选出触发词前后窗口内的核心实体进一步过滤,同时考虑选择句法树上最近的实体或者根据实体与触发词的位置关系选取非核心实体。此外,还将尝试利用 LSTM 模型、BILSTM 模型挖掘实体特征。

参考文献

[1] CHIEU H L,NG H T. A maximum entropy approach to information extraction from semi-structured and free text[C]//Proceedings of the 18th National Conference on Artificial Intelligence. Edmonton, Canada: AAAI, 2002:786-791.

[2] AHN D. The stages of event extraction [C]// Proceedings of Association for Computational Linguistics 2006 Workshop on Annotating and Reasoning About Time and Events. Sydney,Australia:ACL,2006:1-8.

[3] LI Qi, JI Heng, HUANG Liang. Joint event extraction via structured prediction with global features [C]// Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. Stroudsburg, USA:ACL,2013:73-82.

[4] NGUYEN T H,GRISHMAN R. Event detection and domain adaptation with convolutional neural networks [C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Stroudsburg, USA:ACL,2015:365-371.

[5] CHEN Yubo,XU Liheng,LIU Kang,et al. Event extraction via dynamic multi-pooling convolutional neural networks [C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Stroudsburg,USA:ACL,2015:167-176.

[6] NGUYEN T H, CHO K, GRISHMAN R. Joint event extraction via recurrent neural networks[C]//Proceedings of the North American Chapter of the Association for Computational Linguistics-Human Language Technologies. Stroudsburg,USA:ACL,2016:300-309.

[7] FENG Xiaocheng, HUANG Lifu, TANG Duyu, et al. A language-independent neural network for event detection[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Stroudsburg,USA:ACL,2016:66-71.

(下转第 224 页)

计算机工程 www.ecice06.com

(上接第 217 页)

[8] LIU Shulin, CHEN Yubo, LIU Kang, et al. Exploiting argument information to improve event detection via supervised attention mechanisms[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, USA: ACL, 2017:1789-1798.

[9] FINKEL J R, GRENAGER T, MANNING C. Incorporating non-local information into information extraction systems by gibbs sampling[C]//Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics. Stroudsburg, USA: ACL, 2005:363-370.

[10] MASLENNIKOV M, CHUA T. A multi resolution framework for information extraction from free text[C]//Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics. Stroudsburg, USA: ACL, 2007:592-599.

[11] JI Heng, GRISHMAN R. Refining event extraction through cross-document inference[C]//Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, USA: ACL, 2008:254-262.

[12] GUPTA P, JI Heng. Predicting unknown time arguments based on cross-event propagation[C]//Proceedings of the Association for Computational Linguistics-International Joint Conference on Natural Language Processing. Stroudsburg, USA: ACL, 2009:369-372.

[13] LIAO Shasha, GRISHMAN R. Using document level cross-event inference to improve event extraction[C]//Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, USA: ACL, 2010:789-797.

[14] HONG Yu, ZHANG Jianfeng, MA Bin. Using cross-entity inference to improve event extraction[C]//Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, USA: ACL, 2010:1127-1136.

[15] 李培峰,周国栋,朱巧明. 基于语义的中文事件触发词抽取联合模型[J]. 软件学报,2016,27(2):280-294

[16] MIKOLOV T, CHEN Kai, CORRADO G, et al. Efficient estimation of word representations in vector space [EB/OL]. [2018-03-30]. https://arxiv.org/abs/1301.3781.

编辑 刘盛龄

计算机工程 www.ecice06.com