

# 基于知识图谱表示学习的谣言早期检测方法

皮德常<sup>1</sup>, 吴致远<sup>1</sup>, 曹建军<sup>2</sup>

(1. 南京航空航天大学计算机科学与技术学院, 江苏南京 211106;  
2. 中国人民解放军国防科技大学第六十三研究所, 江苏南京 210007)

**摘要:** 社交网络谣言是严重危害社会安全的一个重要问题。目前的谣言检测方法基本上都依赖用户评论数据。为了获取可供模型训练的足量评论数据, 需要任由谣言在社交平台上传播一段时间, 这就扩大了谣言的危害。本文提出了一种基于知识图谱表示学习的谣言检测方法。该方法不依赖用户评论数据。首先基于PN-KG2REC算法得到实体和关系的表示; 然后将待检测三元组中的实体和关系表示进行拼接, 得到三元组表示; 最后对三元组的向量表示进行分类, 并根据分类结果判断待检测三元组描述内容的真假性。采用公开数据的实验结果表明, 本文提出的谣言检测方法在不依赖用户评论数据的前提下, 能够有效地对谣言进行早期检测。

**关键词:** 谣言检测; 社交媒体; 知识图谱; 表示学习; 文本挖掘

**基金项目:** 科技创新2030—“新一代人工智能”重大项目(No. 2021ZD0113103)

**中图分类号:** TP393; TP309

**文献标识码:** A

**文章编号:** 0372-2112(2023)02-0385-11

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20210121

## Early Rumor Detection Method Based on Knowledge Graph Representation Learning

PI De-chang<sup>1</sup>, WU Zhi-yuan<sup>1</sup>, CAO Jian-jun<sup>2</sup>

(1. College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 211106, China;  
2. The Sixty-third Research Institute, National University of Defence Technology, Nanjing, Jiangsu 210007, China)

**Abstract:** Rumor on social network is an important issue that seriously harms social security. At present, most machine learning based rumor detection methods use comment data as auxiliary information for rumor detection, aiming to achieve higher accuracy. However, in order to obtain enough comment data for training models, it is inevitable that rumors will spread on social platforms for a period of time, which further expands the harm. This paper proposes a rumor detection method based on knowledge graph representation learning. This method uses the facts in the knowledge graph to detect rumors instead of relying on comment data. This method first obtains the representations of entities and relations based on proposed algorithm PN-KG2REC; and then concatenates the representations of entity and relation to obtain the representations of triplets; finally, it determines the authenticity of the contents described by the triplets according to the classification results. Experimental results show that the proposed method can effectively detect rumors without relying on comment posts in the early.

**Key words:** rumor detection; social media; knowledge graph; representation learning; text mining

**Foundation Item(s):** National Key Research and Development Program of China (No.2021ZD0113103)

### 1 引言

移动互联网技术的飞速发展以及手机、平板和个人电脑等终端设备的普及, 使得人们能够在第一时间将自己获取的信息分享到社交平台上。然而由于信息获取渠道的多样性以及个人知识储备的有限性, 个人在发布信息时, 并不能保证信息是真实无误的。因此,

当一条信息发布在社交平台上时, 为了避免虚假的信息被传播扩散, 需要对信息的真假性进行及时验证。但是, 由于信息刚刚发布时关注度不高或是发布者自身的关注者较少, 往往缺少用户的评论数据, 可利用的辅助信息较少。目前主流的基于用户评论数据的谣言检测方法不再适用, 检测的难度也随之增加。如果等到待

检测的信息在社交媒体上传播一段时间、用户评论数据足够多之后,再使用基于用户评论数据的谣言检测方法进行检测,有可能会扩大谣言带来的危害.因此,如何在不依赖用户评论数据的前提下,对社交媒体的谣言进行检测,是当下面临的难点和挑战.

判断发布的信息是否为谣言,本质上是在判断该信息所描述的内容是否正确,符合事实.针对信息发布早期,评论数据缺失的问题,国内外学者展开了大量的研究,最主流的是基于谣言内容的检测方法,这些方法主要分为两类:(1)人工设计复杂的语义线索和特征用于谣言检测<sup>[1,2]</sup>; (2)根据帖子的语义信息判断帖子描述的内容是否为谣言<sup>[3,4]</sup>,例如,如果信息是负面的,具有煽动性的语义,则认为该帖子是谣言.但是这两类方法都有其局限性,前者需要设计合乎逻辑的规则和特征,后者判断信息是否为谣言的标准过于片面,并且从实验结果来看,两类方法检测的准确率都不理想.针对上述问题,本文借助知识图谱实现谣言检测.对于待检测的帖子,可以通过实体识别和关系抽取将其描述的内容表示为一系列实体和关系构成的三元组.而知识图谱天然地以(头实体,关系,尾实体)的形式对客观现实进行描述,可以用来辅助判断抽取的三元组真假性.

本文提出了一种新颖的基于知识图谱表示学习的谣言早期检测方法.该方法首先基于提出的 PN-KG2REC 算法得到实体和关系的表示;然后将待检测三元组中的实体和关系表示进行拼接,得到三元组的表示;最后对三元组的向量表示进行分类,并根据分类结果判断待检测三元组描述内容的真假性.本文方法可以作为现有谣言检测方法的补充.例如,在帖子刚发布时,可以使用本文方法对帖子的真假性进行检测,得到一个早期的判断结果.等帖子传播一段时间、用户评论数据足够多的时候,使用现有的基于评论数据的谣言检测方法再次检测帖子的真假性.采用公开实验数据的研究结果表明,本文提出的谣言检测方法在不依赖用户评论数据的前提下,能够有效地检测早期谣言.

本文贡献概括为以下三个方面:

(1)通过随机游走生成“实体-关系”序列,引入知识图谱结构信息,通过超矩形(hyper-rectangle)表征具有相似信息的实体集合,考虑了实体属性的相似性和差异性,提升了实体和关系表示的能力;

(2)本文提出了两种用于生成超矩形区域外实体的策略,通过生成区域外实体,引入了非事实信息;

(3)采用公开数据集上的实验结果表明,提出的谣言检测方法能够有效地检测谣言,且检测准确率优于目前的前沿方法.

## 2 相关工作

目前多数学者将谣言定义为:在缺乏事实依据支撑的情况下传播开来的信息<sup>[5,6]</sup>.主流方法是基于用户评论数据的谣言检测方法.这些学者将用户的评论数据按照时间的先后顺序排列构成时间序列,通过对时间序列进行建模研究谣言的时序特性,进而根据时序特性对谣言进行分类.为了提取谣言传播过程中的时序特征,文献[7]提出了一种新颖的时间序列模型 DSTS (Dynamic Series-Time Structure),该模型将评论数据构成的时间序列分割成固定的时隙,根据每一个时隙中的文本内容和用户信息构建特征,并通过比较相邻时隙之间特征的差异,研究在信息传播过程中的不同时间节点,其上下文环境会如何变化.为了更好地捕捉谣言传播的时序特性和动态变化,文献[8]引入循环神经网络 RNN (Recurrent Neural Network)对评论数据构成的时间序列进行建模,每一个时隙的特征作为 RNN 单元的输入,RNN 网络的输出是关于整个时间序列的向量表示,最终根据时间序列的向量表示对谣言进行分类.为了捕捉用户评论数据的结构特性,一些学者根据用户回复、评论和转发关系构建源帖子为根的树型传播结构,利用信息的传播结构和传播模式进行谣言检测.文献[9]通过引入递归神经网络<sup>[10]</sup>,自动对传播树的特征进行细粒度的提取,提高了模型检测精度.

也有学者研究在不依赖用户评论数据的前提下如何进行谣言检测.其中,最常见的是基于内容的检测.这类方法认为谣言和非谣言在内容上存在差异.文献[1]设计了多种语义线索,对于待检测的信息,首先按照线索集合中定义的语义线索提取文本内容的特征,然后根据提取的特征来鉴别文本的真假.文献[2]引入了一些社交平台特有的元素,设计了细粒度的语义线索,如推文中“#”符号和“@”符号的数量等,其中“#”符号用于标识一个话题,“@”符号用于向特定的用户发送信息.但是上述方法需要设计合乎逻辑的语言规则或特征,这需要耗费大量的人力和时间,且只能针对特定平台或主题,不具备一般性.为了降低特征工程的成本以及提取更深层次的语义特征,一些学者使用机器学习算法和深度学习算法对谣言的文本特征进行提取<sup>[11]</sup>.文献[12]首先基于词袋模型将文本内容表示为词频向量,然后将词频向量作为朴素贝叶斯和支持向量机等分类模型的输入,根据分类结果判断待检测的信息真假性.为了降低人工设计特征的成本,文献[4]基于卷积神经网络 CNN<sup>[12]</sup>提出一种提取文本特征的神经网络结构 FNDNet<sup>[4]</sup>.该网络可以自动从文本内容中提取出对谣言分类有帮助的特征,通过堆叠网络的层数使得模型可以学习到更深层次的语义信息,进而实现谣言检测.

上述基于内容的检测方法需要设计合乎逻辑的规则和特征,并且判断信息是否为谣言的标准过于片面.针对这一问题,一些学者借助知识图谱<sup>[13]</sup>实现对社交媒体谣言的检测.文献[14]使用实体识别和关系抽取技术,将数据集中的新闻数据抽象为(头实体,关系,尾实体)的形式,例如,新闻“拜登赢得2020年美国大选”表示为三元组(拜登,获胜,2020年美国大选);然后根据数据集中的真实新闻数据,构建真实信息的知识图谱,根据虚假新闻数据构建虚假信息知识图谱;接着使用TransE算法<sup>[15]</sup>分别训练上述知识图谱,学习知识图谱中的实体和关系的向量表示,通过融合策略将学习到最大偏移向量进行拼接;最后提出了一个分类模型.文献[15]同样使用实体识别和关系抽取技术,在知识图谱上进行随机游走<sup>[16]</sup>生成游走序列,并采用Word2vec算法<sup>[17]</sup>训练词向量的方式,根据游走序列中的上下文预测目标实体或目标关系,得到知识图谱中实体和关系的向量表示,最后将三元组的向量表示作为随机森林分类器的输入,根据输出结果判断三元组所描述内容真假性.

### 3 相关定义与问题描述

**定义1 实体识别和关系抽取** 实体识别和关系抽取的目的是,从非结构化的文本数据中发掘实体以及实体之间的关系,从而将非结构化的文本数据表示为结构化的三元组数据 $(h, r, t)$ ,其中 $h$ 表示头实体, $t$ 表示尾实体, $r$ 表示实体之间的关系.

**定义2 知识图谱表示学习** 知识图谱表示学习的目的是,将知识图谱中的实体 $e \in \mathcal{E}$ 和关系 $r \in \mathcal{R}$ 嵌入低维连续的向量空间,使得嵌入后的实体表示和关系表示能够满足知识图谱所描述的正确的事实 $\mathcal{F}^+$ ,从而简化知识图谱上的计算任务.知识图谱表示学习算法一般都可以分为三个步骤.

第一,定义实体 $e \in \mathcal{E}$ 和关系 $r \in \mathcal{R}$ 的表示形式,即实体和关系以什么样的形式嵌入低位连续的向量空间;第二,定义关于(头实体,关系,尾实体)三元组 $(h, r, t)$ 的得分函数,该得分函数用来衡量三元组 $(h, r, t)$ 的真假性,当三元组 $(h, r, t)$ 描述的是一个客观存在的事实时,即 $(h, r, t) \in \mathcal{F}^+$ ,代入得分函数会得到一个较高的得分,反之,当三元组 $(h, r, t)$ 描述的是一个错误的事实时,即 $(h, r, t) \in \mathcal{F}^-$ ,代入得分函数将得到一个较低的得分,其中 $\mathcal{F}^- \subseteq \mathcal{E} \times \mathcal{R} \times \mathcal{E}$ 为错误的事实集合;第三,通过优化算法,极大化正确三元组 $(h, r, t) \in \mathcal{F}^+$ 的得分和极小化错误三元组 $(h, r, t) \in \mathcal{F}^-$ 的得分,最终得到训练好的实体和关系嵌入表示.

**问题描述:**对于社交平台上发布的一则帖子,我们希望在无用户评论数据的前提下,根据已有知识图谱

判断待检测帖子内容的真假.由于待检测的帖子所描述的内容可以被抽象为(头实体,关系,尾实体)的形式,而知识图谱中又包含了实体之间的关系,因此,基于知识图谱进行谣言检测可以转化为两个子问题:第一,如何对知识图谱中的实体和关系进行表示;第二,如何根据实体和关系的表示,判断待检测三元组描述的内容的真假性.

## 4 本文提出的谣言检测方法

为了尽可能减少谣言带来的危害,需尽早地检测谣言.本文提出了基于知识图谱表示学习的谣言早期检测算法.本节首先介绍如何在知识图谱上进行随机游走生成游走序列,以及如何对知识图谱中的实体和关系进行表示;然后阐述用来衡量三元组真假性的距离函数和两种用于生成负样本的策略;最后,提出PN-KG2VEC模型的训练方法,并介绍如何根据知识图谱中的实体和关系的表示对谣言进行分类.

### 4.1 知识图谱上的随机游走

对于知识图谱 $\mathcal{G} = \{\mathcal{E}, \mathcal{R}, \mathcal{F}^+\}$ ,其中 $\mathcal{E}$ 表示知识图谱中实体的集合, $\mathcal{R}$ 表示知识图谱中关系的集合, $\mathcal{F}^+ \subseteq \mathcal{E} \times \mathcal{R} \times \mathcal{E}$ 为知识图谱所描述的正确的事实集合.知识图谱中的每一个顶点都对应一个实体 $e$ ,每一条边都对应一种关系,且任意一条边都可以通过一个正确的事实三元组 $(h, r, t) \in \mathcal{F}^+$ 唯一确定.选择知识图谱 $\mathcal{G}$ 中的任意顶点(实体)作为起始顶点,从起始顶点开始进行 $L$ 次随机游走,可以得到长度为 $2L+1$ 的游走序列 $\mathcal{S} = \{s^{(0)}, s^{(2)}, \dots, s^{(2L+1)}\}$ ,其中序列元素 $s^{(2^{*}l-1)}$ ( $l_1 = 1, 2, \dots, L+1$ )为游走序列中的实体,序列节点 $s^{(2^{*}l_2)}$ ( $l_2 = 1, 2, \dots, L$ )为游走序列中的关系.一个长度为 $2L+1$ 的游走序列 $\mathcal{S} = \{s^{(0)}, s^{(2)}, \dots, s^{(2L+1)}\}$ ,可以看作由 $T$ 个子序列 $\{s^{(2l-1)}, s^{(2l)}, s^{(2l+1)}\}$ 构成,其中 $l = 1, 2, \dots, L$ ,每一个子序列 $\{s^{(2l-1)}, s^{(2l)}, s^{(2l+1)}\}$ 都对应一个正确的事实三元组 $(h, r, t) \in \mathcal{F}^+$ .对于一个子序列 $\{s^{(2l-1)}, s^{(2l)}, s^{(2l+1)}\}$ ,假设序列元素 $s^{(2l-1)}$ 对应实体 $e_i$ ,元素 $s^{(2l+1)}$ 对应实体 $e_j$ ,元素 $s^{(2l)}$ 对应第 $k$ 种类型的关系 $r_k$ ,则从顶点 $e_i$ 开始经由边 $(e_i, r_k, e_j)$ 游走到顶点 $e_j$ 的概率如式(1)所示:

$$p(\Psi(s^{(2l+1)}) = j | \Psi(s^{(2l-1)}) = i, \Phi(s^{(2l)}) = k) = \begin{cases} \frac{1}{Z_i}, & \text{if } (e_i, r_k, e_j) \in \mathcal{F}^+ \\ 0, & \text{if } (e_i, r_k, e_j) \in \mathcal{F}^- \end{cases} \quad (1)$$

其中, $\Psi$ 为知识图谱中的顶点(实体)到顶点序号的映射,例如,当 $s^{(2l-1)} = e_i$ 时, $\Psi(s^{(2l-1)}) = i$ ; $\Phi$ 为知识图谱中的边到关系的映射,例如,当 $s^{(2l)} = r_k$ 时, $\Phi(s^{(2l)}) = k$ ;  $Z_i$ 为以实体 $e_i$ 为顶点的所有出边的个数之和,即 $Z_i = \sum_{j=1}^{|\mathcal{E}|} \sum_{k=1}^{|\mathcal{R}|} \text{count}(e_i, r_k, e_j)$ ,  $\text{count}(e_i, r_k, e_j)$ 为计数函数,如果

$(e_i, r_k, e_j) \in \mathcal{F}^+$ , 则  $\text{count}(e_i, r_k, e_j)$  为 1, 否则为 0. 显然, 由式(1)可知, 实体顶点  $e_i$  和实体顶点  $e_j$  之间的边越多, 越有可能从顶点  $e_i$  游走到顶点  $e_j$ .

#### 4.2 实体和关系的嵌入表示

为了提取知识图谱中包含的事实信息, 需要对知识图谱中的实体和关系进行嵌入表示. 传统的知识图谱表示学习算法, 将实体表示为向量空间中的独立点, 不能同时有效地表征实体之间的复杂关系, 如一对多关系、多对一关系, 以及组合关系和对称关系等. 以 TransE 算法为例, 假设用户 A 和用户 B 具有朋友关系, 令  $e_A$  为用户 A 的向量表示,  $e_B$  为用户 B 的向量表示,  $r_{\text{friend}}$  为朋友关系的向量表示, 则根据 TransE 算法需要同时满足  $\|e_A + r_{\text{friend}} - p_B\| = 0$  和  $\|e_B + r_{\text{friend}} - p_A\| = 0$ , 只有当  $r_{\text{friend}}$  为零向量且  $e_A = e_B$  时, 两者才能同时成立, 但是 A 和 B 是不同的两个人, 它们的向量表示  $e_A$  和  $e_B$  应该位于向量空间中的不同位置, 因此  $e_A = e_B$  显然是不合理的.

本文使用超矩形来表征具有相似信息的实体的集合. 每一个实体都会被表示为向量空间中的一个点, 但具有相似信息的实体会被包含在一个用来表征该信息的超矩形中, 假设低维连续向量空间为  $d$  维, 超矩形  $q$  中所包含实体的向量表示的集合  $\text{Rec}_q$  的定义如式(2)所示:

$$\text{Rec}_q = \{e_i \in \mathbb{R}^d : \text{Cen}(q) - \text{Off}(q) \preceq e_i \preceq \text{Cen}(q) + \text{Off}(q)\} \quad (2)$$

其中, 超矩形  $q$  的向量表示  $(\text{Cen}(q), \text{Off}(q)) \in \mathbb{R}^{2d}$  包含两个部分,  $\text{Cen}(q) \in \mathbb{R}^d$  表示超矩形  $q$  的中心向量, 用来描述超矩形的位置,  $\text{Off}(q) \in \mathbb{R}^d$  表示超矩形  $q$  的正偏移向量, 用来描述超矩形的大小.  $e_i \in \mathbb{R}^d$  为实体  $e_i$  在向量空间中的表示. 符号  $\preceq$  表示元素层级的小于或等于, 例如,  $e_A = [e_A^1, e_A^2, \dots, e_A^d]^T \in \mathbb{R}^d \preceq e_B = [e_B^1, e_B^2, \dots, e_B^d]^T \in \mathbb{R}^d$  等价于对于任意的  $i, i = 1, 2, \dots, d$ , 满足  $p_A^i \leq p_B^i$ .

对于知识图谱中任意实体  $e_i$ , 其向量表示  $e_i \in \mathbb{R}^d$  为向量空间中的一个点, 可以看作一个中心向量为  $e_i$ 、大小为 0 (即偏移向量为零向量) 的超矩形, 写成超矩形的形式为  $\text{Rec}_{e_i} = (\text{Cen}(e_i) = e_i, \mathbf{0}) \in \mathbb{R}^{2d}$ , 其中  $\mathbf{0}$  表示元素全为 0 的  $d$  维向量. 对于知识图谱中的任意边对应的关系  $r$ , 其向量表示  $r = (\text{Cen}(r), \text{Off}(r)) \in \mathbb{R}^{2d}$  同样由两部分组成, 其中  $\text{Cen}(r)$  表示实体经过关系  $r$  的投影后, 中心向量的改变量;  $\text{Off}(r)$  表示实体经过关系  $r$  的投影后, 偏移向量的改变量. 因此, 对于知识图谱中的 (头实体, 关系, 尾实体) 三元组  $(e_h, r, e_t)$ , 考虑投影  $\text{Rec}_{e_h} + r$ , 可以得到一个新的超矩形  $q_i$ , 其向量表示的计算方式如式(3)所示:

$$\begin{aligned} q_i &= (\text{Cen}(q_i), \text{Off}(q_i)), \\ \text{Cen}(q_i) &= \text{Cen}(e_h) + \text{Cen}(r), \\ \text{Off}(q_i) &= \text{Off}(e_h) + \text{off}(r) \end{aligned} \quad (3)$$

其中, 新得到的超矩形  $q_i$  的向量表示也包含两部分, 中心向量  $\text{Cen}(q_i)$  由头实体  $e_h$  的中心向量和关系  $r$  的中心向量相加得到, 偏移向量  $\text{Off}(q_i)$  由头实体  $e_h$  的偏移向量和关系  $r$  的偏移向量相加得到. 如图 1 所示, 实体  $e_h$  经过关系  $r$  投影得到新的超矩形  $q_i$ .

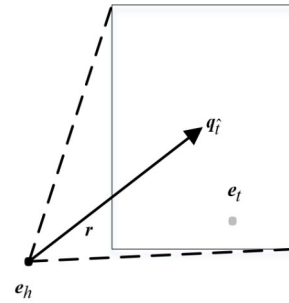


图 1 投影

图 1 中,  $e_h$  为头实体的中心向量,  $r$  为关系的表示,  $q_i = (\text{Cen}(q_i), \text{Off}(q_i))$  为投影后得到的超矩形  $q_i$  的表示,  $e_t$  为尾实体的表示. 超矩形  $q_i$  的中心向量  $\text{Cen}(q_i)$  由头实体  $e_h$  的中心向量  $e_h$  和关系  $r$  的中心向量  $\text{Cen}(r)$  相加得到. 超矩形  $q_i$  的偏移向量  $\text{Off}(q_i)$  由关系  $r$  的偏移向量  $\text{Off}(r)$  确定.

在 4.1 节通过知识图谱上的随机游走得到不同长度的游走序列.  $T$  次随机游走, 可以得到长度为  $2L+1$  的游走序列  $\mathcal{S} = (s^{(1)}, s^{(2)}, \dots, s^{(2L+1)})$ , 其中序列元素  $s^{(2^*l_1-1)}$  ( $l_1 = 1, 2, \dots, L+1$ ) 表示游走序列中的第  $l_1$  个顶点, 对应着知识图谱中的实体, 序列节点  $s^{(2^*l_2)}$  ( $l_2 = 1, 2, \dots, L$ ) 表示游走序列中的第  $l_2$  条边, 对应着知识图谱中的关系. 因此, 将一个长度为  $2L+1$  的游走序列中的元素替换为实体和关系后, 可以得到一个新的实体-关系序列  $\mathcal{S}'$ , 如式(4)所示:

$$\mathcal{S}' = (e_{\Psi(s^{(1)})}, r_{\Phi(s^{(2)})}, \dots, e_{\Psi(s^{(2L+1)})}, r_{\Phi(s^{(2L)})}, \dots, e_{\Psi(s^{(2L+1)})}) \quad (4)$$

其中,  $\Psi$  为知识图谱中的顶点到实体序号的映射, 例如, 当  $s^{(2^*l_1-1)} = e_i$  时,  $\Psi(s^{(2^*l_1-1)}) = i$ .  $\Phi$  为知识图谱中的边到关系序号的映射, 例如, 当  $s^{(2^*l_2)} = r_k$  时,  $\Phi(s^{(2^*l_2)}) = k$ . 图 1 展示了长度为 3 的游走序列的投影过程. 一般长度为  $2L+1$  的序列  $\mathcal{S}'$  的投影过程, 多次投影后得到的超矩形  $q_i$  的计算方式如式(5)~(7)所示:

$$q_i = (\text{Cen}(q_i), \text{Off}(q_i)) \quad (5)$$

$$\text{Cen}(q_i) = \text{Cen}(e_{\Psi(s^{(1)})}) + \sum_{l=1}^L \text{Cen}(r_{\Phi(s^{(2^*l)})}) \quad (6)$$

$$\text{Off}(q_i) = \text{Off}(e_{\psi(s^{(0)})}) + \sum_{l=1}^i \text{Off}(r_{\phi(s^{(2l)})}) \quad (7)$$

最终得到的超矩形向量表示包含两部分,即中心向量  $\text{Cen}(q_i)$  和偏移向量  $\text{Off}(q_i)$ . 中心向量  $\text{Cen}(q_i)$  由序列  $S'$  中的前  $2L$  个元素的中心向量相加得到,偏移向量  $\text{Off}(q_i)$  由序列  $S'$  中的前  $2L$  个元素的偏移向量相加得到,且因为实体的偏移向量为零向量,即  $\text{Off}(e_{\psi(s^{(0)})}) = \mathbf{0}$ ,所以实际上  $\text{Off}(q_i)$  由序列  $S'$  中的第 1 个元素到第  $2L$  个元素的偏移向量相加得到的. 图 2 展示了不同长度的游走序列对应的投影过程.

2018 年图灵奖的获得者为 Geoffrey Hinton, Yann LeCun 和 Yoshua Bengio, 因此可以得到长度为 3 的实体-关系序列即 (2018 图灵奖, 获得者, Geoffrey Hinton), (2018 图灵奖, 获得者, Yann LeCun) 和 (2018 图灵奖, 获得者, Yoshua Bengio). 实体 (2018 图灵奖) 经过关系 (获奖者) 投影后, 得到的超矩形向量表示  $q_1 = (\text{Cen}(\text{2018 图灵奖}) + \text{Cen}(\text{获奖者}), \text{Off}(\text{2018 图灵奖}) + \text{Off}(\text{获奖者}))$  应该能够包含 Geoffrey Hinton, Yann LeCun 和 Yoshua Bengio 这 3 个实体.

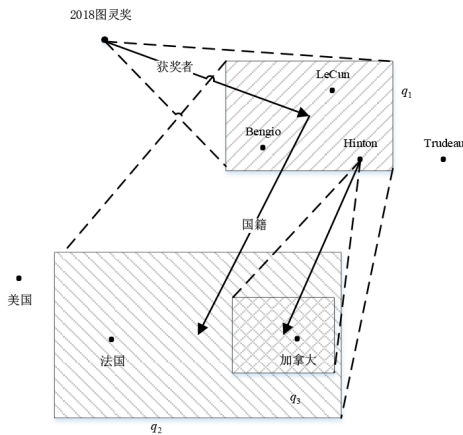


图 2 不同长度的游走序列的投影

加拿大总统 Justin Trudeau 并没有获得图灵奖, 因而不在超矩形  $q_1$  内. Geoffrey Hinton 的国籍是加拿大, 因此投影后得到的超矩形  $q_2$  应该包含加拿大这一实体. 同理, 考虑长度为 5 的实体-关系序列 (2018 图灵奖, 获得者, Geoffrey Hinton, 国籍, 加拿大), (2018 图灵奖, 获得者, Yann LeCun, 国籍, 法国) 和 (2018 图灵奖, 获得者, Yoshua Bengio, 国籍, 法国), 投影后得到的超矩形向量表示  $q_3 = (\text{Cen}(\text{2018 图灵奖}) + \text{Cen}(\text{获奖者}) + \text{Cen}(\text{国籍}), \text{Off}(\text{2018 图灵奖}) + \text{Off}(\text{获奖者}) + \text{Cen}(\text{国籍}))$  应该包含法国和加拿大这两个实体, 而不会包含美国这一实体.

### 4.3 距离函数

对于长度为  $2L + 1$  的实体-关系序列  $S'$ , 根据 4.2 节

描述的投影方式可以得到一个超矩形  $q$ , 我们希望超矩形  $q$  能够包含序列  $S'$  中目标实体  $e_{\psi(s^{(2L+1)})}$ . 因此, 需要定义一个关于超矩形的向量表示  $q$  和实体的向量表示  $v$  的距离函数  $\text{dist}(q, v)$ , 通过极小化向量空间中目标实体  $e_{\psi(s^{(2L+1)})}$  到超矩形  $q$  的距离  $\text{dist}(e_{\psi(s^{(2L+1)})}, q)$ , 训练模型参数, 得到知识图谱中实体和关系的向量表示.

本文采用文献 [18] 的距离函数, 该函数由区域内的距离和区域外的距离两部分组成, 如式 (8) 所示:

$$\begin{aligned} \text{Dist}(S') &= \text{dist}(v, q) \\ &= \text{dist}_{\text{outside}}(v, q) + \alpha \cdot \text{dist}_{\text{inside}}(v, q) \end{aligned} \quad (8)$$

其中,  $S'$  表示长度为  $2L + 1$  的实体-关系序列,  $e$  为序列  $S'$  的目标实体,  $q$  为根据序列  $S'$  投影得到的超矩形.  $e$  为实体的向量表示,  $q$  为超矩形的向量表示.  $\text{dist}_{\text{outside}}$  为区域外的距离,  $\text{dist}_{\text{inside}}$  为区域内的距离,  $\alpha$  为平衡参数, 用来调节区域外的距离和区域内的距离的比重.  $\text{dist}_{\text{outside}}$  和  $\text{dist}_{\text{inside}}$  的定义如式 (9) 和式 (10) 所示:

$$\text{dist}_{\text{outside}}(v, q) = \|\text{Max}(\text{Cen}(v) - q_{\text{max}}, \mathbf{0}) + \text{Max}(q_{\text{min}} - \text{Cen}(v), \mathbf{0})\|_1 \quad (9)$$

$$\text{dist}_{\text{inside}}(v, q) = \|\text{Cen}(q) - \text{Min}(q_{\text{max}}, \text{Max}(q_{\text{min}}, \text{Cen}(v)))\|_1 \quad (10)$$

其中,  $q_{\text{min}} = \text{Cen}(q) - \text{Off}(q)$ ,  $q_{\text{max}} = \text{Cen}(q) + \text{Off}(q)$ .  $\text{Max}(x, \mathbf{0})$  表示如果向量  $x$  某一维度的值小于 0, 则值变为 0, 否则不变.  $\text{Min}(x, \mathbf{0})$  表示如果向量  $x$  某一维度的值大于 0, 则值变为 0, 否则不变.  $\|x\|_1$  表示向量  $x$  的 L1 范数, 即向量  $x$  各维度的绝对值之和. 分析式 (8) 和 (9) 可知, 当实体  $e$  在超矩形  $q$  的外部时, 区域外的距离  $\text{dist}_{\text{outside}}(v, q)$  为实体  $v$  各个维度到超矩形  $q$  对应边界的最短距离之和, 区域内的距离  $\text{dist}_{\text{inside}}(v, q)$  为超矩形  $q$  的偏移向量  $\text{Off}(q)$  的 L1 范数; 当实体  $e$  在超矩形  $q$  的内部时, 区域外的距离  $\text{dist}_{\text{outside}}(v, q)$  为 0, 区域内的距离  $\text{dist}_{\text{inside}}(v, q)$  为实体  $e$  各个维度到中心向量  $\text{Cen}(q)$  对应维度的距离之和. 图 3 展示了距离函数的计算方式.

向量空间的维度为 2, 超参数  $\alpha = 1$ , 超矩形  $q$  的中

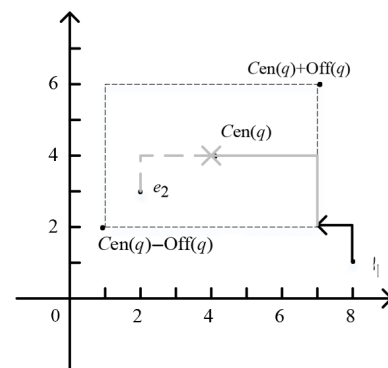


图 3 区域内和区域外距离的计算

心向量为(4,4),偏移向量为(3,2).根据超矩形的定义可知,超矩形的边界为图3中的虚线部分.对于实体 $e_1$ ,其位于超矩形 $q$ 的外部,中心向量 $e_1$ 为(8,1),计算实体 $e_1$ 到超矩形 $q$ 的距离,外部距离的计算方式如黑色实线箭头所示, $x$ 轴方向上的距离为1, $y$ 轴方向上的距离为1,外部距离为2,内部距离计算方式如灰色实线箭头所示,内部距离为 $2+3=5$ ,因此,实体 $e_1$ 到超矩形 $q$ 的距离为 $2+5=7$ ;对于实体 $e_2$ ,其位于超矩形 $q$ 的内部,中心向量 $e_2$ 为(2,3),外部距离为0,内部距离计算方式如灰色虚线箭头所示, $x$ 轴方向上的距离为2, $y$ 轴方向上的距离为1,内部距离为 $2+1=3$ ,因此实体 $e_2$ 到超矩形 $q$ 的距离为 $0+3=3$ .

#### 4.4 区域外实体的选择

对于实体-关系序列 $\mathcal{S}'$ ,我们希望目标实体 $e_{\psi(\mathcal{S}^{2L+1})}$ 距离投影得到的超矩形 $q$ 尽可能近.而对于不应该包括在超矩形内的实体 $e^-$ ,我们希望这些实体距离超矩形 $q$ 的距离尽可能大.在本文中,我们采取两种策略生成区域外的实体 $e^-$ .

策略一:对于实体-关系序列 $\mathcal{S}'=(e_{\psi(\mathcal{S}^{(0)})}, r_{\phi(\mathcal{S}^{(0)})}, \dots, e_{\psi(\mathcal{S}^{(2L-1)})}, r_{\phi(\mathcal{S}^{(2L-1)})}, \dots, e_{\psi(\mathcal{S}^{(2L+1)})})$ ,将其中的关系 $r_{\phi(\mathcal{S}^{(2L)})}, l=1, 2, \dots, L$ ,随机地替换为其他的关系,将序列中的最后一个实体 $e_{\psi(\mathcal{S}^{(2L+1)})}$ (目标实体)随机地替换为其他的实体,得到替换后的实体-关系序列 $\mathcal{S}^-$ .对于序列 $\mathcal{S}^-$ ,投影得到的超矩形 $q^-$ 应该不包括目标实体 $\mathcal{S}^-(2L+1)$ ,其中 $\mathcal{S}^-(i)$ 表示序列 $\mathcal{S}^-$ 中的第 $i$ 个元素,目标实体 $\mathcal{S}^-(2L+1)$ 即为序列中的最后一个实体.将使用第一种策略得到的所有可能的实体-关系序列 $\mathcal{S}^-$ 的集合记作 $D_1^-$ .

策略二:用来生成实体-关系序列 $\mathcal{S}'$ 的知识图谱 $G$ 是基于数据集中的正样本构建的,知识图谱 $G$ 中的三元组描述的都是客观事实,而数据集中还存在大量的非事实信息,例如三元组(中国,首都,南京)描述的是一个错误的信息.因此,在训练知识图谱中实体和关系的向量表示时,数据集中表示错误信息的三元组,可以被用来生成区域外的实体 $e^-$ .具体的做法是,对于任意一个实体-关系序列 $\mathcal{S}'$ ,根据序列的起始元素 $e_{\psi(\mathcal{S}^{(0)})}$ ,在数据集的负样本中查找所有以 $e_{\psi(\mathcal{S}^{(0)})}$ 作为头实体的三元组(以 $e_{\psi(\mathcal{S}^{(0)})}$ 作为头实体的三元组可以视为长度为3的实体-关系序列),将所有查找到的三元组的集合记作 $D_2^-$ ,由于存在查找不到的情况,因此集合 $D_2^-$ 可能为空.

因此,对于任意一个生成实体-关系序列 $\mathcal{S}'$ ,可以通过第一种策略和第二种策略分别得到集合 $D_1^-$ 和集合 $D_2^-$ .对于集合 $D_1^-$ 和集合 $D_2^-$ 中的任意一个长度为 $2L+1$ 实体-关系序列 $\mathcal{S}^-$ ,投影得到的超矩形 $q^-$ 应该不包括目标实体 $\mathcal{S}^-(2L+1)$ ,即希望距离 $\text{dist}(\mathcal{S}^-(2L+1), q^-)$ 尽可能大.

#### 4.5 PN-KG2REC的训练方式

将提出的知识图谱表示学习算法命名为PN-KG2REC.其中,KG2REC的意义为,知识图谱中的实体 $e$ 和关系 $r$ 的表示由中心向量和偏移向量两部分组成,实体 $e$ 的空间表示可以看作位置为 $\text{cen}(e)$ 、大小为 $\text{off}(e)$ 的超矩形,实体 $e$ 经过关系 $r$ 的投影也是一个超矩形.PN的意义是,在生成域外实体的过程中,使用了数据集的正样本(positive)和负样本(negative).

PN-KG2REC算法的目标函数如式(11)所示:

$$\mathcal{L} = -\log(\sigma(\gamma - \text{Dist}(\mathcal{S}^+))) + \frac{1}{|D_1^-|} \sum_{\mathcal{S}^- \in D_1^-} \log(\gamma - \text{Dist}(\mathcal{S}^-)) + \frac{1}{|D_2^-|} \sum_{\mathcal{S}^- \in D_2^-} \log(\gamma - \text{Dist}(\mathcal{S}^-)) \quad (11)$$

其中, $\text{Dist}(\mathcal{S}')$ 为式(3)~(5)所计算的距离,表示长度为 $2L+1$ 的实体-关系序列 $\mathcal{S}'$ 中,目标实体到投影超矩形的距离. $\sigma(x)=1/(1+e^{-x})$ 为sigmoid激活函数,参数 $\gamma$ 为定值,对于实体-关系序列 $\mathcal{S}'$ ,我们希望 $\text{Dist}(\mathcal{S}')$ 尽可能小,因此当 $\text{Dist}(\mathcal{S}')$ 比 $\gamma$ 大时,计算 $-\log(\sigma(\gamma - \text{Dist}(\mathcal{S}^+)))$ 会得到较大的损失值;对于实体-关系序列 $\mathcal{S}^-$ ,我们希望 $\text{Dist}(\mathcal{S}^-)$ 尽可能大,因此当 $\text{Dist}(\mathcal{S}^-)$ 比 $\gamma$ 小时,计算 $\log(\sigma(\gamma - \text{Dist}(\mathcal{S}^-)))$ 会得到较大的损失值. $|D_1^-|$ 和 $|D_2^-|$ 分别表示集合 $D_1^-$ 和集合 $D_2^-$ 中实体-关系序列的个数.

算法PN-KG2REC的训练方式如算法1所示.

##### 算法1 PN-KG2REC,一种知识图谱表示学习算法

**输入:** 知识图谱 $\mathcal{G}=\{\mathcal{E}, \mathcal{R}, \mathcal{F}^+\}$ ,训练数据集中的负样本三元组,随机游走的最大长度 $L$ ,每一个实体作为游走起点的游走序列个数 $M$ .

**输出:** 知识图谱中实体和关系的表示.

开始

1. 按照4.2节的方式对实体和关系的表示进行初始化
2. 游走序列的集合 $\mathcal{S}=\emptyset$
3. for 实体 $e$  in  $\mathcal{E}$  do
4.     for  $m=1, 2, \dots, M$  do
5.         随机选择区间 $[1, L]$ 之间的正整数 $l$ ,以实体 $e$ 作为起点
6.         进行 $l$ 次随机游走,得到长度为 $2l+1$ 的游走序列 $S$
7.         将游走序列 $S$ 添加到集合 $\mathcal{S}$
8.     end for
9. end for
10. for  $S$  in  $\mathcal{S}$  do
11.     根据式(4)将游走序列 $S$ 转换为实体-关系序列 $\mathcal{S}'$
12.     分别使用策略一和策略二生成集合 $D_1^-$ 和集合 $D_2^-$
13.     根据式(11)计算目标函数的值
14.     使用Adam优化器极小化目标函数的值,更新参数
15. end for
16. 输出知识图谱中实体和关系的表示

结束

算法 1 展示了 PN-KG2REC 算法训练得到实体和关系表示的过程,算法第 1 行将对实体和关系的表示随机初始化,但需要保证实体的偏移向量为零向量. 算法第 2 行创建游走序列的集合  $\mathcal{S}$ . 算法第 5~7 行以知识图谱中的每一个实体作为起始顶点,进行随机游走,将得到的游走序列添加到集合  $\mathcal{S}$  中. 算法第 11~14 行是具体的训练过程,首先将游走序列转换为实体-关系序列,然后使用策略一和策略二生成区域外实体,并通过 Adam<sup>[19]</sup> 优化器极小化目标函数的值,从而更新训练参数. PN-KG2REC 训练结束后输出实体和关系的表示.

#### 4.6 谣言的分类

对于社交媒体上的一则待检测的帖子,首先需要使用实体识别和关系抽取技术,将该帖子描述的内容抽取为一个或者多个三元组的形式. 对于任意一个三元组  $(e_h, r, e_t)$ ,将三元组中实体和关系的中心向量进行拼接,作为三元组的向量表示. 本文使用 XGBoost 算法<sup>[20]</sup>作为分类器,对三元组的向量表示进行分类:如果被分类为正例,则表示该三元组描述是正确的;如果被分类为负例,则表示该三元组描述的是错误的事实. 因为一个帖子描述的内容可能被抽取为多个三元组,因此,只有在所有的三元组都被分类为正例的情况下,才可以认为待检测的帖子是正确的. 否则,只要有一个三元组被分类为负例,就认为该帖子为谣言.

为了使 XGBoost 分类器能够正确地根据三元组的向量表示进行分类,需要使用训练集数据训练 XGBoost 分类器. 对于训练数据集中的任意三元组  $(h, r, t)$ ,如果该三元组描述的是一个客观存在的事实时,即  $(h, r, t) \in \mathcal{F}^+$ ,则赋予该三元组一个正样本标签;反之,当三元组  $(h, r, t)$  描述的是一个错误的事实时,即  $(h, r, t) \in \mathcal{F}^-$ ,则赋予该三元组一个负样本标签. 最终,通过使用训练数据集中正样本和负样本训练 XGBoost 分类器,使之能够正确地判断三元组所描述的内容.

## 5 实验及结果分析

### 5.1 数据集和评价指标

采用公开数据集 Semantic Web Challenge 2019<sup>[21]</sup> 对提出的谣言检测方法进行验证. 该数据集包含 25 000 个三元组,三元组的真实性已经有专业的工作人员进行验证分类,其中正样本 12 500 个,负样本 125 000 个,包含 11 990 个不同的实体以及 5 种不同的关系. 正样本三元组描述的是客观事实,可以作为非谣言数据,负样本三元组描述的是错误的事实,可以作为谣言数据. 需要注意的是,由于该数据集中的正样本和负样本都是以(头实体,关系,尾实体)的形式呈现,因此无需进行信息抽取. 如果数据集中的谣言和非谣言为文本数据,则需要使用实体识别和关系抽取技术,将

文本信息表示为(头实体,关系,尾实体)的形式. 实验数据集的统计信息如表 1 所示.

表 1 Semantic Web Challenge 2019 数据集统计信息

名称	数量
实体数量	11 990
关系类型数量	5
三元组总数	25 000
正样本占比	50%
负样本占比	50%

本文按照 8:2 的比例将实验数据集切分为训练集和测试集,并且保证测试集中出现的实体和关系都包含在训练集中. 本文只使用训练集中的正样本三元组构建知识图谱,从而保证不会泄露测试数据集中的事实信息.

实验数据集中正负样本的比例为 1:1,不存在类别不平衡问题,因此使用正确率作为评价指标.

### 5.2 比较方法

本文提出的谣言检测方法其核心思想是基于知识图谱表示学习算法 PN-KG2REC 得到实体和关系的表示,然后使用 XGBoost 算法对文本中包含的(头实体,关系,尾实体)三元组进行分类. 为了检验本文提出的 PN-KG2REC 算法在谣言检测任务上的有效性,将 PN-KG2REC 算法与一些其他知识图谱表示学习算法在谣言检测任务上的表现做了详细的比较. 在进行比较时,本文使用不同的知识图谱表示学习算法训练得到实体和关系的表示,然后对进行谣言分类,观察并分析分类结果. 对比的知识图谱表示学习算法包括以下 6 种.

TransE 算法<sup>[22]</sup>:将知识图谱中的实体和关系表示为同一特征空间中的向量,对于描述正确事实的三元组  $(h, r, t)$ ,头实体的向量表示  $h \in \mathbb{R}^d$  和尾实体的向量表示  $t \in \mathbb{R}^d$  可以通过关系的向量表示  $r \in \mathbb{R}^d$  进行连接(允许存在一定的误差),即满足  $h + r \approx t$ . 其目标函数如下式(12)所示:

$$\mathcal{L} = \sum_{(h, r, t) \in G} \sum_{(h', r', t') \in G'} [\gamma + \|h + r - t\|_2^2 - \|h' + r' - t'\|_2^2]_+ \quad (12)$$

其中,  $[x]_+$  表示如果  $x$  大于 0 则结果为  $x$ , 否则为 0;  $\gamma$  为手动设置的超参数;  $h, r, t$  分别表示三元组中实体和关系对应的向量;  $(h', r', t')$  是关于  $(h, r, t)$  的负采样;  $\|x\|_2^2$  表示向量  $x$  的 L2 范数的平方.

TransH 算法<sup>[23]</sup>:将实体  $h$  表示为空间中的向量  $h \in \mathbb{R}^d$ ,将关系  $r$  表示为超平面上的向量  $r$ ,超平面的法向量为  $w_r \in \mathbb{R}^d$ . 对于描述正确事实的三元组  $(h, r, t)$ ,TransH 算法计算头实体和尾实体在超平面上的投影向量,分别为  $h_{\perp} = h - w_r^T h w_r$  和  $t_{\perp} = t - w_r^T t w_r$ ,使得  $h_{\perp} + r \approx t_{\perp}$ . 其目标函数如式(13)所示:

$$\mathcal{L} = \sum_{(h,r,t) \in G} \sum_{(h',r',t') \in G'} \left[ \gamma + \|\mathbf{h}_\perp + \mathbf{r} - \mathbf{t}_\perp\|_2^2 - \|\mathbf{h}'_\perp + \mathbf{r}' - \mathbf{t}'_\perp\|_2^2 \right] \quad (13)$$

其中,  $\mathbf{h}_\perp$  和  $\mathbf{t}_\perp$  分别为头实体的表示和尾实体的表示在超平面  $\mathbf{w}_r \in \mathbb{R}^d$  上的投影向量.

TransR 算法<sup>[24]</sup>: 将实体和关系分别嵌入不同的向量空间. 对于描述正确事实的三元组  $(h, r, t)$ , 算法将头实体的向量表示  $\mathbf{h} \in \mathbb{R}^d$  和尾实体的向量表示  $\mathbf{t} \in \mathbb{R}^d$  通过投影矩阵  $\mathbf{M}_r \in \mathbb{R}^{d \times k}$ , 投影到关系向量  $\mathbf{r} \in \mathbb{R}^d$  所在的空间, 使得  $\mathbf{M}_r \mathbf{h} + \mathbf{r} \approx \mathbf{M}_r \mathbf{t}$ . 其目标函数如式(14)所示:

$$\mathcal{L} = \sum_{(h,r,t) \in G} \sum_{(h',r',t') \in G'} \left[ \gamma + \|\mathbf{M}_r \mathbf{h} + \mathbf{r} - \mathbf{M}_r \mathbf{t}\|_2^2 - \|\mathbf{M}_r \mathbf{h}' + \mathbf{r}' - \mathbf{M}_r \mathbf{t}'\|_2^2 \right] \quad (14)$$

其中,  $\mathbf{M}_r \in \mathbb{R}^{d \times k}$  为实体特征空间到关系特征空间的转换矩阵.

RDF2VEC 算法<sup>[25]</sup>: 首先在知识图谱上进行随机游走生成游走序列, 然后采用 Word2vec 算法训练词向量的方式, 根据游走序列中的上下文预测目标实体或目标关系, 最终得到知识图谱中实体和关系的向量表示.

FML + TML+Hybrid 算法<sup>[14]</sup>: 首先根据数据集构建真实信息的知识图谱和虚假信息知识图谱, 然后分别采用 TransE 算法训练真实信息的知识图谱和虚假信息知识图谱中的实体和关系表示, 并通过融合策略将学习到最大偏移向量进行拼接, 最后根据拼接的向量表示进行分类, 将分类的结果作为检测文本真假性的依据.

PN-KG2REC 即本文提出的算法: 首先在知识图谱上进行随机游走生成“实体-关系”的游走序列, 然后使用超矩形来表征具有相似信息的实体的集合, 并通过两种不同的策略生成区域外的实体, 最终得到的实体和关系表示由中心向量和偏移向量两部分组成.

### 5.3 参数设置

本文实验中对比的算法均采用原论文的参数和设置进行复现. 对于 TransE, TransR, TransD, FML + TML+Hybrid 算法以及本文提出的 PN-KG2REC 算法,  $\gamma$  均设为 1.0, 向量维度均设置为 32. 对于本文提出的 PN-KG2REC, 需要设置参数  $\alpha$ , 固定参数  $\gamma = 1$ , 依次设置参数  $\alpha$  为 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0. 当参数  $\alpha$  取 0.8 时, 谣言检测准确率达到最高值 0.895, 效果最好. 此外, PN-KG2REC 算法需要基于知识图谱上的随机游走生成游走序列, 进行  $L$  次游走会得到长度为  $2L+1$  的序列, 依次设置参数  $L$  为 1, 2, 3, 4, 观察游走序列的最大长度对模型分类效果的影响, 如表 2 所示.

由表 2 发现, 在随机游走次数由 1 增加到 4 的过程中, PN-KG2REC 方法的分类准确率逐渐提升. 当最大

表 2 随机游走的最大次数对分类准确率的影响

随机游走的最大次数	准确率
1	0.882
2	0.886
3	0.895
4	0.900
5	0.895

次数由 4 增至 5 时, 分类准确率降低. 由表 3 可以发现, 虽然随着随机游走次数的增加, PN-KG2REC 方法的分类准确率逐渐提升, 但是模型训练过程中所花费的时间也越来越多.

表 3 随机游走的最大次数与训练时长的关系

随机游走的最大次数	训练时间
1	52 s
2	116 s
3	232 s
4	427 s
5	768 s

因此, 为了平衡分类效果与训练时间成本, 下文在比较 PN-KG2REC 方法与其他方法的分类准确率时, 设置随机游走的最大次数为 3.

### 5.4 分类效果分析

表 4 给出了不同算法的分类效果, 发现本文提出的算法 PN-KG2REC 在谣言检测任务上的准确率优于其它算法.

表 4 基于知识图谱的谣言检测方法分类效果比较

方法	准确率
TransE 算法	0.843
TransH 算法	0.851
TransR 算法	0.855
RDF2VEC 算法	0.878
FML + TML+Hybrid 算法	0.868
PN-KG2REC 算法	0.895

由表 4 可知, 将 TransE, TransH, TransR 和 RDF2VEC 算法, 与 PN-KG2REC 算法应用于谣言检测任务后, 在公开数据集上的分类准确率均超过了 0.8, 证明了基于知识图谱表示学习来进行谣言检测是可行的. TransE 算法的分类准确率为 0.843, 表现最差, 这是因为 TransE 算法对事实信息进行嵌入的方式比较简单, 学习到的实体和关系向量表示不能够表征更为丰富的信息. TransR 算法认为实体往往会包含多种属性, 而不同的关系会关注实体的不同属性, 因此 TransR 算法将知识图谱中的实体和关系嵌入不同的向量空间, 并为每一种关系的分配了一个转移矩阵. 相同的实体可以通过不同的转移矩阵投影到关系空间的不同位置. TransR

算法学习得到的实体和关系表示比 TransE 算法具有更好的表征能力,因此基于 TransR 的谣言检测方法比 TransE 算法取得了约 1% 的提升. TransH 算法的分类准确率高于 TransE 算法,低于 TransR 算法. TransH 算法的想法与 TransR 算法类似,区别在于 TransH 算法为每一个关系分配了一个超平面,根据相同的实体在不同超平面上的投影位置不同,来体现不同的关系关注不同的属性.

RDF2VEC 算法的分类准确率高于 TransE, TransH 和 TransR 算法,提升约 2%. 这表明, RDF2VEC 算法通过在随机游走序列上应用 Word2vec 算法,将知识图谱的语义信息和结构信息嵌入实体和关系的表示中,使得实体和关系具有更强的表征能力.

FML+TML+Hybrid 算法分类准确率为 0.868,这是因为 FML+TML+Hybrid 算法同时考虑了真实信息的知识图谱和虚假信息知识图谱. 该方法的分类准确率低于本文算法,这是因为它在对真实信息的知识图谱和虚假信息知识图谱中的实体和关系进行表示时,采用了 TransE 算法,表征能力相对较弱.

本文提出的 PN-KG2REC 算法在谣言检测任务上的准确率比 TransR 算法提升了 4%,比 RDF2VEC 算法提升了 1.7%,比 FML+TML+Hybrid 算法提升了 2.7%. 这是因为 PN-KG2REC 算法首先通过随机游走生成的“实体-关系”序列描述知识图谱中包含的事实信息,然后引入中心向量和偏移向量,使具有相同属性的实体被包含在相同的超矩形内,最后根据两种不同的策略生成区域内和区域外实体,用以模型的训练. PN-KG2REC 算法通过随机游走考虑了知识图谱的语义信息和结构信息,通过引入超矩形体现了对实体不同属性的关注,通过区域外实体的生成策略引入了负样本的信息,因而学习得到的实体和关系表示能包含更为丰富的信息,从而提升了分类效果.

### 5.5 区域外实体选择方法的比较

4.4 节提出两种选择区域外实体的策略,其中策略一基于随机采样的方式生成区域外的实体,策略二根据数据集中的负样本生成区域外的实体. 为了验证策略的有效性,本文分 4 种情况讨论了不同的策略对分类结果的影响,实验结果如表 5 所示.

- (1) 无策略:不使用任何策略选择区域外实体.
- (2) 策略一:只使用策略一选择区域外实体.
- (3) 策略二:只使用策略二选择区域外实体.
- (4) 策略一+策略二:使用策略一和策略二选择区域外实体.

由表 5 可知,在不使用任何策略选择区域外实体的情况下,模型分类准确率只有 0.878,是 4 种情况中最低的. 单独使用策略一或者策略二选择区域外实体,可

以提升模型分类准确率,但提升的幅度有限,使用策略一的模型相较于未使用策略的模型,在分类准确率上提升了 0.3%,使用策略二的模型提升了 0.4%.

表 5 区域外实体的选择策略对分类准确率的影响

策略	正确率
无策略	0.878
策略一	0.881
策略二	0.882
策略一+策略二	0.895

同时使用策略一和策略二的模型,取得了最好的分类效果,分类准确率为 0.895,这表明,通过策略一和策略二选择与区域内实体具有不同属性的区域外实体,有助于 PN-KG2REC 算法更好地理解和区分实体的不同属性,从而使得学习到的实体和关系表示具有更强的表征能力,因而在谣言分类任务上的分类准确率会更好.

## 6 总结和展望

为了尽早地检测谣言,降低其带来的危害,本文提出了一种不依赖用户评论的谣言早期检测算法,并采用公开数据集对算法进行了大量的对比和分析. 实验结果表明,本文提出的模型在谣言检测任务上的准确率比其他最新相关方法更高. 为了验证区域外实体选择策略的有效性,分 4 种情况讨论了不同策略的组合对分类结果的影响. 实验结果表明,综合使用提出的两种策略选择区域外实体,能够取得更好的分类效果. 需要说明的是,本文提出的基于知识图谱表示学习的谣言早期检测方法可以作为现有的谣言检测方法的补充,可以与基于用户评论数据的谣言检测方法配合使用,来提高谣言检测的准确率.

本文提出的谣言检测方法虽然取得了不错的实验效果,但依然存在可改进的空间:(1)借助知识图谱进行谣言检测,需要将帖子描述的内容表示为三元组的形式,因此如何设计合理有效的实体识别和关系抽取算法,对文本中的实体和关系进行正确地提取,还值得深入研究;(2)待检测的帖子所描述的内容可能没有被知识图谱覆盖,需要进一步设计合理的策略来解决这一问题;(3)谣言与反谣言的群体行为传播动力学模型及其对抗博弈是一个融合了社会心理学的研究<sup>[26]</sup>,把本文提出的谣言早期检测模型与之融合,实现对谣言与反谣言群体行为的早期预测,也是一个值得关注的问题.

### 参考文献

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

- [2] ZHAO Z, RESNICK P, MEI Q. Enquiring minds: Early detection of rumors in social media from enquiry posts[C]// Proceedings of the 24th International Conference on World Wide Web. New York: ACM, 2015: 1395-1405.
- [3] OSHIKAWA R, QIAN J, WANG W Y. A survey on natural language processing for fake news detection[C]// Proceedings of the 12th Language Resources and Evaluation Conference. Bertin: ELRA, 2020: 6086-6093.
- [4] KALIYAR R K, GOSWAMI A, NARANG P, et al. FND-Net - a deep convolutional neural network for fake news detection[J]. Cognitive Systems Research, 2020, 61: 32-44.
- [5] ZUBIAGA A, AKER A, BONTCHEVA K, et al. Detection and resolution of rumours in social media: A survey [J]. ACM Computing Surveys (CSUR), 2018, 51(2): 1-36.
- [6] SHARMA K, QIAN F, JIANG H, et al. Combating fake news: A survey on identification and mitigation techniques [J]. ACM Transactions on Intelligent Systems and Technology, 2019, 10(3): 1-42.
- [7] MA J, GAO W, WEI Z, et al. Detect rumors using time series of social context information on microblogging websites[C]// Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. New York: ACM, 2015: 1751-1754.
- [8] MA J, GAO W, MITRA P, et al. Detecting rumors from microblogs with recurrent neural networks[C]// Proceedings of the 25th International Joint Conference on Artificial Intelligence. New York: IJCAI, 2016: 3818-3824.
- [9] MA J, GAO W, WONG K F. Rumor detection on Twitter with tree-structured recursive neural networks[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2018: 1980-1989.
- [10] TAI K S, SOCHER R, MANNING C D. Improved semantic representations from tree-structured long short-term memory 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: ACL, 2015: 1556-1566.
- [11] MIHALCEA R, STRAPPARAVA C. The lie detector: Explorations in the automatic recognition of deceptive language[C]// Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the AFNLP. Stroudsburg: ACL, 2009: 309-312.
- [12] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4): 541-551.
- [13] JI S, PAN S, CAMBRIA E, et al. A survey on knowledge graphs: Representation, acquisition and applications[C]// Proceedings of the 30th AAAI conference on artificial intelligence. Vancouver: AI Access Foundation, 2020: 1-25.
- [14] PAN J Z, PAVLOVA S, LI C, et al. Content based fake news detection using knowledge graphs[C]// Proceedings of the 17th International Semantic Web Conference. Berlin: Springer-Verlag, 2018: 669-683.
- [15] PISTER A, ATEMEZING G A. Knowledge graph embedding for triples fact validation[C]// Proceedings of the 18th International Semantic Web Conference. Berlin: Springer-Verlag, 2019: 21-24.
- [16] PEROZZI B, AL-RFOU R, SKIENA S. Deepwalk: Online learning of social representations[C]// Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2014: 701-710.
- [17] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[C]// Proceedings of 1st International Conference on Learning Representations. Scottsdale: ICLR, 2013: 1-12.
- [18] REN H, HU W, LESKOVEC J. Query2box: Reasoning over knowledge graphs in vector space using box embeddings[C]// Proceedings of the 8th International Conference on Learning Representations. Scottsdale: ICLR, 2020: 1-10.
- [19] HE X, ZHAO K, CHU X. AutoML: A survey of the state-of-the-art[J]. Knowledge-Based Systems, 2021, 212: 106622.
- [20] CHEN T, GUESTRIN C. Xgboost: A scalable tree boosting system[C]// Proceedings of the 22nd ACM International Conference on Knowledge Discovery And Data Mining. New York: ACM, 2016: 785-794.
- [21] ISWC 2019[EB/OL]. (2019)[2021]. <https://dice-group.github.io/semantic-web-challenge.github.io/>.
- [22] BORDES A, USUNIER N, GARCIA-DURAN A, et al. Translating embeddings for modeling multi-relational data [J]. Advances in neural information processing systems, 2013, 26: 2787-2795.
- [23] WANG Z, ZHANG J, FENG J, et al. Knowledge graph embedding by translating on hyperplanes[C]// Proceedings of the 28th AAAI Conference on Artificial Intelligence. Vancouver: AI Access Foundation, 2014: 1112-1119.
- [24] LIN Y, LIU Z, SUN M, et al. Learning entity and relation

embeddings for knowledge graph completion[C]//Proceedings of the 29th AAAI Conference on Artificial Intelligence. Vancouver: AI Access Foundation, 2015: 2181-2187.

- [25] RISTOSKI P, PAULHEIM H. Rdf2vec: Rdf graph embeddings for data mining[C]//Proceedings of the International Semantic Web Conference. Berlin: Springer-Verlag, 2016: 498-514.
- [26] XIAO Y, CHEN D, WEI S. et al. Rumor propagation dynamic model based on evolutionary game and anti-rumor [J]. Nonlinear Dynamics 2019, 95: 523-539.

#### 作者简介



皮德常 男,1971年11月生.南京航空航天大学教授、博士生导师.主要研究方向为大数据、文本挖掘和隐私保护.  
E-mail: dc.pi@nuaa.edu.cn



吴致远 男,1996年6月生.南京航空航天大学硕士研究生.主要研究方向为文本挖掘、图神经网络.  
E-mail: aguiller2018@163.com



曹建军(通讯作者) 男,1975年6月生.国防科技大学第六十三研究所副研究员、硕士生导师.主要研究方向为数据清洗、情报分析.  
E-mail: jianjuncao@yeah.net