

基于贴文级特征融合的社交网络谣言检测方法

余潇龙 郭天成 陈阳 王新

(复旦大学计算机科学技术学院 上海 201203)

(上海市智能信息处理重点实验室 上海 201203)

摘要 现有的谣言检测方法,对贴文语义和贴文发布者信息、贴文传播状态等社交网络属性之间的关联利用不足,导致检测性能受限。针对这一不足,提出一种基于贴文级特征融合的谣言检测模型 PF-HAN。模型通过带注意力机制的双向长短期记忆网络(Bi-LSTM)生成各条贴文的语义表示,并与对应贴文的社交网络特征进行拼接以保留两者的对应关系。融合得到的贴文综合表示以序列形式输入分层注意力网络提取时序特征,生成最终的事件表示用于谣言判别。实验结果表明,模型在新浪微博和 Twitter 数据集中进行谣言检测任务时 F1 值达到 0.956 和 0.740,且能以高准确率完成谣言早期检测任务。

关键词 深度学习 自然语言处理 谣言检测 特征融合

中图分类号 TP391

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.08.027

SOCIAL NETWORKS RUMOR DETECTION APPROACH BASED ON POST-LEVEL FEATURE FUSION

Yu Xiaolong Guo Tiancheng Chen Yang Wang Xin

(School of Computer Science, Fudan University, Shanghai 201203, China)

(Shanghai Key Laboratory of Intelligent Information Processing, Shanghai 201203, China)

Abstract The existing rumor detection methods largely neglect the correlation between post semantics, post publishers and post propagation status, which lead to low detection rates. To solve this problem, this paper proposes a rumor detection approach PF-HAN based on post-level feature fusion. The model used a Bi-LSTM with attention mechanism to generate the semantic representation of each post, and spliced it with the social network features of the corresponding post to preserve the correspondence between them. The integrated representation of the posts obtained by the fusion was input into the hierarchical attention network in the form of sequence to extract the temporal features and generate the final event representation for rumor discrimination. Experimental results over Sina Weibo and Twitter show that the F1 value of the model reaches 0.956 and 0.740 when the model performs the rumor detection task and it can complete the early rumor detection task with high accuracy.

Keywords Deep learning Natural language processing Rumor detection Feature fusion

0 引言

谣言(rumor)是关于某个对象、事件或问题的未经证实、没有事实证据的主张或声明。谣言会引起各种严重的问题。它们会损害个人或组织的声誉,引起骚

乱和动荡,使公司忽然陷入灾难,还会动摇金融市场。因此,对谣言的有效管控非常有必要。

谣言传播是一个历史悠久的问题。谣言最早通过“口耳相传”的方式传播,后通过报纸广播等媒介传播。近年来,随着在线社交网络的发展,越来越多的人在社交平台上发布和接收信息,谣言也在这些平台传

播,且传播的速度更快、范围更广。多项研究^[1-3]表明在线社交网络已成为谣言传播的重要媒介。

由于谣言的危害性,在线社交网络平台和政府机构都会进行谣言检测和披露。传统的人工谣言检测依赖专家知识并消耗大量人力,检测速度慢且检测范围有限^[1]。因此自动化谣言检测技术的发展很迫切,也吸引了大量研究者的关注。

目前,多数研究者将社交网络谣言检测视为一个分类问题,最常见的分类方式为将待检测对象归入谣言和非谣言中的一类。根据检测对象的不同,现有工作可分为针对单条社交网络上的贴文进行判别和针对包含了一组相关贴文的舆论事件进行判别两种。本文的研究对象是谣言事件。

不同研究用于谣言判别使用的信息存在差异,部分研究在检测时只使用文本信息^[4-5],部分研究额外利用谣言的传播结构^[6-7],还有部分研究使用图片等多媒体数据辅助检测^[8]。

由于在线社交网络谣言以在线社交网络为平台传播,相比传统谣言增加了在线社交网络属性可供分析。例如,发布贴文的用户账号信息,发布贴文所使用的设备,贴文被其他用户点赞、转发的次数,因为贴文在社交网络平台上传播而产生的数据等都属于在线社交网络属性。

现有的利用文本信息和在线社交网络属性进行谣言检测的方法,虽然同时使用了这两类信息,但基本将这两类信息在较高层次进行结合。例如文献[9]和文献[10]将判定事件中用户声誉分布、发送贴文的设备构成作为特征,与贴文情绪倾向等文本特征组成特征向量输入分类器。文献[6]提取事件的社交网络特征作为查询向量,用于优化基于贴文文本信息构建的谣言事件表示。文献[11]使用文本信息和用户信息分别构建事件表示,之后共同决策进行谣言判别。

这些方法无法将一条特定贴文的文字语义与贴文发布者、贴文传播情况等进行关联。在这些模型中,声誉较高用户对声誉较低用户的贴文发布质疑信息和低声誉用户对高声誉用户发布的贴文进行质疑,会得到相同的事件表示。这一缺陷损害了模型的表达能力,限制了检测性能的提升。

针对这一问题,本文提出一种基于贴文级特征融合的分层注意力网络谣言检测模型(PF-HAN),该模型在单条贴文层面融合贴文的语义特征、贴文发布者特征、贴文的传播特征,将信息之间的对应关系保留,之后通过分层 LSTM-注意力网络获取待判定事件表

示,对事件是否为谣言进行判别。

为了验证谣言检测模型的有效性,本文使用新浪微博和 Twitter 两个具有代表性的在线社交网站的真实数据对谣言检测模型进行测试,并与多个典型的基线模型进行对比。为了探究模型中各结构对谣言检测性能的影响,本文设计消融实验并进行测试。

1 相关工作

早期的在线社交网络谣言检测方法基于传统机器学习。它们首先运用特征工程手段,从谣言相关贴文的文字内容中提取情感倾向^[12],从用户个人资料中提取用户影响力、声誉特征^[13-15],从帖子的传播方式^[10,16]、发布贴文的位置^[10]等提取特征。然后将待判定事件的特征输入决策树、支持向量机等分类器,训练模型后进行事件分类。由于各特征的计算是独立的,且事件特征多由其中贴文的对应特征取平均获得,这些工作对贴文文本内容和社交网络属性之间的对应关系利用较少。

随着深度学习的发展和自然语言处理技术的实用化,涌现了一批仅凭文本信息进行谣言检测的工作。Ma 等^[4]使用 RNN 提取待判定事件的文本表示,首次将深度学习模型用于在线社交网络谣言检测。Chen 等^[5]将注意力机制引入 RNN,提升了检测能力。Chen 等^[17]提出了一种基于 CNN 的残差注意力网络,它通过基于微调注意力机制的残差框架捕获事件中贴文间的长期依赖。

由于谣言的发布者常会学习模仿真实信息的写作风格以逃避检测,而谣言的传播模式与正常消息存在差异,研究者试图通过利用传播结构检测谣言。Liu 等^[18]通过谣言和非谣言在传播树深度和宽度方面的差异检测谣言。Ma 等^[7]提出了一种使用递归神经网络的树状模型,它通过自顶向下和自底向上两个树结构的递归神经网络将消息的传播过程在模型中直接体现。Bian 等^[19]以传播和散布是谣言的关键特征为理论基础,提出了基于双向图卷积网络的谣言检测模型。胡斗等^[20]提出了一种基于多关系传播树的检测模型,加强了对传播树中兄弟节点间关系的利用。

研究者也尝试在深度学习模型中加入社交网络属性用于检测。Guo 等^[6]构建了事件的社交网络特征,作为检测模型中注意力层的查询向量以提取更能体现事件谣言相关特性的文本片段。廖祥文等^[21]将局部用户信息和时间段内的贴文的时序信息相结合,提升了谣言检测的性能。Ruchansky 等^[11]构建了事件

参与用户的评分矩阵,与事件语义表示共同进行模型决策。

2 基于贴文级特征融合的谣言检测模型

2.1 问题定义和形式化

社交网络谣言检测任务旨在研究如何从社交网络中检测谣言,实质上是一个二分类问题,该任务的形式化定义如下:给定一个事件集合 $E = \{e_1, e_2, \dots, e_m\}$ 和一个类别集合 $L = \{l_1, l_2\}$, 其中: e_i 代表第 i 个待判定事件; m 为事件总数; l_1, l_2 分别代表谣言和非谣言这两个类别标签。事件 e_i 包含若干条贴文 $p_{i,j}$, 贴文 $p_{i,j}$ 包含贴文文本 $t_{i,j}$ 和贴文在线社交网络属性 $s_{i,j}$, 固有 $e_i = \{(p_{i,j})\} = \{(t_{i,j}, s_{i,j})\}$ 。社交网络谣言检测任务是学习一个分类模型 f , 将每个事件 e_i 映射成一个类别标签 l_j , 即 $f: e_i \rightarrow l_j$ 。模型输入包含若干条贴文的事件, 输出该事件对应的标签(谣言/非谣言)。

2.2 模型介绍

本文提出的基于贴文级特征融合的分层注意力网络谣言检测模型结构如图 1 所示。

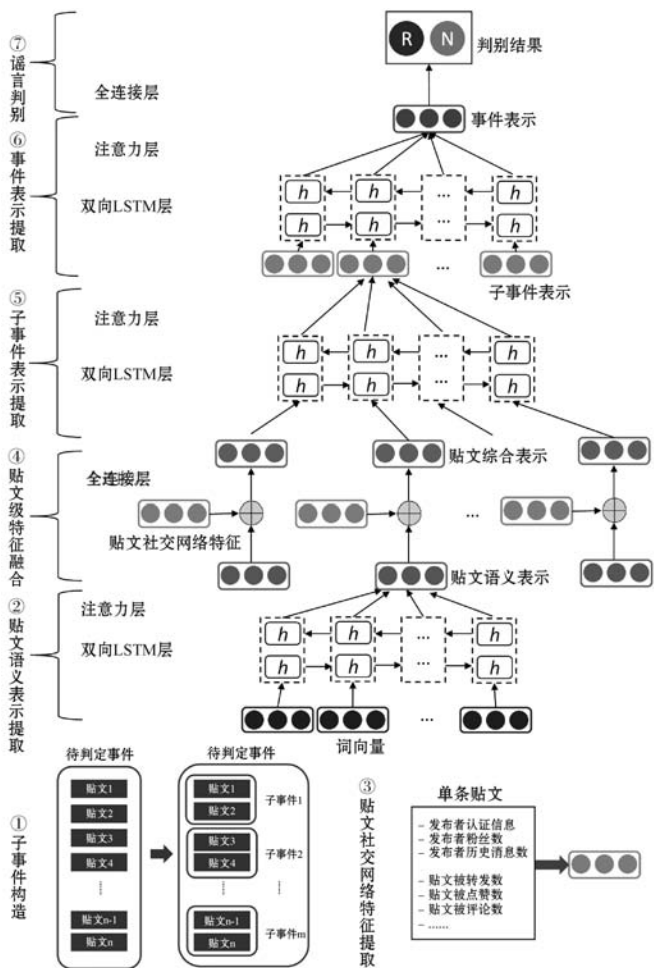


图 1 PF-HAN 模型结构

2.2.1 子事件分割

一个具备一定讨论热度的谣言事件往往包含大量的相关的贴文。将所有贴文直接输入网络意味着极大的计算负荷,并且容易引发长程依赖问题。现有的谣言检测模型往往选择将事件分为多个子事件,每个子事件包含一定数量的贴文,再进行后续运算。

该模块使用被广泛使用的子事件分割方法^[5-6],将原始事件分割成具有相近贴文数量的多个子事件。

2.2.2 贴文语义表示提取

在该模块中,我们首先将事件内的每条贴文文本中各单词嵌入到低维语义空间,用词向量进行表示。词向量 $w_{i,j,k}$ 对应第 i 个子事件中的第 j 条贴文的第 k 个单词。为了结合单词的上下文信息提取单词语义,我们将贴文词向量序列输入双向 LSTM 网络。网络接收贴文中各单词的词向量 $w_{i,j,k}$ 作为各时刻的输入,输出隐含状态 $h_{i,j,k}$ 。

构成贴文的各单词对贴文语义的贡献是存在差异的。因此,相比使用平均池化层处理隐层序列,我们引入了注意力机制,凸显被评估为对谣言检测意义较大的单词对贴文语义表示的影响。具体计算式如下:

$$\mathcal{F}_w(h_{i,j,k}) = v_w^T \tanh(W_{hw} h_{i,j,k} + b_w) \quad (1)$$

$$a_{i,j,k} = \frac{\exp(\mathcal{F}_w(h_{i,j,k}))}{\sum_j \exp(\mathcal{F}_w(h_{i,j,k}))} \quad (2)$$

$$c_{i,j} = \sum_k a_{i,j,k} h_{i,j,k} \quad (3)$$

式中: \mathcal{F}_w 为单词的评分函数,用于衡量各单词的重要程度; $h_{i,j,k}$ 为隐含状态; v_w 为权重向量; W_{hw} 为权重矩阵; b_w 为偏置项,它们在训练开始时被随机初始化; $a_{i,j,k}$ 为经标准化后的词权重; $c_{i,j}$ 为加权求和后得到的贴文语义表示。

2.2.3 贴文社交网络特征提取

在该模块中,我们逐贴文提取贴文的社交网络特征,包括贴文发布者特征和贴文传播学特征 2 类,具体特征和定义参见表 1 和表 2。

表 1 贴文发布者特征

特征	定义
个人资料完整度	是否填写个性签名,是否设置头像
信息认证情况	是否为网站认证用户
影响力	被关注人数
信息获取欲望	关注其他用户数
活跃度	发布的历史贴文数
账号注册时长	截至贴文发布日的账号注册天数

表2 贴文传播学特征

特征	定义
原创性	贴文是否原创
社区赞同度	贴文被点赞数
讨论热度	贴文被回复数
传播强度	贴文被转发数

由于在线社交网络依赖于具体的服务商(平台),因此不同平台上可以获取的特征存在差异,需要根据实际情况进行社交网络特征提取。

2.2.4 贴文级特征融合

在该模块中,我们在贴文层级融合贴文语义和贴文社交网络特征,使贴文的表示更适合用于后续的谣言事件检测任务。我们将贴文的语义表示向量 $\mathbf{c}_{i,j}$ 和社交特征向量 $\mathbf{z}_{i,j}$ 进行拼接,并使用全连接层将其映射到新的语义空间,得到贴文的综合表示 $\mathbf{p}_{i,j}$,具体计算公式如下:

$$\mathbf{p}_{i,j} = \mathbf{W}_f(\mathbf{c}_{i,j} \oplus \mathbf{z}_{i,j}) + \mathbf{b}_f \quad (4)$$

式中: $\mathbf{p}_{i,j}$ 为贴文综合表示向量; \mathbf{W}_f 为空间变换矩阵; $\mathbf{c}_{i,j}$ 为贴文语义表示向量; $\mathbf{z}_{i,j}$ 为贴文社交网络特征向量; \mathbf{b}_f 为偏置向量。

根据我们的调研,本文是最先使用该融合方式的谣言事件检测工作,后续的实验将验证该融合方式的有效性。

2.2.5 子事件表示提取

该模块接收各子事件中贴文的综合表示 $\mathbf{p}_{i,j}$,输入双向 LSTM-注意力网络后得到子事件表示 \mathbf{u}_i 。

双向 LSTM 网络以子事件 i 中贴文的综合表示 $\mathbf{p}_{i,j}$ 作为各时刻的输入,输出隐含状态 $\mathbf{h}_{i,j}$ 。语义和发送者不同的贴文对子事件含义的贡献存在差异,所以我们继续通过注意力网络获取子事件 \mathbf{u}_i 。具体计算式如下:

$$\mathcal{F}_p(\mathbf{h}_{i,j}) = \mathbf{v}_p^T \tanh(\mathbf{W}_{hp}\mathbf{h}_{i,j} + \mathbf{b}_p) \quad (5)$$

$$a_{i,j} = \frac{\exp(\mathcal{F}_p(\mathbf{h}_{i,j}))}{\sum_j \exp(\mathcal{F}_p(\mathbf{h}_{i,j}))} \quad (6)$$

$$\mathbf{u}_i = \sum_j a_{i,j} \mathbf{h}_{i,j} \quad (7)$$

式中: \mathcal{F}_p 为贴文的评分函数,用于衡量各贴文的重要程度; $\mathbf{h}_{i,j}$ 为隐含状态; \mathbf{v}_p 为权重向量; \mathbf{W}_{hp} 为权重矩阵; \mathbf{b}_p 为偏置项,它们在训练开始时被随机初始化; $a_{i,j}$ 为经标准化后的贴文权重; \mathbf{u}_i 为加权求和后所得到的子事件表示。

虽然存在为子事件表示的注意力层增加事件相关的查询向量的工作^[6],但本文后续研究表明,在进行了贴文级特征融合后,不加入额外查询向量的基础注意

力网络可以获得更优的性能。

2.2.6 事件表示提取和模型求解

与子事件表示提取模块类似,该模块接收子事件表示序列 \mathbf{u}_i 输入双向 LSTM-注意力网络后得到待判定事件的表示向量 \mathbf{e} 。

通过全连接层将向量 \mathbf{e} 投射到包含谣言和非谣言两类的目标空间中:

$$\hat{\mathbf{e}} = \tanh(\mathbf{W}_c \mathbf{e} + \mathbf{b}_c) \quad (8)$$

式中: \mathbf{W}_c 、 \mathbf{b}_c 为权重矩阵和偏置向量。

使用 softmax 层计算谣言分布:

$$p = \text{softmax}(\hat{\mathbf{e}}) \quad (9)$$

我们使用预测概率分布和真实概率分布之间的交叉熵误差作为损失函数。为了避免过拟合,我们对原损失函数进行 L2 正则化,最终损失函数如下:

$$L = - \sum_{e \in E} [y^{(e)} \ln(p_r^{(e)}) + (1 - y^{(e)}) \ln(1 - p_r^{(e)})] + \frac{\lambda}{2} \|\theta\|_2^2 \quad (10)$$

式中: $y^{(e)}$ 为事件 \mathbf{e} 的真实标签,谣言事件和非谣言事件分别对应 1 和 0; $p_r^{(e)}$ 为预测得到的谣言类概率; E 为训练集中的所有事件; θ 为模型的参数集合; λ 为正则化系数。

3 实验

3.1 实验数据集

本文使用 Ma 等^[4]发布的谣言事件检测数据集进行实验。该数据集已被多篇文章^[5-6,11,21]使用,是在线社交网络谣言检测领域的经典数据集。该数据集包含大量来自新浪微博和 Twitter 这两个具有全球影响力的微博客平台的真实舆论事件及事件标签。新浪微博的具体贴文数据可从公开数据集中直接获取, Twitter 贴文的具体数据通过 Twitter API 访问。由于部分贴文已被删除或设置访问权限导致无法获取,数据集包含贴文数降低了 21.4%。最终数据集统计信息如表 3 所示。

表3 数据集统计信息

事件	新浪微博	Twitter
事件数	4 664	992
谣言事件数	2 313	498
非谣言事件数	2 351	494
事件平均贴文数	816	867
事件平均时长/h	1 811.1	2 523.8

3.2 实验环境

本文的实验环境如表 4 所示。

表 4 实验环境配置表

配置项	配置情况
CPU	Xeon(R) CPU E5-2680 v3 @ 2.50 GHz
内存容量/GB	128
GPU	GeForce GTX 1080Ti
显存容量/GB	12
操作系统	CentOS Linux release 7.5.1804
开发平台	Python 3.6.10
深度学习框架	PyTorch 1.4.0

3.3 实验设计与参数设置

为了验证模型的有效性,本文选择了 2 种使用文本信息的模型 ML-GRU^[4]、CallAtRumor^[5] 和 3 种同时使用文本信息和在线社交网络属性的模型 DTC^[9]、SVM-TS^[22]、HSA^[6] 作为基线方法,在新浪微博和 Twitter 两个数据集上和本文提出的 PF-HAN 模型进行谣言检测性能对比。本文采用了准确率(Accuracy)、查准率(Precision)、查全率(Recall)、F1 值(F1-score)共 4 个评价指标,与基线方法论文^[5-6]保持一致。

本文具体的实验参数设置如下:为了较好地与基线方法比较,本文在子事件分割时使用了与基线方法论文^[5-6]相同的参数设置,即子事件微博序列长度为 5,最少子事件数为 2。单条贴文最大单词数位 32,单个事件最大贴文数为 1 000。对于中文数据集和英文数据集分别使用 300 维、200 维预训练 GloVe 词向量^[23],并通过全连接层统一维度到 128。全连接层和 LSTM 层的单元数为 128,各层 dropout 率为 0.1。我们使用 Adam 优化器作为通过损失函数更新模型内部参数的优化器。相对于传统的梯度下降方法,Adam 提出了基于动量的梯度动态估计,在计算上十分高效,并且效果也更优秀。优化器使用动态学习率,初始学习率为 0.002,衰减系数为 0.6,衰减周期为 25。训练数据的批大小为 64。为了减少训练时的显存占用,使用梯度累加技术,实验单次批大小为 16,累计 4 次反向传播后调用一次优化器。迭代次数在新浪微博和 Twitter 中分别设置为 50、25。

为了尽量减少数据集分割造成的实验结果波动,本文参考文献[17,24]的划分方法使用 5 折交叉验证进行实验性能评估。各参数指标为 5 次实验的平均值。

3.4 实验结果与分析

表 5 和表 6 记录了各模型在新浪微博和 Twitter 数据集上的谣言检测效果。

表 5 实验结果表-新浪微博

模型	准确率	查准率	查全率	F1 值
ML-GRU	0.920	0.905	0.940	0.922
CallAtRumor	0.923	0.926	0.916	0.921
DTC	0.922	0.919	0.927	0.922
SVM-TS	0.915	0.912	0.916	0.914
SAT	0.933	0.932	0.932	0.932
PF-HAN	0.957	0.954	0.959	0.956

表 6 实验结果表-Twitter

模型	准确率	查准率	查全率	F1 值
ML-GRU	0.674	0.648	0.703	0.673
CallAtRumor	0.701	0.690	0.738	0.712
DTC	0.717	0.723	0.710	0.715
SVM-TS	0.626	0.641	0.585	0.610
SAT	0.692	0.694	0.693	0.692
PF-HAN	0.728	0.711	0.775	0.740

本文提出的 PF-HAN 模型在新浪微博和 Twitter 数据集上分别达到 0.957、0.728 的准确率和 0.956、0.740 的 F1 值,超过了所有基线方法。基线模型 SAT 凭借 RNN 对事件时序特征的捕捉和社交注意力机制对在线社交网络特征的利用在新浪微博实验中性能较优,但在数据噪声更多^[4]的 Twitter 数据集中表现不佳。相对地,PF-HAN 在 Twitter 数据集中依然可保持较高的性能,且对谣言的查全率显著高于基线模型。

3.5 早期谣言检测

在谣言检测相关的研究中,一个非常重要的目标是在尽可能早的时间检测到谣言^[24]。为了评估模型的早期谣言检测性能,我们设置了一组截止时间点,测试集中发送时间晚于截止时间点的贴文会被移除以模拟谣言传播的早期场景。图 2 和图 3 展现了早期谣言检测实验结果。

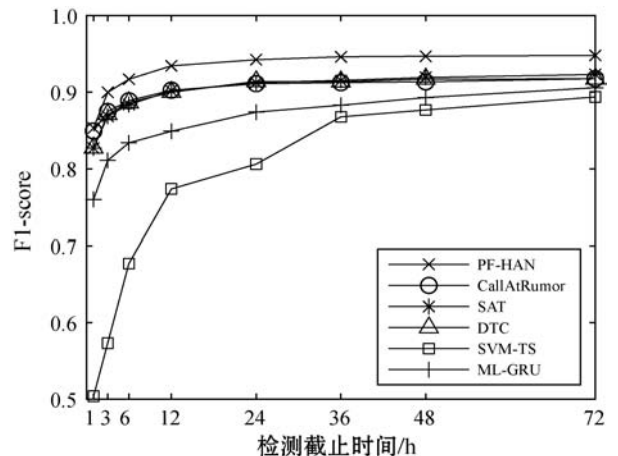


图 2 早期谣言检测实验结果-新浪微博

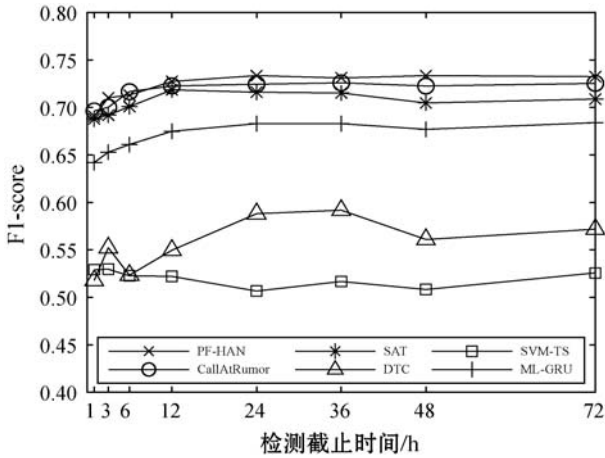


图3 早期谣言检测实验结果-Twitter

在新浪微博和 Twitter 数据集中,本文模型在各早期检测节点都取得了高于基线方法的 F1 值。在新浪微博的实验中,PF-HAN 在 6 h 的 F1 值已超过基线方法中最优的 SAT 在 72 h 达到的水平。在 Twitter 的实验中,PF-HAN 在 5 种基线方法在 48 h 的截止点出现 F1 值下滑的现象,而 PF-HAN 的检测性能依然稳定,说明 PF-HAN 相比基线方法具有更强的鲁棒性。

3.6 模型子结构影响

为了分析模型中各结构对谣言检测性能的影响,本文设计以下 PF-HAN 的变种模型:

- (1) HAN-textOnly: 不融合社交网络特征,只使用文本信息。
- (2) PF-HAN-linear: 在贴文级特征融合时使用线性叠加的方式进行特征融合。
- (3) EF-HAN: 在事件层面进行特征融合。
- (4) PF-HAN-SA: 在贴文级进行特征融合后,使用社交注意力结构^[6]替代本文的注意力结构。

表 7 和表 8 记录了各变种模型的谣言检测效果。

表 7 不同网络结构的影响-新浪微博

模型	准确率	查准率	查全率	F1 值
PF-HAN	0.957	0.954	0.959	0.956
HAN-textOnly	0.923	0.926	0.916	0.921
PF-HAN-linear	0.917	0.913	0.920	0.916
EF-HAN	0.934	0.952	0.908	0.929
PF-HAN-SA	0.938	0.932	0.944	0.938

表 8 不同网络结构的影响-Twitter

模型	准确率	查准率	查全率	F1 值
PF-HAN	0.728	0.711	0.775	0.740
HAN-textOnly	0.701	0.690	0.738	0.712
PF-HAN-linear	0.698	0.695	0.722	0.708
EF-HAN	0.718	0.698	0.745	0.721
PF-HAN-SA	0.720	0.702	0.753	0.727

在两个数据集中,没有使用社交网络特征的 HAN-textOnly 性能显著低于 PF-HAN,说明在贴文层级融合社交网络特征可以有效提升谣言检测性能。但是,使用不当的方式融合社交网络特征反而会降低模型性能,实例就是消融实验中使用线性融合策略的 PF-HAN-linear 性能低于 HAN-textOnly。在事件层面融合特征的 EF-HAN 性能优于 HAN-textOnly 但劣于 PF-HAN,表明在事件层面融合特征对谣言检测性能的提升幅度低于贴文级特征融合,验证了贴文层面进行特征融合的策略优于在事件层面进行特征融合。更换注意力策略的 PF-HAN-SA 的预测表现劣于 PF-HAN,说明 PF-HAN 模型在注意力策略选用上是合理而恰当的。

4 结 语

本文针对现有谣言检测模型对贴文语义和贴文社交网络特征之间的对应关系利用不足的问题,提出一种基于贴文级特征融合的分层注意力网络谣言检测方法 PF-HAN。与性能最好的基线方法相比,本文提出的模型在新浪微博数据集上将 F1 值提高 2.4 个百分点达到 0.956,在 Twitter 数据集上将 F1 值提高 2.5 个百分点达到 0.740,可以更有效地对谣言事件进行检测。在未来的工作中,我们将尝试在贴文层面完成特征融合的基础上,将贴文的传播结构信息加入模型,用于提高模型对谣言事件的检测效果。

参 考 文 献

- [1] Zubiaga A, Aker A, Bontcheva K, et al. Detection and resolution of rumours in social media: A survey [J]. ACM Computing Surveys, 2018, 51(2): 1-36.
- [2] Alzanin S M, Azmi A M. Detecting rumors in social media: A survey [J]. Procedia Computer Science, 2018, 142: 294-300.
- [3] 陈燕方,李志宇,梁循,等. 在线社会网络谣言检测综述 [J]. 计算机学报, 2018, 41(7): 1648-1677.
- [4] Ma J, Gao W, Mitra P, et al. Detecting rumors from microblogs with recurrent neural networks [C]//25th International Joint Conference on Artificial Intelligence, 2016: 3818-3824.
- [5] Chen T, Li X, Yin H Z, et al. Call attention to rumors: Deep attention based recurrent neural networks for early rumor detection [C]//Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2018: 40-52.
- [6] Guo H, Cao J, Zhang Y Z, et al. Rumor detection with hierarchical social attention network [C]//27th ACM International Conference on Information and Knowledge Management, 2018: 943-952.

- [7] 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.
- [8] Wang W Y. “Liar, Liar Pants on Fire”: A new benchmark dataset for fake news detection [C]//55th Annual Meeting of the Association for Computational Linguistics, 2017: 422–426.
- [9] Castillo C, Mendoza M, Poblete B. Information credibility on twitter [C]//20th International Conference Companion on World Wide Web, 2011: 675–684.
- [10] Kwon S, Cha M, Jung K, et al. Prominent features of rumor propagation in online social media [C]//13th International Conference on Data Mining, 2013: 1103–1108.
- [11] Ruchansky N, Seo S, Liu Y. CSI: A hybrid deep model for fake news detection [C]//26th ACM International Conference on Information and Knowledge Management, 2017: 797–806.
- [12] Zimbra D, Ghiassi M, Lee S. Brand-related twitter sentiment analysis using feature engineering and the dynamic architecture for artificial neural networks [C]//49th Hawaii International Conference on System Sciences, 2016: 1930–1938.
- [13] Wang Y, Lin X M, Wu L, et al. Effective multi-query expansions: Collaborative deep networks for robust landmark retrieval [J]. *IEEE Transactions on Image Processing*, 2017, 26(3): 1393–1404.
- [14] Wang Y, Lin X M, Wu L, et al. Effective multi-query expansions: Robust landmark retrieval [C]//23rd ACM International Conference on Multimedia, 2015: 79–88.
- [15] Zafarani R, Liu H. 10 Bits of surprise: Detecting malicious users with minimum information [C]//24th ACM International on Conference on Information and Knowledge Management, 2015: 423–431.
- [16] Liu X M, Nourbakhsh A, Li Q Z, et al. Real-time rumor debunking on twitter [C]//24th ACM International on Conference on Information and Knowledge Management, 2015: 1867–1870.
- [17] Chen Y X, Hu L, Sui J, et al. Attention-residual network with CNN for rumor detection [C]//28th ACM International Conference on Information and Knowledge Management, 2019: 1121–1130.
- [18] Liu Y H, Jin X L, Shen H W, et al. Do rumors diffuse differently from non-rumors? A systematically empirical analysis in sina weibo for rumor identification [C]//Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2017: 407–420.
- [19] Bian T, Xiao X, Xu T Y, et al. Rumor detection on social media with bi-directional graph convolutional networks [C]//34th AAAI Conference on Artificial Intelligence, 2020: 549–556.
- [20] 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.
- [21] 廖祥文,黄知,杨定达,等.基于分层注意力网络的社交媒体谣言检测[J].*中国科学:信息科学*,2018,48(11):1558–1574.
- [22] Ma J, Gao W, Wei Z Y, et al. Detect rumors using time series of social context information on microblogging websites [C]//24th ACM International on Conference on Information and Knowledge Management, 2015: 1751–1754.
- [23] Pennington J, Socher R, Manning C D. Glove: Global vectors for word representation [C]//Conference on Empirical Methods in Natural Language Processing, 2014: 1532–1543.
- [24] Zhao Z, Resnick P, Mei Q Z. Enquiring minds: Early detection of rumors in social media from enquiry posts [C]//24th International Conference on World Wide Web, 2015: 1395–1405.

~~~~~

(上接第 181 页)

- [ 15 ] 聂永辉,高磊,唐威. Hilbert-Huang 变换在电力系统暂态信号分析中的应用[J].*电力系统及其自动化学报*, 2009, 21(4): 63–69.
- [ 16 ] Wu Z H, Huang N E. Ensemble empirical mode decomposition: A noise assisted data analysis method [J]. *Advances in Adaptive Data Analysis*, 2009, 1(1): 1–41.
- [ 17 ] Yeh J R, Shieh J S, Huang N E. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method [J]. *Advances in Adaptive Data Analysis*, 2014, 2(2): 135–156.
- [ 18 ] 高佳程,田蕴卿,朱永利,等.基于完全集合经验模态分解和排列熵的局部放电信号的小波包去噪方法[J].*电力系统及其自动化学报*, 2018, 30(3): 1–7.
- [ 19 ] 赵荣珍,孙泽金.融合 CEEMD\_MPE 和 GK 模糊聚类的故障识别方法[J].*振动工程学报*, 2020, 33(3): 629–635.
- [ 20 ] Aziz W, Arif M. Multiscale permutation entropy of physiological time series [C]//9th International Multitopic Conference, 2005: 1–6.
- [ 21 ] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders [C]//25th International Conference on Machine Learning, 2008: 1096–1103.
- [ 22 ] 许倩文,吉兴全,张玉振,等.栈式降噪自编码网络在变压器故障诊断中的应用[J].*电测与仪表*, 2018, 55(17): 62–67.
- [ 23 ] 张建财,高军伟.基于变分模态分解和多尺度排列熵的滚动轴承故障诊断[J].*噪声与振动控制*, 2019, 39(6): 181–186.
- [ 24 ] 熊国良,甄灿壮,张龙,等.基于 CEEMDAN 多尺度排列熵的轴承故障智能识别 Fisher-GG 聚类方法[J].*噪声与振动控制*, 2020, 40(6): 1–7, 26.