

文章编号: 1003-0077(2023)09-0161-14

基于事件-词语-特征异质图的微博谣言检测新方法

王友卫¹, 凤丽洲², 王炜琦¹, 侯玉栋¹

(1. 中央财经大学 信息学院, 北京 100081; 2. 天津财经大学 统计学院, 天津 300222)

摘要: 当前微博谣言检测研究大多基于微博原文、评论内容及其相互关系, 忽略了情感特征、语法特征及语言特征等重要因素的影响。为此, 该文提出了一种基于事件-词语-特征异质图的微博谣言检测新方法。首先, 在传统方法基础上引入情感、语法、心理等方面的知识, 提出文本特征的概念以有效挖掘微博事件中蕴含的情感特征、语法特征以及语言特征。然后, 综合微博评论、文本词语及文本特征对谣言检测结果的影响, 构建用于谣言检测的事件-词语-特征异质图。最后, 利用 GraphSAGE 和异质图注意力网络在节点表达方面的优势提出新的节点信息聚合方法, 以此在区分节点类型重要性的同时降低节点集规模带来的影响。实验结果表明, 该方法能有效提高微博事件表示的准确性; 相对于传统机器学习方法和典型的深度学习方法而言, 该方法在谣言检测精度上具有明显优势。

关键词: 谣言检测; 文本特征; 异质图注意力网络

中图分类号: TP391

文献标识码: A

Weibo Rumor Detection Based on Heterogeneous Graph of Event-Word-Feature

WANG Youwei¹, FENG Lizhou², WANG Weiqi¹, HOU Yudong¹

(1. School of Information, Central University of Finance and Economics, Beijing 100081, China;

2. School of Statistics, Tianjin University of Finance and Economics, Tianjin 300222, China)

Abstract: Most of the current Weibo rumor detection methods are based on the original Weibo texts, comment contents and their interrelationships, ignoring the influence of important factors such as emotional features, lexical features and language features. In this paper, a new method for Weibo rumor detection based on a heterogeneous graph of event-word-feature is proposed. Firstly, the knowledge from the aspects of emotion, lexical and psychology is introduced to effectively mine the emotional features, lexical features and language features of Weibo events. Then, an event-word-feature heterogeneous graph for rumor detection is constructed to combine the effects of Weibo comments, text words and text features. Finally, utilizing the advantages of GraphSAGE and heterogeneous graph attention network, and a new node information aggregation method is proposed to distinguish the importance of node types while reducing the impact of node set size. The experimental results show that the rumor detection accuracy of the proposed method has obvious advantages over that of traditional machine learning based methods and typical deep learning based methods.

Keywords: rumor detection; text features; heterogeneous graph attention network

0 引言

随着社交媒体的发展, 网络谣言给社会带来了严重的影响, 并逐渐引起了公众关注, 成为了国内外

学者的研究热点。党的二十大报告中指出, “健全网络综合治理体系, 推动形成良好网络生态”。可见, 实现谣言检测对于促进网络空间建设、维护社会稳定快速发展具有重大的现实意义。

谣言检测任务通过模型将正常文档与含有谣言

收稿日期: 2022-11-25 **定稿日期:** 2023-04-23

基金项目: 教育部人文社科项目(19YJCZH178); 国家自然科学基金(61906220); 国家社会科学基金(18CTJ008); 中央财经大学新兴交叉学科建设项目

的异常文档区分开,属于文本分类领域中的重要子问题。谣言检测的相关方法可分为三类^[1]:①基于外部知识的方法;②基于关系网络的方法;③基于文本内容的方法。基于外部知识的方法主要利用专家系统或集体智慧对谣言文档进行判别,该方法需要耗费大量的人工成本建立知识图谱,因此相关研究较少。基于关系网络的方法通过消息的传播特点、传播者追加的评论文本以及传播者的社会背景对谣言进行检测^[2]。但是,此类方法的检测准确度与传播时间成正比,无法在谣言传播的初期对其进行较好的识别。基于文本内容的方法认为谣言与非谣言在表达习惯、讨论主题以及行文风格上存在一定差别,因此可以通过从文本中提取可供分类的向量化信息实现谣言检测。基于文本内容的检测方法可进一步分为基于传统分类器的检测方法与基于深度学习的检测方法两类。前者方法通过匹配文本中出现的人工特征来构建文本内容的 one-hot 向量,之后将特征向量输入支持向量机(Support Vector Machine, SVM)、随机森林(Random Forest, RF)、逻辑回归(Logistic Regression, LR)等机器学习模型,以此实现对谣言信息的识别。此类方法中最常用的特征是文本极性以及组合特征,如通过人工词典构建的情绪特征和语言学家构建的句式特征等^[3]。但此类方法在提取特征时单纯依靠规则或者人工经验,因此相对于基于深度学习的检测方法而言无法较好地表征文本中的潜在语义信息。

近年来,基于深度学习的文本分类方法已被广泛应用于谣言检测任务中。Nguyen等^[4]与Singh等^[5]分别将谣言数据的向量化结果带入到卷积神经网络(Convolutional Neural Network, CNN)与循环神经网络(Recurrent Neural Network, RNN)中,相对于传统分类器而言有效提高了谣言检测的精度。Ma等^[6]提出了一种基于RNN的评论上下文学习方法,通过学习微博事件的连续表示来实现谣言信息检测。Song等^[7]和王友卫等^[8]将所有转发信息视为一个序列,通过CNN实现可信早期谣言检测研究,有效缩短了谣言检测的时间跨度。Chen等^[9]结合知识图谱提出一种基于图的谣言文本生成模型G2S-AT-GAN。该模型使用基于注意力的图卷积神经网络(Graph Convolutional Network, GCN)和生成对抗网络(Generative Adversarial

Network, GAN)来生成不同主题的谣言文本,在解决谣言数据不平衡问题的同时提高了谣言检测的性能。但上述方法大多仅关注源信息和评论信息的内容,并没有有效利用评论、用户之间的相互关系,为此,杨延杰等^[10]利用消息转发关系构建评论转发图,通过两个融合门控机制的图卷积网络模块来聚合邻居节点信息以生成节点的表示,有效利用了源博文的影响力与任意帖子之间的多角度影响。Wu等^[11]提出了基于图神经网络全局嵌入的谣言检测模型和基于图神经网络集成学习的谣言检测模型。通过在有限时间步内在相邻节点之间交换信息来更新节点表示,有效提高了谣言检测的准确性。Bian等^[12]提出了一种双向图卷积网络(Bi-GCN),根据谣言的转发关系建立自上而下和自下而上的图传播结构,较好地实现了谣言信息的嵌入表达,不足之处在于图中每个评论节点只能融合一个传播方向上的邻居节点信息。为了更好地考虑用户信息的影响,Lu和Li通过引入文本发布者的社交关系进行谣言检测^[13]。Zhang等^[14]借助立场检测任务,提出了一种基于多模态融合和元知识共享的谣言检测方法。该方法使用注意力机制计算评论权重,较好地区分了不同评论的重要性。

通过研究发现,虽然上述方法已获得较好的谣言检测效果,但仍存在以下问题:①大多数方法在利用评论文本内容时仅关注词语特征信息,忽略了词语情感特征、语法特征、语言特征等重要因素的影响;②现有算法普遍根据原始博文和转发评论之间的关系建立图结构,忽略了原始评论之间的语义关联性,因此难以针对新发布的博文进行检测,限制了模型的泛化能力。

为解决上述问题,本文以微博为研究对象,利用图神经网络在图表示学习任务方面的优势,提出了基于事件-词语-特征异质图的微博谣言检测新方法RD_EWF。具体而言,本文创新点如下:

(1) 综合考虑情感特征、语法特征以及语言特征对于谣言检测的影响,在评论内容信息基础上提出文本特征的概念。在此基础上,将微博事件、文本词语、文本特征作为节点构建事件-词语-特征异质图,解决了传统方法单纯利用文本内容信息导致的模型表达能力不足的问题。

(2) 综合考虑事件-事件之间、评论-词语之间、评论-特征之间以及词语-词语之间的相互作用,提

出基于 GraphSAGE 和异质图注意力网络(Heterogeneous Graph Attention Network, HGAT)的图节点表示学习方法 GS_HGAT, 以此区分不同类型节点的影响, 实现对微博事件节点的归纳式表达, 提高模型的泛化能力。

1 相关理论

1.1 图神经网络(Graph Neural Network, GNN)^[15]

GNN 是被广泛应用于图分析任务的一类神经网络, 现已广泛应用于社交网络、推荐系统、生物科技等领域。给定一个属性图 $G=(V, E)$ (V 为节点集、 E 为边集) 及其特征矩阵 $\mathbf{X}=\{\mathbf{x}_i\}$, 其中 \mathbf{x}_i 是节点 $v_i \in V$ 的 d 维特征向量, GNN 的目标是学习每个节点 v_i 的表示 \mathbf{h}_i , 考虑第 m 层 GNN, 节点 v_i 在第 m 层的表示向量定义为:

$$\mathbf{h}_i^m = \text{Update}(\mathbf{h}_i^{m-1}, \text{agg}(\{\mathbf{h}_j^{m-1} : v_j \in N(v_i)\})) \quad (1)$$

其中, \mathbf{h}_i^{m-1} 是节点 v_i 在第 $m-1$ 次迭代后的特征向量, 其中 $\mathbf{h}_i^0 = \mathbf{x}_i$, $N(v_i)$ 是 v_i 的邻居节点集, agg 和 Update 分别是 GNN 层的信息聚合函数和更新函数。对于节点级任务, 可直接将节点表示 \mathbf{h}_i^m 用于下游任务。对于图级任务, 如式(2)所示, 可通过输出函数 Readout 聚合节点特性以获得整个图的表示 \mathbf{h}_G 。

$$\mathbf{h}_G = \text{Readout}(\mathbf{h}_i^m \mid v_i \in V) \quad (2)$$

1.2 中文语言查询和词数统计(Chinese Linguistic Inquiry and Word Count, C_LIWC)^[16]

C_LIWC 词典是在语言查询和词数统计词典(Linguistic Inquiry and Word Count, LIWC)基础上形成的词典。Pennebaker 等人研究建立的 LIWC 词典主要用于对文本描述中的单词进行统计分析。该词典包含约 4 500 个从社会学、健康学以及心理学方面挖掘的情绪和认知方面单词, 现已成为英文情绪分析研究应用的重要依据。台湾科技大学人文社会学科研究人员根据中文特性将 LIWC 词典翻译改编为中文版本 C-LIWC。C-LIWC 包含语言特征 30 类(如副词、介词等)、心理特性 42 类(如正向情绪词、负向情绪词等), 共 72 个类别、6 862 个词。其中, 与情绪相关的类别包括 positive emotion、negative emotion、anxiousness、anger 和 sad-

ness。C_LIWC 词典中每个词都有一个或多个类别属性, 如“担忧”同时属于 Negative Emotion 类和 Anxiousness 类。

2 研究方法

2.1 问题描述

以微博为研究对象, 相关定义如下^[8]:

定义 1 微博事件 定义微博事件集合 $E = \{E_i\} (0 \leq i < M, M \text{ 为微博事件个数})$ 。其中, $E_i = \{m_{i,j}\} (0 \leq j \leq n_i)$ 表示第 i 个微博事件, 其中包含源微博 $m_{i,0}$ 及其评论 $m_{i,j} (1 \leq j \leq n_i, n_i \text{ 为源微博对应的评论数})$ 。

定义 2 源微博 源微博是指最开始发出的微博, 该微博不回复其他任何微博。 $m_{i,0}$ 表示第 i 个微博事件的源微博。

定义 3 评论 评论是指直接回复源微博或回复与源微博相关微博的微博。 $m_{i,j} (1 \leq j \leq n_i)$ 表示第 i 个微博事件的第 j 条评论。

给定微博事件 E_i , 本文方法的目的是学习一个分类模型 CM 以输出 E_i 是否为谣言的判定结果, 即: $y = \text{CM}(E_i, \theta)$ (θ 为参数集)。如果 $y = 1$, 说明 E_i 为谣言事件, 否则 E_i 为正常事件。

2.2 方法描述

首先, 对微博语料库中的微博事件进行中文分词, 获得其中所有的事件及其对应的词语; 然后, 引入情感、语法、心理等方面知识, 构建文本特征集, 在此基础上挖掘事件-事件之间、事件-词语之间、词语-特征之间以及词语-词语之间的相互作用, 构建事件-词语-特征异质图; 最后, 为区分不同类型节点的影响, 提出基于 GraphSAGE 和异质图注意力网络的节点聚合方法, 以此获得事件节点的向量表达。本文方法 RD_EWF 执行流程如图 1 所示。其中, M 为事件总数, N 为词语总数, S 为特征总数。具体介绍如下:

2.2.1 文本特征构建

为了综合考虑不同类型特征对谣言检测结果的影响, 本文将构建以下三种文本特征: 情感特征、语法特征以及语言特征。具体如下:

(1) 情感特征

情感特征来自 DUTIR 实验室构建的 Emotion Ontology 情感词典^[17]。该词典是大连理工大学信

息检索研究室在林鸿飞教授的指导下整理和标注的一个中文本体资源,从不同角度描述一个中文词汇或者短语,包括词语词性种类、情感类别、情感强度及极性等信息。如表 1 所示,该词典将词语情感共

分为乐、好、怒等 7 个大类,21 小类,共计 27 466 个词语。为了体现不同情感倾向对于谣言检测结果的影响,本文使用该词典中的 21 个情感小类作为情感特征集 F_s 。

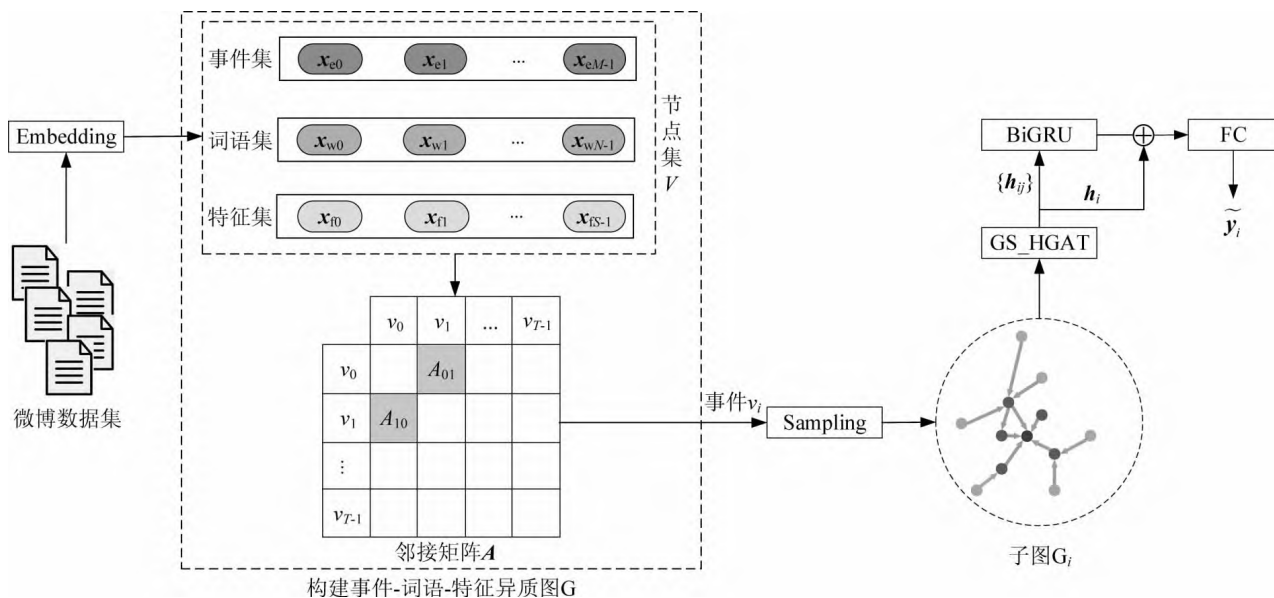


图 1 RD_EWF 执行流程

表 1 Emotion Ontology 情感词典

编号	大类	情感类	例词
1	乐	快乐(PA)	喜悦、欢喜、笑咪咪、欢天喜地
2		安心(PE)	踏实、宽心、定心丸、问心无愧
3	好	尊敬(PD)	恭敬、敬爱、毕恭毕敬、肃然起敬
4		赞扬(PH)	英俊、优秀、通情达理、实事求是
5		相信(PG)	信任、信赖、可靠、毋庸置疑
6		喜爱(PB)	倾慕、宝贝、一见钟情、爱不释手
7		祝愿(PK)	渴望、保佑、福寿绵长、万寿无疆
8	怒	愤怒(NA)	气愤、恼火、大发雷霆、七窍生烟
9	哀	悲伤(NB)	忧伤、悲苦、心如刀割、悲痛欲绝
10		失望(NJ)	憾事、绝望、灰心丧气、心灰意冷
11		疚(NH)	内疚、忏悔、过意不去、问心有愧
12		思(PF)	思念、相思、牵肠挂肚、朝思暮想
13	惧	慌(NI)	慌张、心慌、不知所措、手忙脚乱
14		恐惧(NC)	胆怯、害怕、担惊受怕、胆颤心惊
15		羞(NG)	害羞、害臊、面红耳赤、无地自容

续表

编号	大类	情感类	例词
16	恶	烦闷(NE)	憋闷、烦躁、心烦意乱、自寻烦恼
17		憎恶(ND)	反感、可耻、恨之入骨、深恶痛绝
18		贬责(NN)	呆板、虚荣、杂乱无章、心狠手辣
19		妒忌(NK)	眼红、吃醋、醋坛子、嫉贤妒能
20		怀疑(NL)	多心、生疑、将信将疑、疑神疑鬼
21	惊	惊奇(PC)	奇怪、奇迹、大吃一惊、瞠目结舌

(2) 语法特征

对于信息含量较少的微博文本而言,词性、实体等语法特征可能包含人物、地点、时间、机构、数量、方位等多类信息,其中的部分信息可以作为检测谣言文本的重要依据。一般而言,信息描述越模糊,其为谣言的可能越高。例如,“某大学教授称……”“某知名人士称……”“研究者发现……”等文本并没有明确的人物、机构等名称,因此它们为谣言的可能性较大。中文词法分析(Lexical Analysis of Chinese, LAC)模型^[18]是百度研发的一款联合的词法分析工具,能有效实现中文分词、词性标注、专名识别等功能。鉴于 LAC 模型在处理中文文本方面的优势,

本文使用该模型获得词性标签 24 个、专名实体类别标签 4 个,以此构建语法特征集 F_g 。

(3) 语言特征

语言特征由评论中与用户心理、情绪、认知、用词习惯等相关的词语所归属的类别构成。此类特征能较好地反映用户的情绪变化、心理状态、用词习惯等特点,因此包含与谣言检测相关的重要信息。本文根据 C-LIWC 中文语言分析工具^[16],将其整理的 6 862 个词语所归属的 72 个类别作为语言特征集 F_l 。部分语言特征及其代表性词语如表 2 所示。

表 2 部分语言特征

序号	语言特征	代表性词语
1	posemo	喜欢、热爱
2	negemo	讨厌、悲伤
3	certain	一定、一直
4	incl	而且、额外
5	excl	不管、反而

续表

序号	语言特征	代表性词语
6	percept	抓住、看见
7	negate	不行、不许
8	relativ	至今、那时
9	swear	白痴、坏蛋
10	personal concerns	工作、家庭

2.2.2 谣言检测异质图构建

首先,从训练集中获取事件集 E 、词语集 W 、文本特征集 $F = \{F_s, F_g, F_l\}$ 。在此基础上,建立由 E 、 W 、 F 构成的异质图 $G = \{V, A\}$, A 为 G 中节点对应的邻接矩阵,如图 2 所示,其中, $V = \{E, W, F\}$, $v_{ea} \in E (0 \leq a < 5)$ 、 $v_{wb} \in W (0 \leq b < 5)$ 、 $v_{fc} \in F (0 \leq c < 5)$ 。针对 V 中任意节点对 $v_i, v_j (0 \leq i, j < T, T$ 为图中节点总数),将根据 v_i, v_j 的类型确定它们之间是否存在连边以及对应连边的权重 A_{ij} 。具体如下:

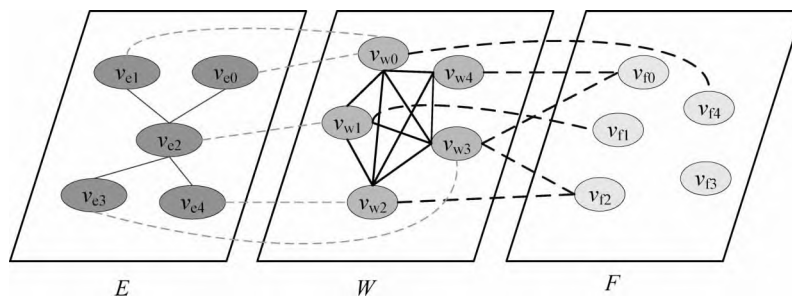


图 2 谣言检测异质图示意

(1) 若 $v_i \in E$ 并且 $v_j \in E$: 根据它们之间的语义相似性建立连边(图 2 中细实线所示),连边权重 A_{ij} 为:

$$A_{ij} = A_{ji} = 1 - \frac{\arccos(\text{cossim}(x_i, x_j))}{\pi} \quad (3)$$

$$\text{cossim}(x_i, x_j) = \frac{x_i \cdot x_j}{\|x_i\| \|x_j\|} \quad (4)$$

$$x_i = \frac{1}{n_{wi}} \sum_{w_k \in v_i} x_{wk}, \quad x_j = \frac{1}{n_{wj}} \sum_{w_k \in v_j} x_{wk} \quad (5)$$

其中, cossim 为余弦相似度函数, \arccos 为反余弦函数, x_{wk} 为词语 w_k 对应的词向量, x_i, x_j 分别为 v_i, v_j 两个事件中所有词语的词向量均值, n_{wi}, n_{wj} 分别为 v_i, v_j 中的词语数。

(2) 若 $v_i \in E$ 并且 $v_j \in W$: 如果 v_j 出现在 v_i 中,则在 v_i, v_j 之间建立连接(图 2 中细虚线所示),权重 A_{ij} 为 v_j 在 v_i 中的归一化 TF-IDF 值,即:

$$A_{ij} = \frac{n_{ij}}{n_{wi}} \times \log \frac{M}{n_{cj}} \quad (6)$$

其中, n_{ij} 为词语 v_j 在事件 v_i 中出现的数量, n_{cj} 为词语 v_j 出现的事件数, M 为事件总数。

(3) 若 $v_i \in W$ 并且 $v_j \in W$,则根据它们的共现情况建立连边(图 2 中粗实线所示)。采用点互信息(Pointwise Mutual Information, PMI)^[19]来计算语料库中词语和词语之间的连接权重 A_{ij} ,定义如下:

$$A_{ij} = A_{ji} = \log_2 \frac{p(v_i, v_j)}{p(v_i)p(v_j)} \quad (7)$$

$$p(v_i) = \frac{n_{ci}}{M}, \quad p(v_j) = \frac{n_{cj}}{M} \quad (8)$$

$$p(v_i, v_j) = \frac{n_{cij}}{M} \quad (9)$$

其中, $p(v_i)$ 为词语 v_i 在事件中出现的概率, $p(v_i, v_j)$ 为词语 v_i 与词语 v_j 在事件中同时出现

的概率, n_{cij} 为词语 v_i, v_j 同时出现的事件数量。

(4) 若 $v_i \in W$ 并且 $v_j \in F$: 如果 v_i 属于 v_j 对应的词语集, 则在 v_i, v_j 之间建立连边(图 2 中粗虚线所示), 令 $A_{ij} = A_{ji} = 1$ 。

2.2.3 节点嵌入

针对训练集中的任意词语 $v_i \in W$, 本文将 Word2Vector^[20] 词嵌入方法得到的结果 $\mathbf{x}_{w_i} \in \mathbb{R}^{1 \times d}$ (d 为嵌入维度) 作为 v_i 的初始化嵌入向量; 针对文本特征 $v_j \in F$, 将与该特征对应的维度为 $|F|$ ($|F|$ 表示集合 F 中元素数量) 的独热向量 $\mathbf{x}_{f_j} \in \mathbb{R}^{1 \times |F|}$ 作为其初始化嵌入向量; 针对事件 $v_k \in E$, 由于其包括原微博 v_{k0} 和相关评论 $\{v_{kl} \mid 1 \leq l < n_k\}$, 考虑到不同评论对于原微博的贡献是不同的, 因此这里先采用软注意力机制获得 v_{kl} 相对于 v_{k0} 的权重 α_l^k , 如式(10)所示。

$$\alpha_l^k = \frac{\exp(\sigma(\lceil \mathbf{e}_{k0} \parallel \mathbf{e}_{kl} \rceil \mathbf{W}_{e1}))}{\sum_{1 \leq m \leq n_k} \exp(\sigma(\lceil \mathbf{e}_{k0} \parallel \mathbf{e}_{km} \rceil \mathbf{W}_{e1}))} \quad (10)$$

其中, \exp 为以 e 为底的指数函数, σ 为 Relu 激活函数, \parallel 为向量拼接操作, \mathbf{W}_{e1} 为参数矩阵, \mathbf{e}_{k0} 为原微博 v_{k0} 中所含词语向量的均值向量, \mathbf{e}_{kl} 为原微博的第 l 条评论中所含词语向量的均值向量。在此基础上, 获得事件 v_k 对应的节点向量 \mathbf{x}_{e_k} , 如式(11)所示。

$$\mathbf{x}_{e_k} = \sigma(\lceil \mathbf{e}_{k0} \parallel \sum_{1 \leq l \leq m} \alpha_l^k \mathbf{e}_{kl} \rceil \mathbf{W}_{e2}) \quad (11)$$

其中, \mathbf{W}_{e2} 为参数矩阵。由于 \mathbf{x}_{f_j} 与 $\mathbf{x}_{w_i}, \mathbf{x}_{e_k}$ 维度不同, 进一步通过前馈神经网络将其映射到一个维度相同的空间内, 即:

$$\mathbf{x}_{fj} = \sigma(\mathbf{x}_{fj} \mathbf{W}_{e3} + \mathbf{b}_f) \quad (12)$$

其中, $\mathbf{W}_{e3} \in \mathbb{R}^{1 \times |F| \times d}, \mathbf{b}_f \in \mathbb{R}^{1 \times d}$ 为训练参数矩阵。

2.2.4 节点采样与聚合

由于 2.2.2 节所构建的谣言检测异质图 G 规模较大, 直接在该图上使用节点分类算法将面临计算开销较大的问题。GraphSAGE 算法^[21] 首先通过采样邻居的策略, 将节点训练由全图训练方式转换为以节点为中心的小批量训练方式, 使得大规模图数据的分布式训练成为可能。此外, GraphSAGE 对邻居节点的聚合操作进行了拓展, 提出平均聚合、LSTM 聚合、池化聚合等方法以提高节点表达的准确性。本文利用 GraphSAGE 的上述优势, 通过采样得到图 G 的子图实现一种高效、可归纳的谣言检测过程。如图 3 所示, 针对每个事件节点 v_i , 首先, 在一阶($k=1$) 采样过程中我们将获取 v_i 邻居中的全部词语节点并获得 v_i 邻居中连边权重最大的 n_s ($n_s=10$) 个事件节点; 然后, 为控制节点集规模, 在

第二、三阶采样中, 分别针对 v_i 的一阶采样结果中的每个节点 v_j , 通过随机采样方法获得 v_j 的 n_s ($n_s=10$) 个邻居节点; 最后, 利用上述采样所得节点集 $V_i = \{V_{ij}\}$ (V_{ij} 为在第 j 阶采样过程中得到的节点集) 及对应邻接矩阵 \mathbf{A}_i 构建子图 G_i 。

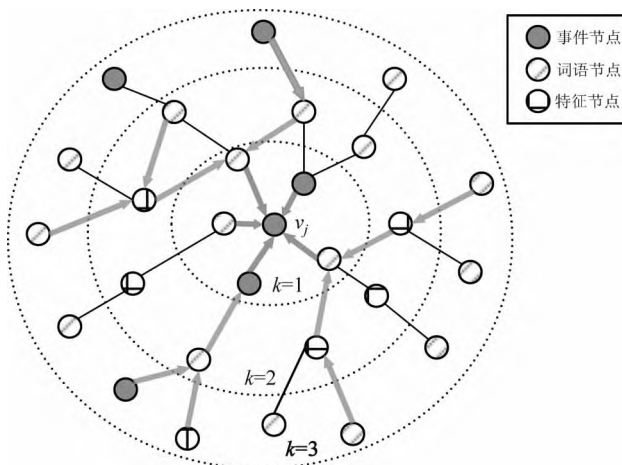


图 3 RD_EWF 的节点采样及聚合

如图 3 所示, RD_EWF 采样方向按照阶段 $k=1, 2, 3$ 依次向外, 而节点聚合过程则与采样方向相反。由于每个节点的邻居节点可能类型不同(如事件节点的邻居可能为词语节点或者事件节点, 而词语节点的邻居可能为词语节点、事件节点或者特征节点), HGAT^[22] 通过异质图注意力网络来考虑不同类型信息的异构性, 并利用双层注意力机制捕获不同邻居节点和不同节点类型对特定节点的重要性。但是, 该方法在所有节点上进行训练, 因此难以适用于规模较大的图结构数据。为此, 本文在 GraphSAGE 基础上结合 HGAT 来为不同类型节点连边赋予注意力权重, 以此在提高节点计算效率的同时区分不同邻居节点对当前节点的影响。在子图 G_i 的第 k 层聚合过程中, 本文基于 GraphSAGE 和 HGAT 的节点聚合过程(GS_HGAT)描述如下:

(1) 按照式(13)计算第 $k-1$ 层节点 v_p^{k-1} 在第 k 层的第 q 个邻居节点 v_{pq}^k 对应的注意力权重 α_{pq}^{k-1} 。

$$\alpha_{pq}^{k-1} = \frac{\exp(\sigma(\lceil \mathbf{h}_p^{k-1} \parallel \mathbf{h}_{pq}^k \rceil \mathbf{W}_n))}{\sum_{v_{pq}^k \in N_k(v_p^{k-1})} \exp(\sigma(\lceil \mathbf{h}_p^{k-1} \parallel \mathbf{h}_{pq}^k \rceil \mathbf{W}_n))} \quad (13)$$

其中, \mathbf{h}_{pq}^k 为 v_{pq}^k 的隐状态向量, \mathbf{h}_p^{k-1} 为 v_p^{k-1} 的隐状态向量, $N_k(v_p^{k-1})$ (为节点 v_p^{k-1} 在图 G_i 中第 k 层的邻居节点集, \mathbf{W}_n 为训练参数矩阵。

(2) 针对第 r ($0 \leq r < R, R$ 为节点类型数量, 本文 $R=3$) 种类型的邻居节点, 按照式(14)计算第 k

-1 层节点 v_p^{k-1} 在第 k 层的第 r 类邻居节点对应的注意力权重 β_{pr}^{k-1} 。

$$\beta_{pr}^{k-1} = \frac{\sigma([\mathbf{h}_p^{k-1} \parallel \mathbf{h}_{pr}^{k-1}]\mathbf{W}_{r1})}{\sum_{l=0, \dots, R-1} \sigma([\mathbf{h}_p^{k-1} \parallel \mathbf{h}_{pl}^{k-1}]\mathbf{W}_{r1})} \quad (14)$$

$$\mathbf{h}_{pr}^{k-1} = \sigma((\mathbf{h}_p^{k-1} \parallel \sum_{v_{pq}^k \in N(v_p^{k-1}), \text{type}(v_{pq}^k)=r} \alpha_{pq}^{k-1} \mathbf{h}_{pq}^k) \mathbf{W}_{r2}) \quad (15)$$

其中, \mathbf{W}_{r1} 、 \mathbf{W}_{r2} 为训练参数矩阵, $\text{type}(v_{pq}^k)$ 获取节点 v_{pq}^k 的类型。

(3) 综合不同类型节点的影响, 按照式(16)实现节点 v_p^{k-1} 对应的隐含状态更新。

$$\mathbf{h}_p^{k-1} = \sigma((\mathbf{h}_p^{k-1} \parallel \sum_{r=0, \dots, R-1} \beta_{pr}^{k-1} \sum_{v_{pq}^k \in N(v_p^{k-1}), \text{type}(v_{pq}^k)=r} \alpha_{pq}^{k-1} \mathbf{h}_{pq}^k) \mathbf{W}_h) \quad (16)$$

其中, \mathbf{W}_h 为训练参数矩阵。

2.2.5 谣言分类

鉴于双向门控循环单元(Bidirectional Gated Recurrent Unit, BiGRU)^[23]能有效提取文本的上下文依赖关系, 为此, 当 $k=1$ 时, RD_EWF 将事件节点 v_i 的邻居中所有词语节点按照其在 v_i 中出现的顺序输入 BiGRU 中以融合词语的序列化信息, 并进一步结合节点 v_i 在聚合过程产生的隐状态向量 \mathbf{h}_i^0 获得其最终的文本嵌入向量 \mathbf{h}_i 。具体定义如式(17)所示。

$$\mathbf{h}_i = \sigma([\mathbf{h}_i^0 \parallel \text{BiGRU}(\{\mathbf{h}_{ij}\})] \mathbf{W}_o) \quad (17)$$

其中, \mathbf{W}_o 为训练参数矩阵, \mathbf{h}_{ij} 为 v_i 的第 j 个邻居节点 v_{ij} 的隐状态向量, $\text{BiGRU}(\{\mathbf{h}_{ij}\})$ 函数输出词语序列 $\{v_{ij}\}$ 对应的句向量。在此基础上, 将 \mathbf{h}_i 输入到全连接层中, 并结合其实际类别及交叉熵函数来使损失最小化, 如式(18)、式(19)所示。

$$\tilde{\mathbf{y}}_i = \text{Softmax}(\mathbf{W}\mathbf{h}_i + \mathbf{b}) \quad (18)$$

$$\text{Loss} = -\frac{1}{\text{Batch_size}} \sum_{v_i \in E_u} \sum_{j=1}^2 \quad (19)$$

$$(y_{ij} \log \tilde{y}_{ij} + (1 - y_{ij}) \log (1 - \tilde{y}_{ij}))$$

其中, \mathbf{W} 为训练参数矩阵, E_u 为第 u 个批事件集, batch_size 为 E_u 中事件数量, y_{ij} 为事件 v_i 的真实类别向量对应的第 j 维的值, \tilde{y}_{ij} 为事件 v_i 的预测向量 $\tilde{\mathbf{y}}_i$ 对应第 j 维的值。

可见, 为提高模型针对图节点的学习能力, RD_EWF 综合考虑了事件、词语、特征三类节点之间的相互影响, 通过引入情感、语法、心理等方面的知识, 解决传统方法单纯利用评论词语信息导致的模型表达能力不足的问题。此外, 为保证模型的高效性及

针对新评论的学习能力, 本文通过随机采样构建谣言检测子图, 利用基于 GraphSAGE 和 HGAT 的节点聚合方法(GS_HGAT)在区分不同类型节点影响的同时提升了模型的泛化学习能力。

3 实验结果与分析

3.1 实验设置

如表 3 所示, 本文使用 Rumdect^[24] 和 CED^[7] 两个公开数据集验证模型的有效性, 采用 7:1:2 的比例将数据集切分为训练集、验证集与测试集。

表 3 实验数据集

详细信息	Rumdect	CED
事件数	4 664	3 387
谣言事件数	2 313	1 538
非谣言事件数	2 351	1 849
用户数量	2 746 818	1 067 410

为验证 RD_EWF 在谣言检测领域的有效性, 将其与 11 个典型基准方法进行对比, 具体包括: ①传统机器学习方法: 朴素贝叶斯(Naive Bayes, NB)^[25]、逻辑回归(Logistic Regression, LR)^[26] 以及随机森林(Random Forest, RF)^[26]; ②基于深度学习的方法: FastText^[27]、TextCNN^[28]、BiGRU-CNN^[29]、TextGCN^[30]、TextING^[31]、I-BERT-LSTM^[32]、dEFEND^[33]、Bi-GCN^[12]、RumorGCN^[34] 及 GLAN^[35]。实验参数设定如下: 节点丢弃率 Dropout_rate = 0.5, 轮次 Epoch = 100, 学习率 Learning_rate = 0.005, 批大小 Batch_size = 100, 词向量维度 Word_dim = 256。为了避免实验误差, 针对每种方法取 50 次实验平均值作为最终的实验结果。

上述方法的参数设置如表 4 所示。

表 4 参数设置

模型名称	模型参数
NB	平滑因子 $\alpha=1$
LR	正则化系数 $C=1$
模型名称	模型参数
RF	决策树数 $n=10$ 最小分裂数 min_split_num = 2
FastText	词窗大小 window_size = 5 负采样个数 negative_num = 5

续表

模型名称	模型参数
TextCNN	卷积核尺寸 filter_size=2,3,4
BiGRU-CNN	卷积核数量 num_filters=256
TextGCN	滑动窗 window=3,
TextING	层数 layer=2
I-BERT-LSTM	BERT 版本 BERT-base-uncased, 词向量 维度 word_dim = 768
dEFEND	注意力层维度 attention_dim=100, 共注 意力层维度 latent_dim = 200
Bi-GCN	层数 layer=2, 丢弃率 dropping_rate=0.2
RumorGCN	层数 layer=2, 丢弃率 dropping_rate=0.2, 相似度阈值 th=0.7
GLAN	卷积核尺寸 filter_sizes 分别等于 3、4、5, 卷积核数量 num_filters=100, 头数 head_num=8

3.2 评价指标

本文采用准确率(Accuracy)与 F_1 值来衡量谣言检测方法的分类效果,定义如式(20)、式(21)所示^[35]。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (20)$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (21)$$

式(20)中 TP 是预测为谣言且实际为谣言的样

本数, FN 是预测为非谣言但实际为谣言的样本数, FP 是实际为非谣言但被预测为谣言的样本数, TN 是实际为非谣言且被预测为非谣言的样本数。式(21)中精确率(Precision)与召回率(Recall)指标定义如式(22)、式(23)所示。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (22)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (23)$$

3.3 隐藏层维度取值影响

为获得最优的隐藏层维度 d , 分别令 $d=128$ 、256、512、1 024, 并统计 RD_EWF 在 Rumdetect 与 CED 数据集上对应的 Accuracy 值和 F_1 值, 结果如图 4 所示。由图知, 当隐藏层的维度小于 256 时, 本文对应的 Accuracy 值和 F_1 值均呈现上升趋势; 当 $d=256$ 时, 本文在 Rumdetect 数据集上获得最大 Accuracy 值(0.938)和 F_1 值(0.934), 在 CED 数据集上获得最大 Accuracy 值(0.916)和 F_1 值(0.914); 当隐藏层的维度大于 256 时, 模型对应的结果呈现下降趋势。究其原因, 过低的嵌入维度可能使得隐藏向量包含的特征信息较少, 导致模型的特征表达能力不足, 而过高的嵌入维度将使得模型出现过拟合或者欠拟合问题, 继而降低了算法的分类性能。由于当 $d=256$ 时本文获得最高的 Accuracy 值和 F_1 值, 因此这里设定默认隐藏层维度为 256。

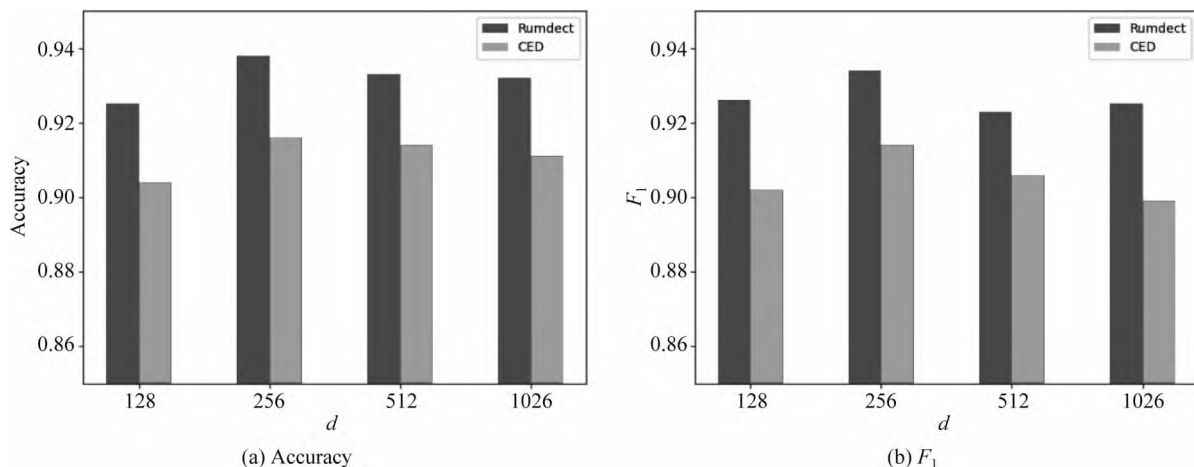


图 4 隐藏层维度的影响

3.4 图节点聚合方法比较

为验证本文提出的基于 GraphSAGE 和 HGAT 的图节点聚合方法(GS_HGAT)在提升节

点表示方面的有效性, 这里将其与以下两种方法进行对比:

(1) **GS**: 使用本文采样方法获得节点子图, 然后直接使用基于 LSTM 聚合器的 GraphSAGE 算

法^[21]生成图节点表达。

(2) **GS_GAT**: 使用本文采样方法获得节点子图, 然后使用图注意力网络(Graph Attention Network, GAT)^[36]生成图节点表达。

在此基础上, 我们将上述不同方法在不同数据集上进行比较, 当采样邻居节点数量 n_s 取 2、4、6、 \dots 、20 时统计不同方法对应的 Accuracy 值和 F_1 值, 结果如图 5、图 6 所示。由图知, 随着 n_s 值的增大, 不同方法对应的结果均呈现出逐渐增加的趋势, 原因在于在图节点信息聚合过程中采样更

多的邻居节点能够丰富节点语义信息, 提高节点特征表示的完整性。但是, 随着 n_s 值继续增加, 不难发现上述方法在不同数据集上的表现均呈现下降趋势, 例如当使用 CED 数据集时, GS 方法在 $n_s=12$ 时取得最大的 Accuracy 值和 F_1 值, 但当 $n_s=20$ 时, 该方法对应结果下降了超过 0.01。可见, 适当增加 n_s 值能提高节点表示学习的准确性, 但是 n_s 值过大容易带来较多的冗余特征信息, 在提高算法计算开销的同时降低信息聚合效果。

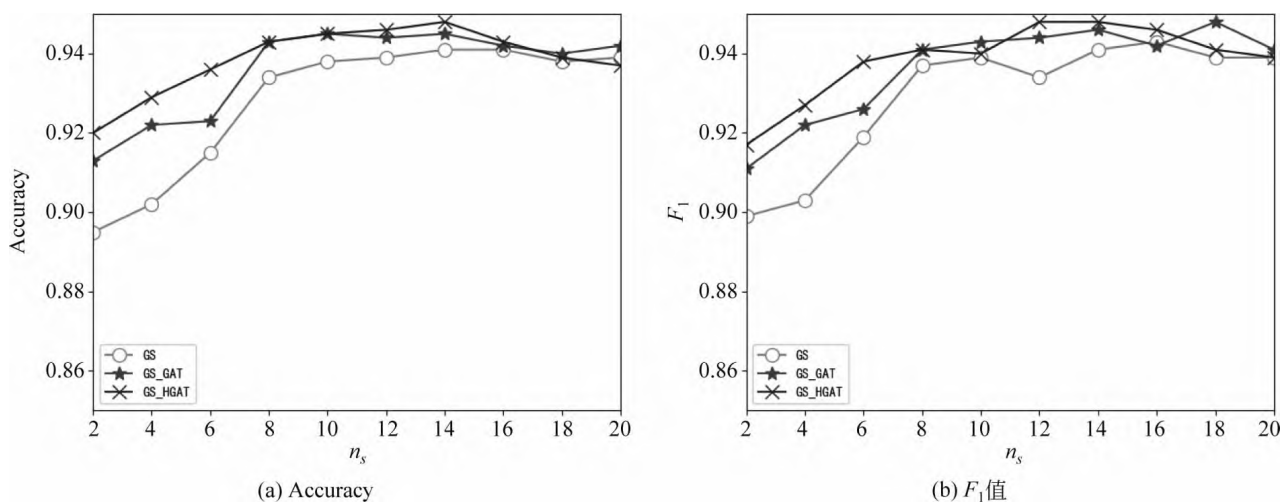


图 5 不同聚合方法在 Rumdict 数据集上的比较

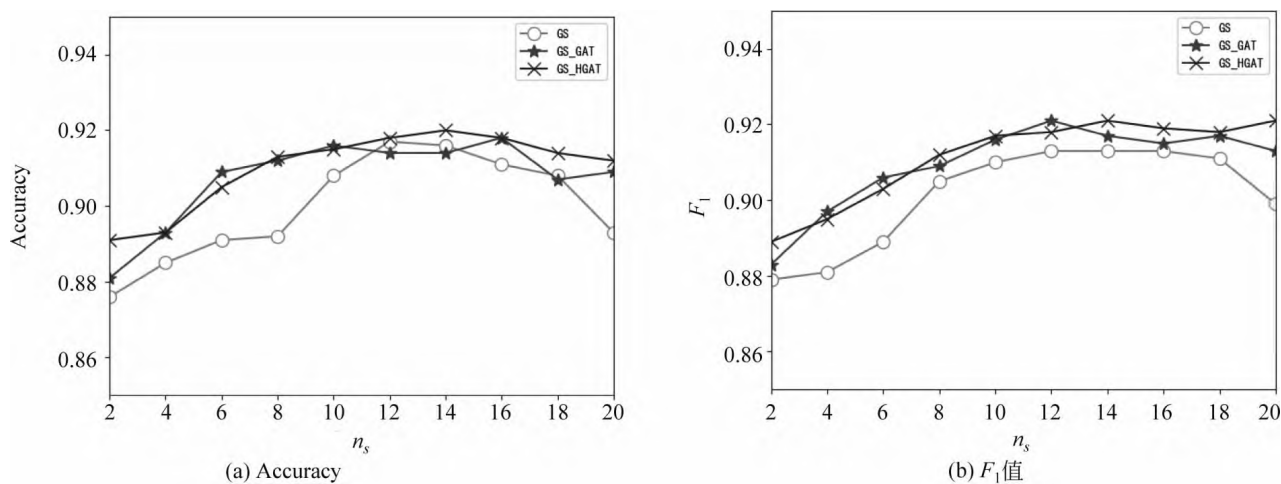


图 6 不同聚合方法在 CED 数据集上的比较

进一步地, 通过对比 GS 与 GS_GAT 发现后者对应的结果普遍高于前者对应的结果, 这是因为 GS 方法在节点聚合过程中仅根据连边权重来获得邻居节点的加权结果, 而 GS_GAT 方法则计算了邻居节点的注意力权重大小, 继而能区分不同邻居节点对聚合结果的贡献程度, 提高节点表

示的准确性。对比 GS_HGAT 和 GS_GAT 时发现, 前者对应的结果普遍偏高。例如, 当使用 Rumdict 数据集时, GS_HGAT 在 $n_s=6$ 时对应的 Accuracy 值比 GS_GAT 方法高出 0.013; 当使用 CED 数据集时, GS_HGAT 在 $n_s=20$ 时对应的 F_1 值比 GS_GAT 算法高出 0.008, 这说明在节点聚合

过程中同时考虑节点权重与节点类型权重,能有效提高模型对于节点表示的学习能力,提升谣言检测效果。

3.5 消融实验

这里在 RD_EWF 的基础上进行调整,衍生出以下几种变体方法并将其与本文进行比较:

(1) **RD_E_noC**: 区别于 RD_EWF,仅使用事件构建谣言检测异质图,并且在计算事件初始化嵌入表达时不考虑评论文本的影响。此外,利用基于平均聚合的 GraphSAGE 算法获得事件的最终嵌入表达,并利用公式(18)、(19)进行模型训练。

(2) **RD_E**: 区别于 RD_EWF,仅使用事件构建谣言检测异质图。此外,利用基于平均聚合的 GraphSAGE 算法获得事件的最终嵌入表达,并利用公式(18)、(19)进行模型训练。

(3) **RD_EW**: 区别于 RD_EWF,该方法仅使用事件及词语构建谣言检测异质图。

(4) **RD_EW_senF**: 区别于 RD_EWF,该方法使用事件、词语以及文本特征中的情感特征构建谣言检测异质图。

(5) **RD_EW_entF**: 区别于 RD_EWF,该方法使用事件、词语以及文本特征中的语法特征构建谣言检测异质图。

(6) **RD_EW_linF**: 区别于 RD_EWF,该方法使用事件、词语以及文本特征中的语言特征构建谣言检测异质图。

在此基础上,我们统计了上述方法在不同数据集上对应的 Accuracy 值和 F_1 值,结果如表 5 所示。由表 5 可知:

表 5 本文方法与不同变体方法的比较

方法	Rumdect		CED	
	Accuracy	F_1	Accuracy	F_1
RD_E_noC	0.883	0.876	0.863	0.867
RD_E	0.922	0.919	0.894	0.896
RD_EW	0.925	0.925	0.901	0.908
RD_EW_senF	0.931	0.927	0.914	0.915
RD_EW_entF	0.925	0.929	0.908	0.906
RD_EW_linF	0.938	0.936	0.918	0.919
RD_EWF	0.944	0.943	0.918	0.923

(1) 与未使用评论信息的 RD_E_noC 方法相比,RD_E 对应的结果明显偏高。例如,当使用 Rumdect 数据集时,RD_E 相对于 RD_E_noC 在 Accuracy 值与 F_1 值方面分别提升 0.039 和 0.043,这说明评论文本对于检测事件是否为谣言具有重要作用。

(2) 对比 RD_EW 和 RD_E 发现,前者在不同数据集上对应的 Accuracy 值与 F_1 值相对后者均偏高,验证了本文使用 BiGRU 融合文本词语的上下文序列化信息对于提升事件节点表达精度的有效性。

(3) 进一步发现,与未使用文本特征的 RD_EW 方法相比,结合部分文本特征的 RD_EW_senF、RD_EW_entF 及 RD_EW_linF 方法对应的 Accuracy 值与 F_1 值普遍偏高。例如,当使用 Rumdect 数据集时,RD_EW_senF 相对于 RD_EW_noF 在 Accuracy 值与 F_1 值方面分别提升 0.006 和 0.002,RD_EW_linF 相对于 RD_EW 在 Accuracy 值与 F_1 值方面分别提升 0.013 和 0.011,说明在事件、词语等信息基础上考虑情感特征、语法特征或者语言特征能从一定程度上提高谣言检测效果。并且,不难发现 RD_EW_linF 相对于 RD_EW 的性能提升程度较另外两种方法更为明显。究其原因,RD_EW_linF 方法使用事件、词语以及文本特征中的语言特征构建谣言检测异质图,其抽取的语言特征中除包含情感特征外,还包含心理特征、认知特征等重要信息,因此相对于 RD_EW_senF(RD_EW_entF),单纯使用情感特征(语法特征)而言更有助于提高节点信息的聚合效果。

当对比 RD_EWF 与其他算法表现时发现,RD_EWF 在不同数据集上的结果均明显高于其他算法,这说明在谣言检测过程中综合考虑词语的情感特征、语法特征、语言特征后的方法性能要优于单纯使用其中一种特征时的方法性能,进一步验证了本文引入的文本特征对于提升谣言检测效果的有效性。

3.6 与现有典型方法的比较

本文将 RD_EWF 与 13 个典型方法进行对比,结果如表 6 所示。其中,所有方法中的最优结果用粗体表示,次优结果使用下划线表示。由表 6 可知:

表 6 不同方法的实验结果 (单位: %)

方法		Rumdetect		CED	
		Accuracy	F_1	Accuracy	F_1
传统方法	NB	78.99	78.76	75.11	75.80
	LR	81.92	78.89	80.76	78.84
	RF	77.89	75.72	77.96	74.57
深度学习 方法	FastText	84.74	83.21	80.73	79.65
	TextCNN	90.38	90.13	89.01	88.60
	BiGRU-CNN	91.36	92.24	89.04	89.36
	TextGCN	91.10	91.75	89.46	90.75
	TextING	91.72	92.54	91.24	90.40
	I-BERT-LSTM	91.86	92.79	91.62	91.23
	dEFEND	<u>94.05</u>	94.17	91.70	91.24
	Bi-GCN	92.62	93.19	91.38	91.71
	RumorGCN	93.55	<u>94.36</u>	91.51	<u>91.97</u>
	GLAN	93.10	93.75	91.46	91.75
	RD_EWF	94.23	94.82	<u>91.67</u>	92.08

(1) 基于深度学习模型的文本分类方法在不同数据集上的表现均优于 NB、LR、RF 等传统分类方法,原因在于前者方法能更好地挖掘谣言信息的隐含特征,而传统分类方法只是使用简单的词袋模型表示事件文本,丢失了文本上下文语义、词语顺序等重要信息。进一步发现,I-BERT-LSTM 结果相对于 TextCNN、BiGRU-CNN、TextGCN、TextING 等方法普遍偏高,这是由于 TextCNN 等方法直接将整个微博事件对应的文本信息作为模型输入,而 I-BERT-LSTM 利用 TextRank 算法获得微博事件对应的文本摘要,因此能有效避免截取有限长度序列带来的关键信息丢失问题。

(2) 当将 dEFEND 与 TextCNN、TextGCN、I-BERT-LSTM 等方法比较时发现,前者对应结果明显偏高,说明将原始微博划分成句子并结合互注意力机制细化句子与相关评论的关系能较好地提升谣言检测效果。

(3) 相对于 Bi-GCN,RumorGCN 对应结果均有所提升,这是因为 Bi-GCN 仅仅考虑传播树中父

子节点之间形成的层间依赖关系,而 RumorGCN 共同显式建模层间依赖关系和兄弟节点之间形成的层内依赖关系,因此能聚合不同依赖关系下的局部邻域信息,继而学习到更准确、更丰富的传播结构特征。

(4) 通过比较 RD_EWF 和其他方法发现,前者除在 CED 数据集上的 Accuracy 值低于 dEFEND 方法对应结果外,在其他情况下均获得最优实验结果。究其原因:①本文构建的谣言检测异质图在传统评论、词语信息的基础上引入情感、语法、心理等知识,综合了事件、词语以及文本特征三方面之间的相互影响,丰富了节点向量表达中所含的特征信息;②本文提出的基于 GraphSAGE 和 HGAT 的节点聚合方法不仅能区分不同邻居节点的重要性,还考虑了不同节点类型对于聚合结果的贡献,避免了 TextING、Bi-GCN、RumorGCN 等方法单纯考虑邻居节点重要性而导致的节点表示不准确的问题。

3.7 微博评论数量影响分析

由于谣言传播较为迅速,因此能否及时地对尚未被评论或者较少被评论过的微博事件进行正确检测是衡量谣言检测算法性能的重要标准。为此,这里根据 3.6 节结果选取了 4 种表现较好的谣言检测算法 I-BERT-LSTM、RumorGCN、dEFEND 和 GLAN,并将它们与本文 RD_EWF 方法进行对比以分析评论数量对不同方法的影响。为了仿真那些尚未被评论或者较少被评论的微博信息,我们随机选择测试集中 10% 的微博事件,然后针对每个事件只保留其中 10% 的评论信息。在此基础上,我们在数据集 Rumdetect 与 CED 上统计了每种方法对应的 Accuracy 值和 F_1 值,结果如图 7、图 8 所示。

由图知,随着测试集中部分微博评论数减少,上述方法对应性能均呈现出不同程度的下降,说明微博评论中蕴含着较多与谣言检测相关的信息。进一步发现,dEFEND 对应结果普遍高于 I-BERT-LSTM、RumorGCN 和 GLAN,说明关注原始微博不同句子之间以及句子与评论之间的语义关联性能较好地保证谣言检测效果。与 I-BERT-LSTM 和 GLAN 相比,RumorGCN 对应结果稍高,可能原因

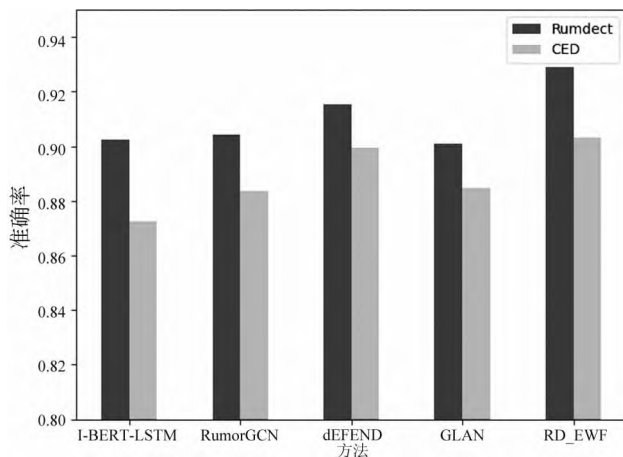


图7 微博评论数量对不同方法的影响

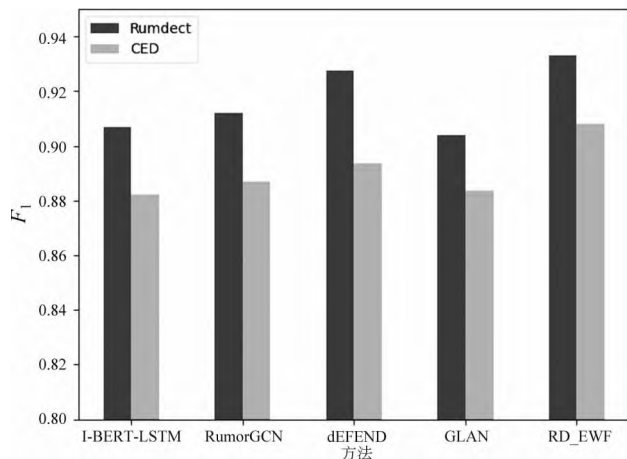


图8 微博评论数量对不同方法的影响

是前两种算法只关注评论信息之间的内容相关性,忽略了评论之间实际转发关系对谣言检测结果的影响。对比 RD_EWF 与其他算法发现,本文对应的 Accuracy 值和 F_1 值在不同数据集上均获得最大值,虽然相对于表 6 中的表现有所下降,但所得结果仍明显高于其他算法。可见,由于本文在考虑原始微博和评论信息的基础上进一步结合了不同事件之间的相关性,因此能在评论较少的情况下根据事件之间的相互联系学习到对分类有用的重要信息,继而提高针对尚未被评论或者有较少评论的微博信息的检测能力。

4 结束语

本文提出了一种基于事件-词语-特征异质图的微博谣言检测新方法 RD_EWF,主要贡献包括:①

在微博原文及评论内容信息基础上,引入了由情感特征、语法特征以及语言特征构成的文本特征的概念,将微博事件、文本词语、文本特征作为节点构建事件-词语-特征异质图,解决了现有方法单纯利用文本内容导致节点信息表达不充分的问题;②综合考虑事件-事件之间、事件-词语之间、词语-词语之间以及词语-特征之间的相互作用,提出基于 GraphSAGE 和异质图注意力网络的节点聚合方法 GS_HGAT,以此区分不同类型节点的影响,在保证模型可归纳的同时提高节点表示的准确性。在两个典型谣言检测数据集上的实验结果表明,RD_EWF 相对于传统文本分类方法及深度学习方法在提升微博谣言检测准确性方面具有明显优势。未来计划将该方法推广至标题党识别、虚假信息识别等相关领域。

参考文献

- [1] KUMAR S, SHAN N. False information on web and social media: A survey[M]. Social Media Analytics: Advances and Applications, Calabasas: CRC press, 2018.
- [2] 张志勇, 荆军昌, 李斐, 等. 人工智能视角下的在线社交网络虚假信息检测、传播与控制研究综述[J]. 计算机学报, 2021, 44(11): 2261-2282.
- [3] 曾雪强, 华鑫, 刘平生, 等. 基于情感轮和情感词典的文本情感分布标记增强方法[J]. 计算机学报, 2021, 44(6): 1080-1094.
- [4] NGUYEN T N, LI C, NIEDDREE C. On early-stage debunking rumors on twitter: Leveraging the wisdom of weak learners[C] // Proceedings of International Conference on Social Informatics, 2017: 141-158.
- [5] SINGH J P, KUMAR A, RANA N P, et al. Attention-based LSTM network for rumor veracity estimation of tweets[J]. Information Systems Frontiers, 2022, 24: 459-474.
- [6] MA J, GAO W, MITRA P, et al. Detecting rumors from microblogs with recurrent neural networks [C] // Proceedings of IJCAI, 2016: 3818-3824.
- [7] SONG C, YANG C, CHEN H, et al. CED: Credible early detection of social media rumors [J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 33(08): 3035-3047.
- [8] 王友卫, 童爽, 凤丽洲, 等. 基于图卷积网络的归纳式微博谣言检测新方法 [J]. 浙江大学学报(工学版),

- 2022, 56(5): 956-966.
- [9] CHEN X, ZHU D, LIN D, et al. Rumor knowledge embedding based data augmentation for imbalanced rumor detection [J]. *Information Sciences*, 2021, 580: 352-370.
- [10] 杨延杰, 王莉, 王宇航. 融合源信息和门控图神经网络的谣言检测研究[J]. *计算机研究与发展*, 2021, 58(7): 1412-1424.
- [11] WU Z, PI D, CHEN J, et al. Rumor detection based on propagation graph neural network with attention mechanism [J]. *Expert Systems with Applications*, 2020, 158: 113595.
- [12] BIAN T, XIAO X, XU T, et al. Rumor detection on social media with bi-directional graph convolutional networks [C] // *Proceedings of AAAI*, 2020: 549-556.
- [13] LU Y J, LI C T. GCAN: Graph-aware co-attention networks for explainable fake news detection on social media [C] // *Proceedings of ACL*, 2020: 505-514.
- [14] ZHANG H, QIAN S, FANG Q, et al. Multi-modal meta multi-task learning for social media rumor detection[J]. *IEEE Transactions on Multimedia*, 2021, 24: 1449-1459.
- [15] ZHOU J, CUI G, HU S, et al. Graph neural networks: A review of methods and applications[J]. *AI Open*, 2020, 1: 57-81.
- [16] 牛耘, 潘明慧, 魏欧, 等. 基于词典的中文微博情绪识别[J]. *计算机科学*, 2014, 9: 253-258.
- [17] 徐琳宏, 林鸿飞, 潘宇, 等. 情感词汇本体的构造[J]. *情报学报*, 2008, 27(2): 180-185.
- [18] JIAO Z, SUN S, SUN K. Chinese lexical analysis with deep bi-gru-crf network[J]. *arXiv preprint arXiv: 1807.01882*, 2018.
- [19] LUO X, LIU Z, SHANG M, et al. Highly-accurate community detection via pointwise mutual information-incorporated symmetric non-negative matrix factorization[J]. *IEEE Transactions on Network Science and Engineering*, 2020, 8(1): 463-476.
- [20] YUAN X, WANG S, WAN L, et al. SSF: Sentence similar function based on Word2Vector similar elements[J]. *Journal of Information Processing Systems*, 2019, 15(6): 1503-1516.
- [21] HAMILTON W, YING Z, LESKOVEC J. Inductive representation learning on large graphs [C] // *Proceedings of NIPS*, 2017: 1025-1035.
- [22] LINMEI H, YANG T, SHI C, et al. Heterogeneous graph attention networks for semi-supervised short text classification[C] // *Proceedings of EMNLP-IJCNLP*, 2019: 4821-4830.
- [23] ZHANG J, JIANG Y, WU S, et al. Prediction of remaining useful life based on bidirectional gated recurrent unit with temporal self-attention mechanism[J]. *Reliability Engineering & System Safety*, 2022, 221: 108297.1-108297.10.
- [24] MA J, GAO W, WONG K F. Detect rumors in microblog posts using propagation structure via kernel learning[C] // *Proceedings of ACL*, 2017: 708-717.
- [25] WONGKAR M, ANGDRSEY A. Sentiment analysis using naive Bayes algorithm of the data crawler; Twitter [C] // *Proceedings of ICIC*, 2019: 1-5.
- [26] SHAH K, PATEL H, SANGHVI D, et al. A comparative analysis of logistic regression, random forest and KNN models for the text classification[J]. *Augmented Human Research*, 2020, 5(1): 1-16.
- [27] JOULIN A, GRAVE E, BOJANOWSKI P, et al. Bag of tricks for efficient text classification [J]. *arXiv preprint arXiv: 1607.01759*, 2016.
- [28] KIM Y. Convolutional neural networks for sentence classification [C] // *Proceedings of EMNLP*, 2014: 1746-1751.
- [29] 程艳, 尧磊波, 张光河, 等. 基于注意力机制的多通道 CNN 和 BiGRU 的文本情感倾向性分析[J]. *计算机研究与发展*, 2020, 57(12): 2583-2595.
- [30] YAO L, MAO C, LUO Y. Graph convolutional networks for text classification [C] // *Proceedings of AAAI*, 2019: 7370-7377.
- [31] ZHANG Y, YU X, CUI Z, et al. Every document owns its structure: Inductive text classification via graph neural networks [C] // *Proceedings of ACL*, 2020: 334-339.
- [32] 许诺, 赵薇, 尚柯源, 等. 基于预训练语言模型的健康谣言检测[J]. *系统科学与数学*, 2022, 42(10): 2582-2589.
- [33] SHU K, CUI L, WANG S, et al. dEFEND: Explainable fake news detection [C] // *Proceedings of SIGKDD*, 2019: 395-405.
- [34] 胡斗, 卫玲蔚, 周薇, 等. 一种基于多关系传播树的谣言检测方法[J]. *计算机研究与发展*, 2021, 58(7): 1395-1411.
- [35] YUAN C, MA Q, ZHOU W, et al. Jointly embedding the local and global relations of heterogeneous graph for rumor detection [C] // *Proceedings of IC-DM*, 2019: 796-805.
- [36] VELIČ KOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks [C] // *Proceedings of ICLR*, 2018: 1-12.



王友卫(1987—),通信作者,博士,副教授,主要研究领域为内容安全、深度学习、数据挖掘。
E-mail: ywwang15@126.com



凤丽洲(1986—),博士,副教授,主要研究领域为文本分析、数据挖掘。
E-mail: lzfeng15@126.com



王炜琦(2000—),硕士研究生,主要研究领域为数据挖掘、网络内容安全分析。
E-mail: wangweiqi132@163.com

中国中文信息学会开展“第九届中国科协青年人才托举工程项目”申报工作的通知

按照《中国科协办公厅关于开展第九届中国科协青年人才托举工程项目被托举人遴选工作的通知》(科协办函创字〔2023〕71号),中国中文信息学会现开展第九届“中国科协青年人才托举工程”候选人的遴选工作。中国科协青年人才托举工程于2015年正式启动,旨在探索创新青年科技人才选拔培养机制,重点关注在关键共性技术、前沿引领技术、现代工程技术、颠覆性技术、“卡脖子”技术等方面潜心研究、守正创新的青年科技工作者,重点关注积极投身国家重大战略部署、基础研究、应用基础研究、工程技术研究等的青年科技工作者,重点关注为推动科技经济融合发展作出积极贡献的青年科技工作者,重点关注在国内开展学术研究取得突出成绩、在国内期刊发表高水平研究成果等的青年科技工作者,重点关注在临床案例、科研仪器、工程技术等领域案例库中发布高水平案例成果的青年科技工作者。项目对入选者每年资助10万元,资助期为3年。现将遴选有关事项通知如下:

申报条件

1. 在中华人民共和国境内工作学习的中国籍公民,自觉遵守中华人民共和国法律法规;
2. 申请人年龄在32岁以下(1991年1月1日后出生(含));
3. 中国中文信息学会个人会员,中文信息处理领域的基层一线科技工作者,积极参与学会活动;
4. 具有良好的科研诚信和学风道德,自觉践行新时代科学家精神;具有坚实的理论基础、较强的创新能力、良好的科研潜质;
5. 曾被取消被托举人入选资格的,青年人才托举工程往届入选者,人力资源和社会保障部博士后创新人才支持计划、留学回国人员创业启动支持计划入选者和其他国家级人才计划入选者,不作为遴选对象;准备出国一年以上的人员,不作为遴选对象;
6. 申请人所在单位或工作团队能够提供相应的配套资金

和政策支持,共同促进其成长。

资助名额

资助名额不超过4人,其中科协资助2人,自筹资助2人。

推荐方式

1. 采取同行推荐和专家评审的人才选拔方式;
2. 申请人应由不少于3名同行专家联名推荐,其中至少1名推荐专家作为提名人并承担托举及导师责任,至少2名推荐专家与申请人具有相同研究领域;
 - 1) 推荐专家为两院院士、中国中文信息学会理事,且最多1名推荐专家与申请人属于同一单位;
 - 2) 每名推荐专家当年度最多可推荐1人,提名人最多可提名1名候选人。

评审流程及申报要求

1. 提交申请材料 请于2023年9月15日前提交《申请表》及相关证明材料。电子版发送至学会邮箱,标题为“CIPS青托+姓名+单位”;纸质版材料一式三份(含原件一份)寄送至学会办公室。(逾期不予受理)
2. 专家评审 学会将择期召开专家评审会,对候选人材料进行评审,共遴选出最多4名入选者。
3. 评审公示 评审结果及评审专家主要信息将在学会网站进行公示,接受会员和社会公众监督。
4. 科协审核 公示无异议报中国科协审核发布后,最终入选第九届“中国科协青年人才托举工程”项目。

联系方式

学会邮箱: cips@iscas.ac.cn

通信地址:北京市海淀区中关村南四街4号中科院软件园7号楼201室,学会办公室

办公电话: 010-62661047

(中国中文信息学会)