



计算机科学与探索
Journal of Frontiers of Computer Science and Technology
ISSN 1673-9418, CN 11-5602/TP

《计算机科学与探索》网络首发论文

题目：基于文本语义增强和评论立场加权的网络谣言检测
作者：朱奕，王根生，金文文，黄学坚，李胜
网络首发日期：2024-04-11
引用格式：朱奕，王根生，金文文，黄学坚，李胜. 基于文本语义增强和评论立场加权的网络谣言检测[J/OL]. 计算机科学与探索.
<https://link.cnki.net/urlid/11.5602.TP.20240410.2115.002>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于文本语义增强和评论立场加权的网络谣言检测

朱奕¹, 王根生^{2,3+}, 金文文³, 黄学坚², 李胜¹

1. 江西财经大学 财税与公共管理学院, 南昌 330013

2. 江西财经大学 信息管理学院, 南昌 330013

3. 江西财经大学 人文学院, 南昌 330013

+ 通信作者 E-mail: wgs74@126.com

摘要：社交网络方便人们信息交流的同时也为谣言的传播提供了新的温床。由于社交媒体帖子通常十分精简，因此大多数基于内容语义特征的谣言检测方法面临着语义信息不足的挑战。同时，目前基于传播特征的谣言检测方法常常忽略了评论用户的个体特征，未能合理分配不同用户评论的权重。因此，提出一种结合文本语义增强和评论立场加权的网络谣言检测方法。首先，通过外部知识图谱获取帖子中的实体和概念的解释，以提供更多上下文信息，从而增强语义理解。接着，借助点互信息将增强后的文本转化为加权图表示，并利用加权图注意力网络学习帖子的增强语义特征。然后，通过预训练的立场检测模型提取帖子中每条评论的立场信息，并根据评论用户的特征来学习立场信息的权重值。此外，将评论立场的时序数据和相应的评论用户序列数据输入跨模态的 Transformer，以学习评论立场的时序特征。最终，将增强的语义特征与加权的评论立场时序特征进行自适应融合，并输入多层感知机中进行分类。在 PHEME 和 Weibo 两个数据集上的实验结果表明，该方法不仅准确率高于最先进的基线方法 1.6% 以上，而且在早期谣言检测方面，比最好的基线方法提前 12 个小时。

关键词：谣言检测；语义增强；评论立场；图神经网络；知识图谱

文献标志码：A **中图分类号：**TP391

Network Rumor Detection Based on Enhanced Textual Semantics and Weighted Comment Stance

ZHU Yi¹, WANG Gensheng^{2,3+}, JIN Wenwen³, HUANG Xuejian², LI Sheng¹

1. School of Finance, Taxation and Public Administration, Jiangxi University of Finance and Economics, Nanchang 330013, China

2. School of Information Management, Jiangxi University of Finance and Economics, Nanchang 330013, China

3. School of Humanities, Jiangxi University of Finance and Economics, Nanchang 330013, China

Abstract: Social networks, while enabling information exchange among individuals, also serve as fertile grounds for the dissemination of rumors. The succinct nature of social media posts poses a challenge for most rumor detection methods reliant on content semantic features due to the insufficiency of semantic information. Additionally, numerous

基金项目：国家自然科学基金（72061015）；江西省教育厅科技项目（GJJ200539）。

This work was supported by the National Natural Science Foundation of China (72061015), the Science and Technology Projects of Jiangxi Provincial Department of Education (GJJ200539).

rumor detection techniques focusing on propagation features often disregard the unique attributes of commenters, leading to inadequate allocation of weights to different user comments. Thus, a network rumor detection approach is proposed, integrating text semantic enhancement and weighted comment stance. Initially, entities and concepts in posts are elucidated via an external knowledge graph to furnish additional contextual information, thereby augmenting semantic comprehension. Subsequently, leveraging pointwise mutual information, the enhanced text is translated into a weighted graph representation, and a weighted graph attention network is employed to assimilate enhanced semantic features of posts. Stance information for each comment within the post is then extracted using a pre-trained stance detection model, with weight values of stance information being learned based on commenters' characteristics. Furthermore, temporal data of comment stances and corresponding commenter sequences are fed into a cross-modal Transformer to glean the temporal features of comment stances. Ultimately, the enhanced semantic features are adaptively merged with the weighted temporal features of comment stances and fed into a multi-layer perceptron for classification. Experimental results on the PHEME and Weibo datasets demonstrate that this method not only achieves an accuracy improvement of over 1.6% compared to the state-of-the-art baseline method but also outperforms best baseline method by at least 12 hours in early rumor detection.

Key words: rumor detection; semantic enhancement; comment stance; graph neural network; knowledge graph

随着移动互联网的迅速发展, 社交媒体已经成为信息交流的重要渠道。然而, 社交媒体方便人们信息获取和分享的同时, 也为谣言的传播提供了滋生土壤^[1]。谣言的传播不仅妨碍了人们对社交媒体的有效利用, 而且可能造成民众的误解、扰乱社会秩序、甚至影响社会稳定和国家安全^[2]。例如, 在新冠肺炎疫情期间社交媒体上出现了一些“COVID-19 疫苗会改变人的 DNA”、“5G 网络导致新冠疫情传播”和“COVID-19 病毒是人造的”等谣言, 导致一些民众产生了误解和恐慌。为了遏制谣言的传播, 一些机构建立了谣言核实平台, 例如国外的事实核查网站 Snopes 和国内的中国互联网联合辟谣平台。然而, 这些平台在识别谣言时主要利用人工验证的方式, 不仅需要耗费大量的人力和物力, 而且存在严重的时间滞后问题。因此, 研究社交网络谣言自动识别方法对遏制谣言的传播具有重要意义。

早期的网络谣言识别主要采用了基于特征工程的机器学习方法^[3], 其研究焦点集中在选择与谣言相关的特征上。然而, 这些手工特征设计受限于先验知识, 因此在捕捉谣言的深层语义和传播特征方面存在一定的局限, 缺乏全面性和灵活性^[4]。随着深度学习技术的不断进展, 研究者逐渐采用深度学习方法来自动挖掘谣言的语义特征和传播特征, 这使得谣言检测逐步进入了数据驱动的新时代^[5]。

然而, 由于内容长度的限制和大量缩写的使用, 社交媒体上的帖子内容通常十分精简, 导致基于内容语义特征的深度学习方法面临语义信息不足的问题^[6]。同时, 目前大部分基于传播特征的深度学习谣言检测方法忽略了评论用户的个体特征, 未对不同的用户评论分配适当的权重。因此, 本文提出一种结合文本语义增强和评论立场加权的网络谣言检测方法, 简称为 TSCS。首先, 通过外部知识图谱获取帖子文本中的实体和概念的解释, 为帖子提供更多的上下文信息, 以增强语义理解。然后, 通过预训练的立场检测模型提取帖子的每条评论的立场信息, 并根据评论用户特征为不同的立场信息分配适当的权重。最后, 融合增强的语义特征和加权的评论立场信息进行谣言检测。在两个公开的社交媒体谣言数据集 PHEME 和 Weibo 上的实验结果表明, 本文方法优于所有对比的基线方法。本文的主要贡献如下:

(1) 提出了一种融合内容语义特征和评论立场的谣言检测模型, 通过内容特征和传播特征的互补, 提高模型的准确率和时效性。

(2) 提出了一种通过知识图谱丰富帖子上下文的方法, 并利用一种加权图注意力网络学习帖子的增强语义特征。

(3) 提出了一种考虑用户特征的评论立场学

习方法,为不同置信度的用户分配不同的权重,使模型专注于权威用户的立场,避免网络水军的故意炒作。

(4)在两个公开的数据集上进行了综合的实验对比和分析,实验结果显示本文方法优于基线方法。

1 相关研究

早期谣言检测研究主要集中在文本内容方面,旨在发掘谣言和非谣言在文本内容上的区别。然而,谣言的散布者常常故意模仿真实信息的写作风格,给单纯基于内容特征的方法带来了挑战。随后,一些研究者转向利用消息传播特征,以捕获谣言和非谣言在传播过程中的差异,并取得了较好的效果^[7]。目前,谣言检测大致可以分为基于内容特征和基于传播特征的两类方法^[8]。

1.1 基于内容特征的谣言检测

基于内容特征的方法将谣言检测视为文本分类任务,当前的主要趋势是利用深度学习技术自动学习谣言内容的语义特征。例如,Kaliyar 等人^[9]提出了一种基于卷积神经网络(Convolutional Neural Network, CNN)的谣言检测模型 FNDNet,从文本内容中自动提取有助于谣言分类的特征;Ajao 等人^[10]结合长短时记忆网络(Long Short-Term Memory, LSTM)和 CNN,以提取 Twitter 上谣言的语义特征;Alkhodair 等人^[11]采用 LSTM 和 Word2vec 的组合来检测与新兴主题相关的突发新闻谣言;Cheng 等人^[12]提出了一种基于生成对抗网络(Generative Adversarial Network, GANs)的谣言检测模型,通过生成器和鉴别器的相互促进,以增强谣言语义特征的学习。然而,由于社交媒体上帖子通常非常精简,上下文语义信息相对较少,并且谣言的发布者经常模仿真实信息的表达方式,导致仅基于内容语义特征的方法有时无法获得良好的检测效果。为了缓解这个问题,一些研究人员在内容语义特征的基础上引入了其他辅助特征。例如,Xu 等人^[13]引入了一种基于主题驱动的谣言检测模型,利用 CNN 从内容中提取主题信息,并将其与内容语义信息相结合;王根生等人^[14]提出了一种融合知识图谱的谣言实

时检测方法,通过知识蒸馏从外部知识图谱中获取文本内容的背景知识,以增强对帖子的语义理解;Singh 等人^[15]将基于注意力的 LSTM 提取的文本语义特征与用户特征整合用于谣言检测;黄学坚等人^[16]提出了一种融合多元用户特征和内容特征的谣言实时检测模型,利用用户的历史行为数据挖掘用户理性值和专业性两个深层次特征。针对基于文本内容的方法忽视了帖子中的图片或视频信息的问题,刘金硕等人^[17]提出了一种融合文本和图像的多模态谣言检测模型 MSRD;Chen 等人^[18]通过多头自注意力融合机制结合文本和图像的互补信息,并利用相反的潜在主题记忆网络来存储有关谣言真假模式的语义信息。针对多模态分析通常无法保持每种模态对应的独特特征的问题^[19],Yadav 等人^[20]提出了一种基于 Transformer 的多级注意力框架 ETMA,用于多模态虚假新闻检测。刘华玲等人^[21]对多模态学习的虚假新闻相关技术做了全面综述后发现,目前的研究在数据集、模型、语境等方面还存在不少缺陷。

1.2 基于传播特征的谣言检测

基于传播特征的方法主要利用信息的传播结构、评论和转发等信息进行谣言检测^[22-23]。例如,Ma 等人^[24]提出了一种传播树核方法,用于捕捉谣言传播结构的高阶模式;Ma 等人^[25]使用时间序列评论数据构建了基于循环神经网络(Recurrent Neural Network, RNN)的谣言检测模型;文献^[26]和^[27]构建了基于树形递归神经网络(Tree-structured Recursive Neural Networks, RvNN)的谣言检测方法,同时学习评论数据的语义特征和信息传播的结构特征。为了使模型专注于传播数据中具有典型谣言特征的部分,Chen 等人^[28]提出了一种融合注意力机制的谣言检测模型。然而,目前大多数基于传播特征的方法主要关注传播过程中的显式交互关系,忽略了潜在关系的建模问题。为了解决这一问题,胡斗等人^[29]提出了一种基于多关系传播树的谣言检测模型。此外,针对当前基于传播结构的谣言检测方法忽视了传播时序的问题,Huang 等人^[30]提出了一种融合空间结构和传播时序的时空模型。随着图神经网络(Graph Neural Network, GNN)的发展,一些研究人员开始利用各种 GNN 学习传播结构的

表示^[31-32]。例如, Wu 等人^[33]提出了基于传播图神经网络全局嵌入和集成学习的谣言检测模型,通过注意力机制来动态调整传播图中每个节点的权重;Liu 等人^[34]提出了一种结合双注意力机制和图卷积网络(Graph Convolutional Network, GCN)的社交网络谣言检测方法,利用 GCN 提取每条与事件相关的微博的传播结构信息,并通过注意力机制来抑制不相关的交互关系。研究发现,帖子传播过程中的评论信息反映了用户对该帖子的立场,例如赞成、反对、怀疑等,这些立场信息对谣言的判断起到了重要作用^[35],因此文献[36-39]提出了多种基于评论立场的谣言检测方法。为了降低模型对传播特征的依赖,以提高早期谣言检测的性能,Tu 等人^[40]构建了一种融合传播特征和文本内容特征的谣言检测方法,实验结果表明该方法相较于仅基于传播特征的方法,能够至少提前 12 个小时进行谣言检测;Lotfi 等人^[41]提出了一种联合传播特征和用户特征的谣言检测模型,利用图卷积神经网络学习信息传播图和用户交互图;Sun 等人^[42]设计了一种基于双动态图卷积网络(Dual Dynamic Graph Convolutional Networks, DDGCN)的谣言检测方法,将消息传播的动态性和来自外部知识图谱中的背景知识融合到统一的框架中;Xu 等人^[43]提出了一种层次聚合的图神经网络(Hierarchically Aggregated Graph Neural Networks, HAGNN),用于捕获文本内容不同粒度的高级表示,并融合传播结构特征进行谣言检测。

1.3 研究评述

基于传播特征的方法通常需要谣言得到一定范围的传播后才能获得较好的检测效果,不能及时地识别谣言。相比之下,基于内容特征的方法独立于谣言的传播数据,将谣言检测视为文本分类问题,能够实现谣言的即时检测。然而,某些谣言可能不具备典型的文本分类特征,这使得仅依赖内容特征的方法效果较差。最新的谣言检测方法采用基于混合特征的策略,将传播特征、内容特征和用户特征融合在一起,降低模型对传播特征的依赖,提高检测的时效性,这已经成为谣言检测领域的发展趋势^[44]。因此,本文提出一种融合内容语义特征和评论

立场的谣言检测模型,通过内容特征和传播特征的互补,提高模型的准确率和时效性。本文模型不仅针对社交媒体帖子文本内容语义不足的问题,提出了一种通过知识图谱丰富帖子上下文的方法,而且设计了一种考虑用户特征的评论立场学习方法,为不同置信度的用户给与不同的权重,使模型专注于权威用户的立场。

2 模型构建

2.1 问题定义

给定一组帖子集合 $P = \{p_1, p_2, \dots, p_n\}$, $p_i = (T_i, C_i)$, 其中 $T_i = \{w_1, w_2, \dots, w_n\}$ 表示帖子 p_i 的文本信息, $C_i = \{(u_1, t_1, c_1), (u_2, t_2, c_2), \dots, (u_m, t_m, c_m)\}$ 表示帖子 p_i 的传播过程中用户的评论信息,其中 u_m 、 t_m 和 c_m 分别表示第 m 条评论的用户、时间和评论内容信息。目标是学习一个模型 $f: P \rightarrow Y$, 输入帖子 p_i 的文本信息 T_i 和用户评论信息 C_i , 把帖子 p_i 分类到预定义的类别 $Y = \{N, R\}$, N 表示非谣言, R 表示谣言。

2.2 总体架构

本文提出的一种基于文本语义增强和评论立场加权的网络谣言检测模型(TSCS)的总体架构如图1所示。主要包括三个核心模块:文本语义增强学习、评论立场加权学习和结果分类。在文本语义增强学习模块中,首先,通过外部知识图谱获取帖子中实体的概念解释,以丰富帖子的上下文信息。然后,采用点互信息把帖子文本和实体概念转化成图结构表示。最后,利用加权图注意力网络对图结构进行学习,以获得增强的文本语义特征。在评论立场加权学习模块中,首先,通过预训练的立场检测模型 StanceBERTa 提取帖子中每条评论的立场信息,并基于评论用户的特征学习立场信息的权重值。随后,将评论立场时序数据和对应的评论用户序列数据输入到 Cross-modal Transformer 中,以学习综合考虑了用户特征的评论立场时序特征。在结果分类模块中,将增强的语义特征与加权的评论立场时序特征进行自适应融合,然后输入多层感知机进行谣言分类。

图 1 TSCS 的架构

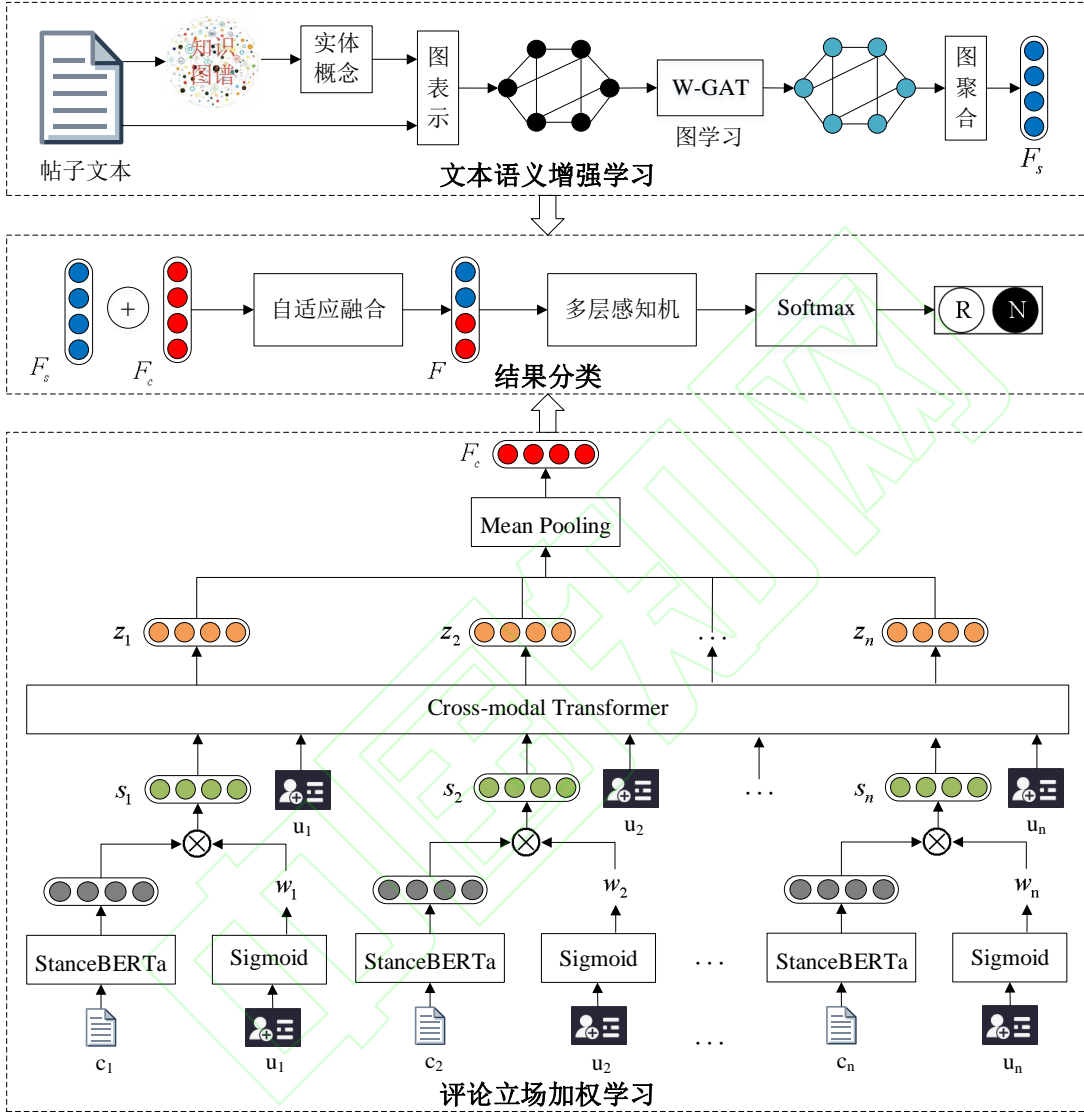


Fig.1 The architecture of TSCS

2.2.1 文本语义增强学习

(1) 实体概念获取

社交媒体上的帖子通常非常精简，上下文信息相对较少，无法提供详细的语义信息。通过外部知识图谱获取帖子中实体的概念，可以增强对帖子内容的语义理解。实体概念获取的具体流程如图 2 所示。首先，通过实体链接（Entity Linking）技术，把帖子中模棱两可的实体提及与知识库中相应的实体进行关联。这一任务包括两个主要步骤，即实

体识别和实体消歧。随后，借助概念化（Conceptualization）技术，从知识图谱中获取每个已识别实体的相关概念信息。例如，对于一段帖子文本“电联表示 5G 网络绝不会导致 COVID-19 的传播”，通过实体链接获得识别的实体集合 $T=\{\text{电联、5G 网络、COVID-19}\}$ 。然后，利用外部知识图谱对集合 T 中的实体进行概念化，获得对应的实体概念 $C_{\text{电联}}=\{\text{国际电信联盟、联合国机构}\}$ 、 $C_{\text{5G 网络}}=\{\text{通讯网络、第五代}\}$ 和 $C_{\text{COVID-19}}=\{\text{新冠疫情、肺炎}\}$ 。

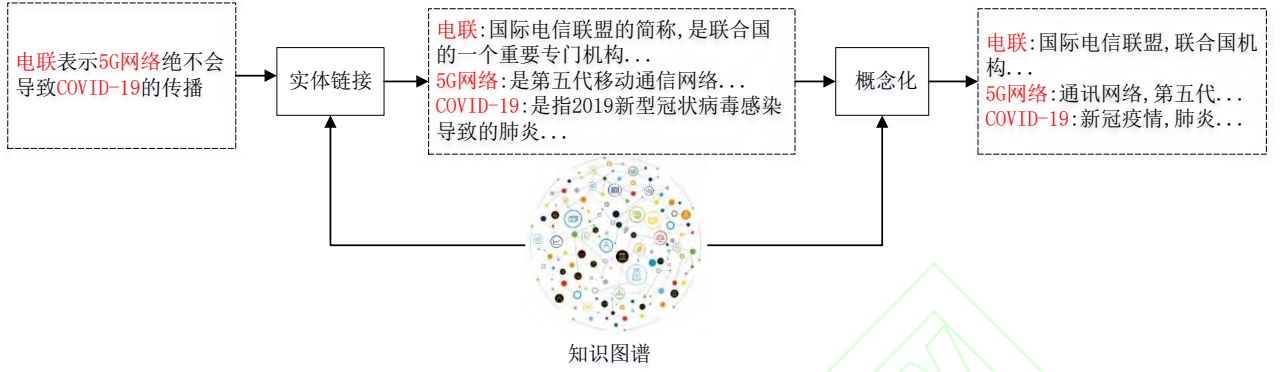


图2 实体概念获取流程

Fig.2 Entity concept acquisition process

(2) 文本图结构表示

目前, 大多数方法依赖于循环神经网络学习文本的语义特征。然而, 社交媒体上的帖子通常采用“#标签#文本@用户[缩写]文本[表情][符号]...”的形式。这些广泛使用的缩写、符号和表情导致了帖子内容的碎片化和离散化。在面对这种碎片化和离散化的文本内容时, 使用序列模型来捕获语义特征可能并不是最理想的选择。因此, 本文提出一种基于图注意力网络的文本语义特征学习方法。首先, 将帖子文本转换成图结构表示形式 $G=(V, E)$, 其中单词充当为节点 V , 而单词之间的相关性则构成了边 E , 边的权重反映了单词之间的关联程度。为了计算单词之间的关联程度, 本文采用点互信息 (Point Mutual Information, PMI) 的方法。具体而言, 使用一个固定大小的窗口在语料库上收集单词的共现统计信息, 并根据公式 (1) ~ (3) 来计算 PMI 值。

$$p(w_i) = \frac{|W(w_i)|}{|W|} \quad (1)$$

$$p(w_i, w_j) = \frac{|W(w_i, w_j)|}{|W|} \quad (2)$$

$$PMI(w_i, w_j) = \text{lb} \frac{p(w_i, w_j)}{p(w_i)p(w_j)} \quad (3)$$

其中, $|W(w_i)|$ 表示包含了单词 w_i 的滑动窗口的数量, $|W|$ 表示滑动窗口的总数量, $|W(w_i, w_j)|$ 表示同时包含单词 w_i 和 w_j 的滑动窗口的数量。这些统计数据不是基于一个特定的帖子内容而是全局的语料库。PMI 值的大小反映了单词 w_i 和 w_j 之间的关联程度, PMI 值越大说明其关联程度越高。只保留 PMI 值

大于零的边, 如公式 (4) 所示。

$$A_{i,j} = \begin{cases} PMI(w_i, w_j) & PMI(w_i, w_j) > 0 \\ 0 & PMI(w_i, w_j) \leq 0 \end{cases} \quad (4)$$

把实体概念作为共现词, 添加到帖子的文本内容中。最终, 通过点互信息把扩充后的帖子文本内容转化成图结构表示 G 和邻接权重矩阵 A 。

(3) 语义特征学习

获得了文本的图结构表示后, 采用一种加权图注意力网络 (Weighted Graph Attention Network, W-GAT) [45] 进行语义特征学习。与传统的图注意力网络 (Graph Attention Network, GAT) 相比, W-GAT 在计算注意力系数时不仅考虑了节点之间的相似性, 还将边的权重作为输入, 从而能够更准确地捕捉全局信息。W-GAT 的计算如公式 (5) 所示。

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in N(i)} a_{i,j}^l W^l h_j^l \right) \quad (5)$$

其中, $N(i)$ 为节点 v_i 的邻居节点集合, σ 为激活函数, h_j^l 为邻居节点 v_j 在第 l 层的输出, W^l 为 W-GAT 第 l 层的学习参数, $a_{i,j}^l$ 为在 l 层中邻居节点 v_j 对 v_i 的权重, 其计算如公式 (6) 所示。利用 Word2vec 作为单词节点 v_i 的初始特征表示 h_i^0 。

$$a_{i,j}^l = \frac{\exp(e_{i,j}^l)}{\sum_{j \in N_i} \exp(e_{i,j}^l)} \quad (6)$$

$$e_{i,j}^l = \alpha(h_i^l, h_j^l, A_{i,j}) \quad (7)$$

其中, $e_{i,j}^l$ 为在第 l 层中节点 v_i 和 v_j 之间的注意力系数, $e_{i,j}^l$ 的计算如公式 (7) 所示。其中, α 为一个可学习的函数, h_i^l 和 h_j^l 分别为节点 v_i 和 v_j 在第 l 层的特征表示, $A_{i,j}$ 表示节点 v_i 和 v_j 的边权重, 即单词

w_i 和 w_j 间的点互信息。通过 l 层的加权图注意力网络学习后, 采用全局平均池化 (Mean Pooling) 来聚合图中所有节点的特征向量 $h_i^{(l+1)}$, 最终得到文本的语义特征 $F_s \in \mathbb{R}^{d^{(l+1)}}$, 其计算过程如公式 (8) 所示, 其中 V 为节点集合。

$$F_s = \frac{1}{|V|} \sum_{i \in V} h_i^{(l+1)} \quad (8)$$

2.2.2 评论立场加权学习

帖子传播过程中的评论信息反映了用户对该帖子的立场, 例如赞成、反对、怀疑等。这些立场信息对谣言的判断起到了重要作用。然而, 目前大多数方法忽略了评论用户的个体特征, 未对不同的用户评论立场分配适当的权重。因此, 本文提出一种考虑用户特征的评论立场学习方法, 为不同置信度的用户给与不同的权重, 以便模型更专注于权威用户的立场。

(1) 评论立场获取

首先, 利用预训练的 StanceBERTa 模型获取每条评论 c_i 的立场信息 SB_i , 如公式 (9) 所示。StanceBERTa 是一种专门用于立场检测任务的语言模型, 用于确定作者对特定目标或主题的态度或观点。它基于 BERT 架构, 这是一种广泛用于自然语言处理任务的预训练语言模型。StanceBERTa 在检测与有争议的话题相关的立场方面表现出强大的性能, 尤其擅长检测细微差别和语言差异。

$$SB_i = \text{StanceBERTa}(c_i) \quad (9)$$

(2) 评论立场加权

其次, 考虑到不同特征用户的立场应该分配不同的权重, 例如有很多粉丝的认证用户的权重应该高于普通的非认证用户。因此选择认证状态、粉丝量、关注量和注册时长等能反映用户权威性的社交特征作为权重影响因素。本文并没有手动设计权重计算公式, 而是通过神经网络自动学习。利用一个全连接层来学习用户的权重值 w_i , 从而得到加权的用户评论立场特征 s_i , 具体计算如公式 (10) 和 (11) 所示。

$$w_i = \text{Sigmoid}(W_u * u_i + b_u) \quad (10)$$

$$s_i = w_i * SB_i \quad (11)$$

其中, W_u 和 b_u 分别为全连接层参数和偏置项, u_i 为

第 i 条评论用户特征, SB_i 为第 i 条评论的立场特征。

(3) 评论立场时序特征学习

然后, 按照评论的时间先后顺序, 把帖子 p_i 的评论立场时序数据 $S_{p_i} = \{s_1, s_2, \dots, s_n\}$ 和对应的评论用户序列数据 $U_{p_i} = \{u_1, u_2, \dots, u_n\}$ 输入 Cross-modal Transformer^[46]中, 学习融合用户特征的评论立场时序特征 $Z \in \mathbb{R}^{n \times d_k}$, 如公式 (12) 所示。

$$Z = \text{Crossmodal_Transformer}(S_{p_i}, U_{p_i}) \quad (12)$$

Cross-modal Transformer 是 Transformer 的扩展, 可以接受来自不同模态的输入数据, 其架构如图 3 所示。

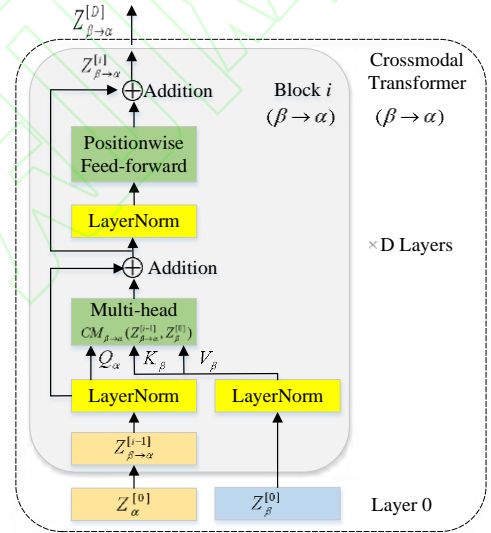


图 3 Cross-modal Transformer 的架构

Fig.3 Overall architecture for Cross-modal Transformer

跨模态注意力是 Cross-modal Transformer 的关键组成部分, 它允许模型在不同模态之间进行信息交互和融合, 其计算形式如公式 (13) 所示, 计算流程如图 4 所示。

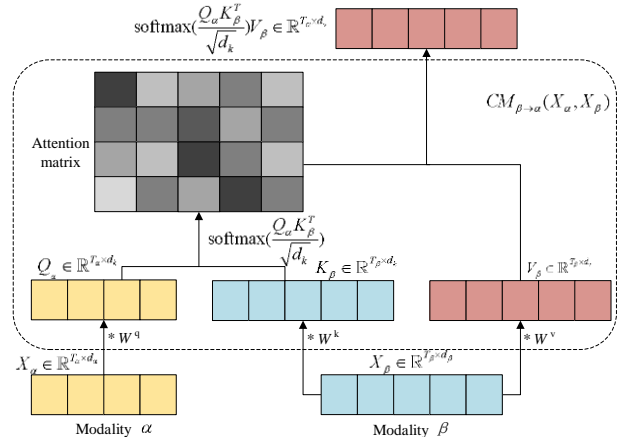


图 4 跨模态注意力

Fig.4 Cross-modal attention

$$CM_{\beta \rightarrow \alpha}(X_\alpha, X_\beta) = \text{softmax}\left(\frac{Q_\alpha K_\beta^T}{\sqrt{d_k}}\right) V_\beta \quad (13)$$

其中, $X_\alpha \in \mathbb{R}^{T_\alpha \times d_\alpha}$ 和 $X_\beta \in \mathbb{R}^{T_\beta \times d_\beta}$ 分别表示模态 α 和模态 β 的输入序列, T_α 和 T_β 为序列长度, d_α 和 d_β 为特征维度。 $Q_\alpha = X_\alpha W^q$ ($Q_\alpha \in \mathbb{R}^{T_\alpha \times d_k}$) 表示 Querys, $K_\beta = X_\beta W^k$ ($K_\beta \in \mathbb{R}^{T_\beta \times d_k}$) 表示 Keys, $V_\beta = X_\beta W^v$ ($V_\beta \in \mathbb{R}^{T_\beta \times d_v}$) 表示 Values, $W^q \in \mathbb{R}^{d_\alpha \times d_k}$, $W^k \in \mathbb{R}^{d_\beta \times d_k}$, $W^v \in \mathbb{R}^{d_\beta \times d_v}$ 为需要学习的参数。最后, 采用全局平均池化, 得到最终的评论立场时序特征 $F_c \in \mathbb{R}^{d_v}$, 如公式 (14) 所示。

$$F_c = \text{Mean_Pooling}(Z) \quad (14)$$

2.2.3 结果分类

通过自适应门将增强的文本语义特征 F_s 与加权的评论立场时序特征 F_c 进行融合, 形成最终的融合特征 F , 如公式 (15) ~ (18) 所示。

$$F_s = \sigma(W_s F_s + b_s) \quad (15)$$

$$F_c = \sigma(W_c F_c + b_c) \quad (16)$$

$$e = \text{sigmoid}(W_e (F_s \oplus F_c) + b_e) \quad (17)$$

$$F = e * F_s + (1 - e) F_c \quad (18)$$

然后, 将融合特征 F 输入多层感知机 (Multi-layer Perceptron, MLP) 中, 最后连接 softmax 输出分类结果, 如公式 (19) 所示。

$$y = \text{softmax}(W * \text{MLP}(F) + b) \quad (19)$$

其中, W 和 b 分别是线性层参数和偏置项。基于最小化交叉熵损失函数对模型进行训练, 其计算如公式 (20) 所示。

$$L = -\frac{1}{N} \sum_{i=1}^N (y_i \log y_i + (1 - y_i) \log (1 - y_i)) + \frac{\lambda}{2} \|W\|_2^2 \quad (20)$$

其中, y_i 和 y_i 分别表示样本 i 的真实标签值和预测值, $\frac{\lambda}{2} \|W\|_2^2$ 为 L2 正则化项, 用于降低模型的过拟合程度。

3 实验分析

3.1 实验设置

3.1.1 实验数据和评价指标

实验采用两个真实社交媒体谣言数据集 PHEME 和 Weibo, 来验证 TSCS 方法的有效性。

PHEME 数据集包含了发布在 Twitter 上的关于五个突发新闻事件的谣言和非谣言, 分别是悉尼劫持事件、渥太华枪击事件、德国之翼坠机事件、查理周刊袭击事件以及弗格森骚乱事件。Weibo 数据集则来源于新浪微博不实信息举报平台, 包含了政治、金融、健康等多领域的谣言和非谣言。这两个数据集提供了帖子的文本内容、用户评论以及用户信息。两个数据集的统计信息如表 1 所示。

表 1 数据集统计

Table 1 Dataset statistics

Datasets	Rumors	Non-rumours	Total
PHEME	1972	3830	5802
WEIBO	2313	2351	4664

采用 3:1:1 的比例进行数据集分层随机抽样, 划分出训练集、验证集和测试集。利用 K 折交叉验证的平均准确率 (Accuracy)、谣言查全率 (Recall)、谣言查准率 (Precision) 和 F1-score 作为性能的评价指标。

3.1.2 实现细节

在文本语义增强学习模块, 采用文献[47]提出的文本实体链接方法, 并利用 YAGO 和 Probase 这两个知识图谱来获取实体的概念解释。YAGO 是一个大规模的语义知识库, 覆盖了多个领域的实体和关系, 可以提供广泛的语义背景支持。Probase 则以概率分布的形式表示常识知识, 帮助模型理解实体之间的关联程度。在点互信息的计算过程中, 分别将英文和中文的滑动窗口大小设置为 6 和 9。采用一个 2 层的加权图注意力网络, 其中第一层和第二层的输出维度分别为 128 和 64。为了提取中文和英文的词向量表示, 使用预训练的词向量库, 即 GoogleNews-Vectors 和 Chinese-World-Vectors, 它们的向量维度均为 300。在评论立场加权学习模块, 设定 Cross-modal Transformer 的层数为 4, 多头个数为 2, 每一层的隐藏层单元数为 64。在结果分类模块中, 采用一个 2 层的感知机, 其中第一层和第二层的输出维度分别为 64 和 32。在模型训练过程中, 采用 Adam 优化器, 学习率设置为 0.001, batch_size 为 16, 并进行 5 折交叉验证。实验基于 Python 3.8 和 PyTorch 深度学习框架来实现, 具体

的模型和训练超参数设置如表 2 所示。

表 2 超参数设置

Table 2 Hyperparameters setting		
参数类型	参数名称	参数值
模型参数	英文滑动窗口大小	6
	中文滑动窗口大小	9
	W-GAT 层数	2
	W-GAT 的输出维度	128,64
	中英文词向量维度	300
	Cross-modal Transformer 层数	4
	Cross-modal Attention 多头个数	2
	隐藏层单元数	64
	感知机层数	2
	感知机输出维度	64,32
训练参数	学习率	1e-3
	正则化参数	1e-5
	Batch size	16
	Dropout rate	0.1
	最大训练轮次	30
	早停等待次数	8
	K-fold 交叉验证	5
	优化器 Optimizer	Adam

3.2 实验结果和分析

3.2.1 超参数实验分析

在点互信息的计算过程中，滑动窗口的大小是一个关键的超参数，本文使用网格搜索来确定最佳滑动窗口大小。借鉴 Word2Vec 模型滑动窗口选择的实践经验，并在考虑到计算资源限制和训练效率的情况下，设定搜索范围为{3, 6, 9, 12, 15, 18, 21}。这个范围可以涵盖大多数任务的需求。TSCS 方法在不同滑动窗口下的实验结果如图 5 和图 6 所示。从图中可以看出，选择过小或者过大的滑动窗口都会对 TSCS 方法的性能产生不利影响。当滑动窗口设置得太小时，上下文信息不足，难以捕捉到两个词之间的关联性。相反，当滑动窗口过大时，会包含大量无关信息，可能淹没了两个词之间的关键信息。在数据集 Weibo 和 PHEME 中，最佳滑动窗口大小分别为 9 和 6。这种差异可能是因为英文通常比中文的语法结构

更为紧凑，表达方式更加简洁。

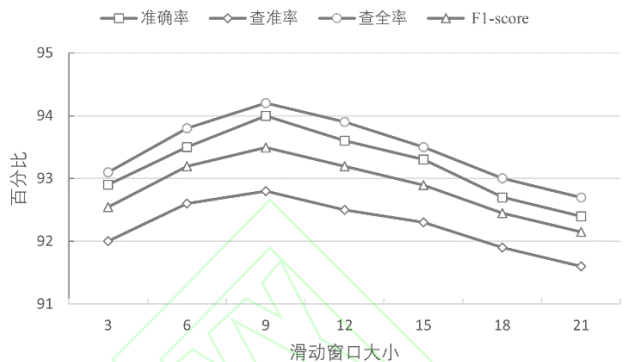


图 5 在 Weibo 上的实验结果
Fig.5 Experimental results on Weibo

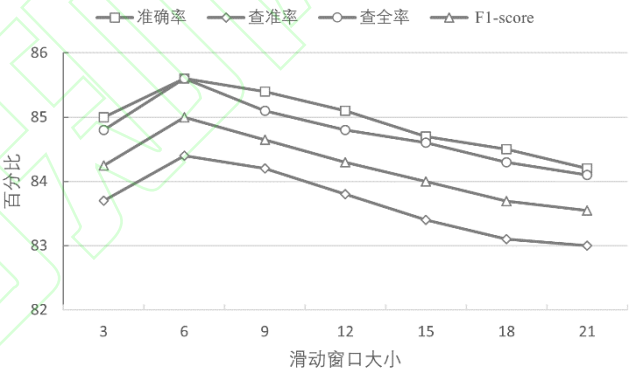


图 6 在 PHEME 上的实验结果
Fig.6 Experimental results on PHEME

3.2.2 方法对比分析

为了验证本文方法的有效性，选择表 3 中的方法作为基线进行对比。在 Weibo 和 PHEME 数据集上的实验结果分别如表 4 和表 5 所示，表中黑体字表示最优的性能。

表 3 基线方法

Table 3 Baseline methods				
基线方法	算法	内容特征	传播特征	用户特征
SVM-TS ^[3]	SVM	✓	✓	✓
GRU-2 ^[25]	GRU		✓	
LSTM-CNN ^[10]	LSTM+CNN	✓		
TDRD ^[13]	CNN	✓		
LSTM-Attention ^[15]	LSTM+Attention	✓		✓
GAN_based ^[12]	GAN	✓		
EGCN ^[32]	GCN	✓	✓	
User-Reply-GCN ^[41]	GCN		✓	✓
GAGK-RD ^[14]	GAT	✓		✓

根据表 4 和表 5 中的实验结果，发现所有数据

驱动的深度学习方法都优于依赖手工特征工程的统计机器学习方法 SVM-TS。这是因为手工设计的特征受到限制，既缺乏灵活性又不够全面，难以捕捉谣言的深层语义和传播特征。深度学习模型则通过数据驱动的方式自动学习适用于任务的高级抽象特征，具有更好的适应性和更丰富的多样性。基于传播特征的方法普遍优于基于内容特征的检测方法。这是因为谣言散布者经常刻意模仿真实消息的写作风格，并且社交媒体上的帖子通常非常精简，上下文信息有限，导致缺乏有效的内容分类特征。谣言和非谣言在传播过程中具有一定的差异，而这种差异不受帖子内容形式的影响。然而，基于传播特征的方法在早期检测可能效果不佳，通常需要谣言在社交网络中获得一定程度的传播才能实现较好的效果。

在内容特征的基础上引入用户特征的 LSTM-Attention 方法的性能优于仅依赖内容特征的 LSTM-CNN、TDRD、和 GAN-based 三个方法，验证了用户特征对谣言识别的积极作用。在传播特征的基础上融合内容特征的 EGCN 和融合用户特征的 User-Reply-GCN 两种方法的性能均超越了仅依赖传播特征的 GRU-2，说明了在传播特征中融合内容特征和用户特征的有效性。此外，集成了背景知识的 GAGK-RD 方法在谣言识别上表现出色，突显了背景知识的重要性。综合而言，本文的 TSCS 方法将用户特征、内容特征和传播特征有机融合，并通过知识图谱增强内容的语义理解，在各项指标上取

得最佳性能，验证了该方法的有效性。

表 4 在 Weibo 上的对比结果

Table 4 Experimental comparison results on Weibo

方法	准确率	查准率	查全率	F1-score
SVM-TS	84.4	86.0	85.3	85.6
LSTM-CNN	85.2	85.7	86.1	85.9
TDRD	86.7	87.0	86.0	86.5
GAN-based	86.8	86.5	89.4	87.9
LSTM-Attention	89.6	89.9	91.6	90.7
GRU-2	90.1	87.6	91.6	89.6
EGCN	91.3	89.9	92.0	90.9
User-Reply-GCN	91.6	90.2	92.3	91.2
GAGK-RD	92.1	91.0	92.4	91.7
TSCS	94.0	92.8	94.2	93.5
Δ_{SOTA}	$\uparrow 1.9$	$\uparrow 1.8$	$\uparrow 1.8$	$\uparrow 1.8$

表 5 在 PHEME 上的对比结果

Table 5 Experimental comparison results on PHEME

方法	准确率	查准率	查全率	F1-score
SVM-TS	78.3	69.2	73.1	71.1
LSTM-CNN	80.4	80.1	81.1	80.6
TDRD	82.7	81.3	78.6	79.9
GAN-based	82.7	81.6	79.1	80.3
LSTM-Attention	83.0	82.3	81.6	81.9
GRU-2	82.6	82.1	81.0	81.5
EGCN	83.5	82.5	82.4	82.4
User-Reply-GCN	83.6	82.2	83.3	82.7
GAGK-RD	84.0	83.2	82.2	82.7
TSCS	85.6	84.4	85.6	85.0
Δ_{SOTA}	$\uparrow 1.6$	$\uparrow 1.2$	$\uparrow 2.3$	$\uparrow 2.3$

3.2.3 消融实验分析

为了验证不同模块对 TSCS 方法的贡献，设计了 8 组消融实验。消融实验①和②分别去除 TSCS 方法中的语义特征 F_s 和评论立场时序特征 F_c ；消融实验③是 TSCS 方法在学习评论立场时序特征的过程中不考虑评论用户的信息；消融实验④是 TSCS 方法在学习文本语义特征的过程中不使用实体概念来增强内容的语义理解；消融实验⑤将 TSCS 方法中基于加权图注意力网络的文本语义特征学习方法替换为传统的基于循环神经网络的方法；消融实验⑥将 TSCS 方法中的加权图注意力神经网络 W-GAT 替换成普通的图注意力神经网络 GAT；消融实验⑦将基于 W-GAT 学习文本语义特征的方法替换成预训练语言模型 BERT；消融实验

⑧把语义特征 F_s 和评论立场时序特征 F_c 自适应的 所示。

融合方式替换成传统的相加方式。实验结果如表 6

表 6 消融实验结果
Table 6 Ablation experiment results

消融实验	Weibo				PHEME			
	Δ acc	Δ pre	Δ rec	Δ F1	Δ acc	Δ pre	Δ rec	Δ F1
① (-) F_s	-1.9	-1.4	-1.2	-1.3	-1.6	-1.6	-1.5	-1.5
② (-) F_c	-2.1	-2.3	-2.2	-2.2	-1.9	-1.8	-1.7	-1.8
③ (-) Users	-1.0	-0.7	-0.8	-0.8	-0.8	-0.9	-0.7	-0.8
④ (-) Context	-0.8	-0.6	-0.7	-0.7	-0.7	-0.8	-0.7	-0.7
⑤ W-GAT→RNN	-0.7	-0.7	-0.6	-0.7	-0.8	-0.8	-0.5	-0.7
⑥ W-GAT→GAT	-0.5	-0.4	-0.4	-0.4	-0.6	-0.6	-0.5	-0.5
⑦ W-GAT→BERT	-0.6	-0.5	-0.6	-0.5	-0.7	-0.7	-0.5	-0.6
⑧ Adaptive→Add	-0.8	-0.7	-0.8	-0.8	-0.9	-0.8	-0.6	-0.7

从表 6 中，可以发现每个模块都发挥着独特的作用，去除或替换任何一个都将对 TSCS 方法的性能产生影响。增强的语义特征 F_s 有效地捕捉了帖子的深层语义信息，因此，消除这一特征将严重影响方法的性能。通过立场检测模型和 Cross-modal Transformer 学习得到的评论立场时序特征 F_c 融合了用户特征和传播特征，对谣言的判断发挥了重要作用，去除它导致方法的性能下降最为显著。用户的认证状态、粉丝量、关注量、注册时间和发文量等特征能够反映用户的权威性，使模型更加专注于权威用户对帖子的立场，因此，消除用户特征也会降低方法的性能。不使用实体概念来增强内容的语义理解会降低方法的性能，因为社交媒体上的帖子通常信息简洁，上下文信息相对较少，难以提供详细的语义信息。将基于图注意力网络的文本语义特征学习方法替换为传统的基于循环神经网络或基于预训练语言模型 BERT 的方法后，TSCS 方法的性能下降，因为社交帖子内容经常呈现碎片化和离散化的特点，而图神经网络能更好地学习文本的非连续语义特征。将 W-GAT 替换为 GAT 后，方法的性能也有所下降，因为 W-GAT 在计算注意力系数时不仅考虑了节点之间的特征相似性，还将边的权重作为输入，从而能够更准确地获得全局信息。通过门控网络的融合方式可以根据不同特征的重要程度自动调整融合权重，比直接相加的融合方式更具有自适应性，因此模型的性能也会更好。

3.2.4 早期检测分析

为了验证 TSCS 方法在早期谣言检测任务中的性能。选取帖子发布后的 9 个时间点（1h、3h、6h、12h、24h、36h、48h、72h、96h）作为检测点。在每个检测点，使用该时间点之前的评论数据作为输入模型的数据。选择基于传播特征的 SVM-TS、GRU-2、EGCN 和 User-Reply-GCN 作为基线方法，对比实验结果如图 7 和图 8 所示。

随着时间的推移，各方法的准确率总体上呈现不同程度的提高。GRU-2、EGCN、User-Reply-GCN 和 TSCS 四个方法相较于 SVM-TS，表现出更快速和更稳定的准确率上升。在经过 36 小时后，它们的准确率基本维持在稳定水平，而 SVM-TS 则需要 76 个小时，这表明了深度学习方法相比于依赖特征工程的方法在早期谣言检测中也具有优势。在 1-24 小时的早期检测阶段，EGCN、User-Reply-GCN 和 TSCS 这三个方法的准确率明显高于 GRU-2 的准确率，因为前面三个方法融合了内容和用户特征，能够减少对传播特征的依赖，从而提高早期的检测性能。TSCS 在各个时间检测点上的准确率均高于其他四个模型，并且在帖子发布后的 24 小时内就实现了良好的检测效果，而其它方法需要 36 小时以上。这说明了本文模型在早期谣言检测任务中的有效性，比基线方法能提前至少 12 小时检测到谣言。

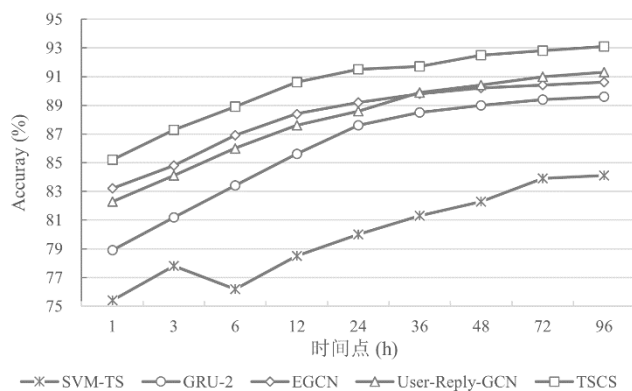


图7 在 Weibo 上的早期检测

Fig.7 Early detection on Weibo

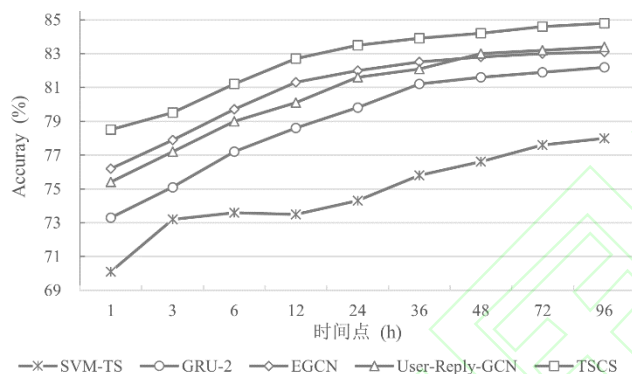


图8 在 PHEME 上的早期检测

Fig.8 Early detection on PHEME

4 结束语

网络谣言检测在遏制谣言传播方面具有重要意义。针对社交媒体帖子语义信息匮乏和大多数方法未考虑评论用户特征的问题,本文提出了一种基于文本语义增强和评论立场加权的网络谣言检测方法。通过外部知识图谱获取帖子文本中的实体概念,以提供更多上下文信息,从而增强语义理解。同时,采用预训练的立场检测模型提取每条评论的立场信息,根据评论用户特征为不同立场信息分配适当的权重。实验结果表明,本文方法在两个公开的社交媒体谣言数据集上不仅准确率高出基线方法 1.6% 以上,而且在早期谣言检测方面,比基线方法至少提前 12 个小时。然而,本文的方法还有改进的空间,例如,只考虑了文本特征而未涵盖可能存在的图片或视频信息,因此如何将多模态内容信息融合到模型之中是下一步的研究方向。

参考文献:

[1] 谢欣彤, 胡悦阳, 刘讚哲, 等. 传播用户代表性特征学习

的谣言检测方法[J]. 计算机科学与探索, 2022, 16(6): 1334-1342.

XIE X T, HU Y Y, LIU X Z, et al. Rumor detection based on representative user characteristics learning through propagation[J]. Journal of Frontiers of Computer Science and Technology, 2022, 16(6): 1334-1342.

[2] ZHENG P, HUANG Z, DOU Y, et al. Rumor detection on social media through mining the social circles with high homogeneity[J]. Information Sciences, 2023, 642: 119083.

[3] 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, Melbourne, Australia, Oct 18-23, 2015. New York: ACM, 2015: 1751-1754.

[4] 黄学坚, 马廷淮, 王根生. 基于分层语义特征学习模型的微博谣言事件检[J]. 数据分析与知识发现, 2023, 7(5): 81-91.

HUANG X J, MA T H, WANG G S. Weibo rumor events detection based on hierarchical semantic feature learning model[J]. Data Analysis and Knowledge Discovery, 2023, 7(5): 81-91.

[5] ZHOU H, MA T, RONG H, et al. MDMN: Multi-task and domain adaptation based multi-modal network for early rumor detection[J]. Expert Systems with Applications, 2022, 195: 116517.

[6] MA T, ZHOU H, TIAN Y, et al. A novel rumor detection algorithm based on entity recognition, sentence reconfiguration, and ordinary differential equation network[J]. Neurocomputing, 2021, 447: 224-234.

[7] FANG L, FENG K, ZHAO K, et al. Unsupervised rumor detection based on propagation tree VAE[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(10): 10309-10323.

[8] 汪建梅, 彭云, 余晨钰. 融合时间序列与卷积神经网络

- 的网络谣言检[J]. 小型微型计算机系统, 2022, 43(5): 1020-1026.
- WANG J M, PENG Y, YU C Y. Network rumor detection combining time series and convolutional neural network[J]. Journal of Chinese Computer System, 2022, 43(5):1020 - 1026.
- [9] KALIYAR R K, GOSWAMI A, NARANG P, et al. FNDNet: A deep convolutional neural network for fake news detection[J]. Cognitive Systems Research, 2020, 61: 32-44.
- [10] AJAO O, BHOWMIK D, ZARGARI S. Fake news identification on Twitter with hybrid CNN and RNN models[C]//Proceedings of the 9th International Conference on Social Media and Society, Copenhagen, Denmark, Jul 18-20, 2018. New York: ACM, 2018: 226-230.
- [11] ALKHODAIR S A, DING S H H, FUNG B C M, et al. Detecting breaking news rumors of emerging topics in social media[J]. Information Processing & Management, 2020, 57(2): 102018.
- [12] CHENG M, LI Y, NAZARIAN S, et al. From rumor to genetic mutation detection with explanations: A GAN approach[J]. Scientific Reports, 2021, 11(1): 5861.
- [13] XU F, SHENG V S, WANG M. Near real-time topic-driven rumor detection in source microblogs[J]. Knowledge-Based Systems, 2020, 207: 106391.
- [14] 王根生, 朱奕, 李胜. 一种融合知识图谱的图注意力神经网络谣言实时检测方法[J/OL]. 数据分析与知识发现 (2023-06-07)[2024-02-10]. <http://kns.cnki.net/kcms/detail/10.1478.G2.20230606.1719.008.html>.
- WANG G S, ZHU Y, LI S. A real-time rumor detection method based on graph attention neural network integrated with knowledge graph[J/OL]. Data Analysis and Knowledge Discovery (2023-06-07) [2024-02-10]. <http://kns.cnki.net/kcms/detail/10.1478.G2.20230606.1719.008.html>.
- [15] SINGH J P, KUMAR A, RANA N P, et al. Attention-based LSTM network for rumor veracity estimation of tweets[J]. Information Systems Frontiers, 2020, 24: 459-474.
- [16] 黄学坚, 王根生, 罗远胜, 等. 融合多元用户特征和内容特征的微博谣言实时检测模型[J]. 小型微型计算机系统, 2022, 43(12): 2518-2527.
- HUANG X J, WANG G S, LUO Y S, et al. Weibo rumors real-time detection model based on fusion of multi user features and content features[J]. Journal of Chinese Computer System, 2022, 43(12): 2518-2527.
- [17] 刘金硕, 冯阔, Jeff Z.Pan, 等. MSRD:多模态网络谣言检测方法[J]. 计算机研究与发展, 2020, 57(11): 2328-2336.
- LIU J S, FENG K, JEFF Z P, et al. MSRD: Multi-modal web rumor detection method[J]. Journal of Computer Research and Development, 2020, 57(11): 2328-2336.
- [18] CHEN J, WU Z, YANG Z, et al. Multimodal fusion network with contrary latent topic memory for rumor detection[J]. IEEE MultiMedia, 2022, 29(1): 104-113.
- [19] HUANG X, MA T, JIA L, et al. An effective multimodal representation and fusion method for multimodal intent recognition[J]. Neurocomputing, 2023, 548: 126373.
- [20] YADAV A, GABAS, KHAN H, et al. ETMA: Efficient transformer-based multilevel attention framework for multimodal fake news detection[J]. IEEE Transactions on Computational Social Systems, 2023, DOI: 10.1109/TCSS. 2023.3255242.
- [21] 刘华玲, 陈尚辉, 曹世杰, 等. 基于多模态学习的虚假新闻检测研究[J]. 计算机科学与探索, 2023, 17(9): 2015-2029.
- LIU H L, CHENG S H, CAO S J, et al. Survey of fake news detection with multi-model learning[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(9): 2015-2029.
- [22] DAVOUDI M, MOOSAVI M R, SADREDDINI M H. DSS: A hybrid deep model for fake news detection using propagation tree and stance network[J]. Expert Systems with Applications, 2022, 198: 116635.
- [23] ZHANG P, RAN H, JIA C, et al. A lightweight propagation

- path aggregating network with neural topic model for rumor detection[J]. *Neurocomputing*, 2021, 458: 468-477.
- [24] MA J, GAO W, WONG K F. Detect rumors in microblog posts using propagation structure via kernel learning[C]// *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, Vancouver, Canada, Jul 30-Aug 4, 2017. Stroudsburg: ACL, 2017:708-717.
- [25] 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, July 9-15, 2016. Palo Alto: AAAI, 2016: 3818-3824.
- [26] 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*, Melbourne, Australia, July 15-20,2018. Stroudsburg: ACL, 2018:1980- 1989.
- [27] MA J, GAO W, JOTY S, et al. An attention-based rumor detection model with tree-structured recursive neural networks[J]. *ACM Transactions on Intelligent Systems and Technology*, 2020, 11(4): 1-28.
- [28] CHEN T, LI X, YIN H, et al. Call attention to rumors: Deep attention based recurrent neural networks for early rumor detection[C]//*Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Melbourne, Australia, Jun 3, 2018. Berlin, Heidelberg: Springer, 2018: 40-52.
- [29] 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.
- HU D, WEI L W, ZHOU W, et al. A rumor detection approach based on multi-relational propagation tree[J]. *Journal of Computer Research and Development*, 2021, 58(7): 1395-1411.
- [30] HUANG Q, ZHOU C, WU J, et al. Deep spatial-temporal structure learning for rumor detection on Twitter[J]. *Neural Computing and Applications*, 2023,35: 12995-13005.
- [31] YU K, JIANG H, LI T, et al. Data fusion oriented graph convolution network model for rumor detection[J]. *IEEE Transactions on Network and Service Management*, 2020, 17(4): 2171-2181.
- [32] BAI N, MENG F, RUI X, et al. Rumor detection based on graph convolutional neural net[J]. *IEEE Access*, 2021, 9: 21686-21693.
- [33] 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.
- [34] LIU X, ZHAO Z, ZHANG Y, et al. Social network rumor detection method combining dual-attention mechanism with graph convolutional network[J]. *IEEE Transactions on Computational Social Systems*, 2022, 10(5): 2350-2361.
- [35] ALSAIF H F, ALDOSSARI H D. Review of stance detection for rumor verification in social media[J]. *Engineering Applications of Artificial Intelligence*, 2023, 119: 105801.
- [36] ZHU Y, WANG G, LI S, et al. A novel rumor detection method based on non-consecutive semantic features and comment stance[J]. *IEEE Access*, 2023, 11: 58016 -58024.
- [37] LI C, PENG H, LI J, et al. Joint stance and rumor detection in hierarchical heterogeneous graph[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 33(6): 2530-2542.
- [38] YE K, PIAO Y, ZHAO K, et al. Graph enhanced BERT for stance-aware rumor verification on social media[C]//*Proceedings of the 30th International Conference on Artificial Neural Networks*, Bratislava, Slovakia, Sep 14-17, 2021. Berlin, Heidelberg: Springer, 2021: 422-435.
- [39] HAN X, HUANG Z, LU M, et al. Rumor verification on social media with stance-aware recursive tree[C]// *Proceedings of the 14th International Conference on Knowledge Science, Engineering and Management*, Tokyo, Japan, Aug 14-16, 2021. Berlin, Heidelberg: Springer, 2021: 149-161.

[40] TU K, CHEN C, HOU C, et al. Rumor2vec: A rumor detection framework with joint text and propagation structure representation learning[J]. Information Sciences, 2021, 560: 137-151.

[41] LOTFI S, MIRZAREZAEI M, HOSSEINZADEH M, et al. Detection of rumor conversations in Twitter using graph convolutional networks[J]. Applied Intelligence, 2021, 51: 4774-4787.

[42] SUN M, ZHANG X, ZHENG J, et al. DDGCN: Dual dynamic graph convolutional networks for rumor detection on social media[C]//Proceedings of the 36th AAAI Conference on Artificial Intelligence, California, USA, Feb 22-Mar 1, 2022. Palo Alto: AAAI, 2022: 4611-4619.

[43] XU S, LIU X, MA K, et al. Rumor detection on social media using hierarchically aggregated feature via graph neural networks[J]. Applied Intelligence, 2023, 53(3): 3136-3149.

[44] CHEN X, ZHOU F, TRAJCEVSKI G, et al. Multi-view learning with distinguishable feature fusion for rumor detection[J]. Knowledge-Based Systems, 2022, 240: 108085.

[45] INAN E. ZoKa: A fake news detection method using edge-weighted graph attention network with transfer models[J]. Neural Computing and Applications, 2022, 34(14): 11669-11677.

[46] TSAI Y H H, BAI S, LIANG P P, et al. Multimodal transformer for unaligned multimodal language sequences[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, Jul 28-Aug 2, 2019. Stroudsburg: ACL, 2019: 6558-6569.

[47] CHEN L, LIANG J, XIE C, et al. Short text entity linking with fine-grained topics[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Torino, Italy, Oct 22-26, 2018. New York: ACM, 2018: 457-466.



朱奕 (1999—), 男, 江西南昌人, 硕士研究生, 主要研究方向为网络舆情。
ZAU Yi, born in 1999, M.S. candidate. His research interest is internet public opinion.



王根生 (1974—), 男, 江西南昌人, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为网络舆情和数据挖掘。
WANG Gensheng, born in 1974, Ph.D., professor, Ph.D. supervisor, CCF member. His research interests include internet public opinion and data mining.



金文文 (1994—), 女, 江西九江人, 硕士研究生, 主要研究方向为网络舆情。
JIN Wenwen, born in 1994, M.S. candidate. Her research interests is internet public opinion.



黄学坚 (1990—), 男, 江西南昌人, 博士研究生, 讲师, CCF 会员, 主要研究方向为网络舆情和数据挖掘。
HUANG Xuejian, born in 1990, Ph.D. candidate, lecturer, CCF member. His research interests include internet public opinion and data mining.



李胜 (1983—), 男, 江西铜鼓人, 博士, 副教授, 博士生导师, 主要研究方向为社会治理。
LI Sheng, born in 1983, Ph.D., associate professor, Ph.D. supervisor. His research interests is social management.