基于传播树的多特征谣言检测方法

张鑫昕[1],潘善亮[1*],茅琴娇[2]
(1. 宁波大学信息科学与工程学院,浙江宁波 315211;2. 宁波工程学院网络空间安全学院,浙江宁波 315211)

摘 要: 目前网络谣言的检测方法主要是从传播路径中寻找信息,大多只采用文本信息作为初始传播特征,因此难以捕捉到丰富的传播结构表示.本文根据谣言的传播路径,提取文本和用户可信度特征,构建一种基于传播树的多特征谣言检测模型.模型通过图卷积网络聚合文本传播信息,使用多头注意力机制挖掘文本传播树的层内依赖关系,同时对用户传播树中的每个用户构建可信度序列,并采用M-Attention模块捕获有效的用户可信度特征.实验结果表明,本文提出的方法在Twitter15、Twitter16和Weibo数据集上的检测准确率达到89.3%、91.7%和96.4%,相比当前最优的传播树模型Bi-GCN(Binary Graph Convolutional Network)分别提升4.8%、4.2%和3%.

关键词: 谣言检测;传播结构;图卷积网络;注意力机制;自然语言处理
基金项目: 浙江省公益性技术应用研究计划项目(No.2017C33001)
中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2024)05-1609-10
电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20220616

A Rumor Detection Approach Based on Multi-Feature Propagation Tree

ZHANG Xin-xin[1], PAN Shan-liang[1*], MAO Qin-jiao[2]
(1. *College of Information Science and Engineering*, *Ningbo University*, *Ningbo*, *Zhejiang* 315211, *China*;
2. *College of Cyberspace Security*, *Ningbo University of Technology*, *Ningbo*, *Zhejiang* 315211, *China*)

Abstract: At present, rumor detection methods on social platforms mainly focus on obtaining information from the propagation path, most of these methods only use text information as the initial propagation feature, which is difficult to capture the rich propagation structure representation. In this paper, according to the propagation path of rumors, text and user credibility features are extracted, and a multi-feature rumor detection model based on propagation tree is constructed. This model aggregates text propagation features through a graph convolutional network, and uses a multi-head attention module to mine the intra-layer dependencies of the text propagation tree. At the same time, a credibility sequence is constructed for each user in the user propagation tree, and the M-Attention module is used to capture effective user credibility features. The experimental results show that the experimental accuracy of Twitter15, Twitter16 and Weibo datasets reaches 89.3%, 91.7% and 96.4%, which are 4.8%, 4.2% and 3% higher than the current optimal propagation tree model Bi-GCN (Binary Graph Convolutional Network) accuracy respectively.

Key words: rumor detection; propagation structure; graph convolutional network; attention mechanism; natural language processing
Foundation Item(s): Zhejiang Province Public Welfare Technology Application Research Program Project (No. 2017C33001)

1 引言

随着微博、微信、豆瓣等网络社交媒体的日益兴起,越来越多的人通过社交平台进行在线互动、获取新闻资讯[1].由于在社交平台上发布消息的成本低、速度快、传播广,导致出于各种不良目的的虚假消息和谣言报道在网络上肆意流传.据智利《第三版时报》网站报道,关于新冠病毒的谣言和阴谋论导致成百上千人死亡.这些谣言引发群众恐慌、危害公共安全、损害民众利益,对社会造成极大的影响.因此,近年来社交媒体上的谣言检测问题引起了越来越多的关注.

传统的谣言检测方法主要探究如何设计有效的特

收稿日期:2022-05-27;修回日期:2023-02-24;责任编辑:王天慧
*通讯作者:潘善亮

征来判别谣言，常见的方法一般侧重于使用文本特征，例如通过n-gram[2]、TF-IDF[3]、bag-of-word[4]对文本内容进行编码．也有一些方法研究原文的写作风格、小标题、词汇、句法结构[5,6]等特征．但这些方法都缺乏灵活性，面对复杂多样的真实环境，不能够很好地表示丰富的语义信息．

近年来，研究者们开始利用帖子在社交媒体中的传播过程，挖掘传播树中的序列特征，采用深度学习的方法构建高层次的传播表示．但是这种方法能够捕获到的传播特征往往具有一定的局限性，主要体现在以下几点：(1)只使用单一的节点特征构建传播树，无法获取丰富的传播表示；(2)只关注传播树中显式的父子传播关系，没有挖掘节点间存在的隐式依赖关系；(3)没有区分传播树中不同类型的节点信息，往往将所有节点信息视为同等重要．

针对上述研究方法存在的局限性，本文构建基于文本内容和用户可信度的多特征传播树，使用注意力模块融合两种不同的传播表示进行谣言检测．首先，采用图卷积网络挖掘父子传播关系，同时通过多头注意力机制挖掘兄弟节点之间的层内依赖关系．然后，使用XLNet[7]模型捕获源帖文本中更长距离的上下文动态特征，增强根节点信息的影响力．此外，以用户可信度作为节点特征，构建用户传播树，提取两种不同类型的用户可信度特征，丰富模型的传播表示，提升模型的检测性能．

本文的贡献可以总结如下：

(1)构建基于图卷积网络的文本传播表示，通过不断聚合邻接节点信息，提取父子传播特征；

(2)提出一种层次特征建模方法，挖掘文本传播树的层内依赖关系，捕获节点间的兄弟传播特征；

(3)构建基于注意力机制的用户传播表示，根据用户历史信息和传播路径生成用户可信度序列，提取传播过程中的用户可信度特征；

(4)提出一种特征融合方法，有效融合文本传播特征和用户传播特征．

2 相关工作

目前，谣言检测的方法主要包括：(1)以文本内容为主要特征的检测方法；(2)以用户信息为主要特征的检测方法；(3)以传播结构为主要特征的检测方法．

(1)基于文本内容的检测方法

传统的谣言检测方法主要依赖于谣言本身的文本内容．Ma等人[8,9]利用循环神经网络中的隐层向量表示谣言，还第一次将多任务学习的思想应用到谣言检测中，把谣言检测问题和立场分类问题组合成一个多任务模型．Chen等人[10]对Ma等人提出的循环神经网络模型进行了改进，捕获了帖子随时间产生的变化特征．夏鑫林等人[11]则将经过词嵌入的文本放入长短期记忆网络，对隐向量特征使用注意力机制加权求和得到最终的文本表示．

(2)基于用户信息的检测方法

基于用户信息的谣言检测方法主要利用用户的个人基本信息构造特征，如个人简介、性别、地区、关注数、历史数据等．Dou等人[12]利用新闻内容和用户历史帖子信息进行谣言检测．Lu等人[13]则将源文本和对应进行转发评论的用户序列构建成一张用户图，使用GCN[14](Graph Convolutional Network)提取特征进行谣言检测．Yuan等人[15]也结合用户关系捕获谣言的潜在特征．

(3)基于传播结构的谣言检测方法

基于传播结构的谣言检测方法通常提取源帖在传播过程中的时间序列和传播结构特征识别谣言．Kwon等人[16]提出通过检验谣言传播的时间、结构和语言特征来确定谣言的特性．Ke等人[17]提出一种混合分类器，通过混合分类器捕获谣言的高阶传播表示．Huang等人[18]构建了一种基于异构图的注意力网络，捕获文本内容的全局语义关系，以及源帖传播的全局结构信息．Shu等人[19]则考虑了传播过程中存在的多种关系，使用矩阵分解的方式获得各个传播节点的嵌入表示．Kang等人[20]考虑了源帖、发布时间、转发帖子、文本主题以及用户之间存在的多种关系，构建异构图挖掘传播特征．Bian等人[21]提出了一个新的双向图卷积网络Bi-GCN，通过自顶向下和自底向上的谣言传播树探索传播结构中的潜在特征．Wei等人[22]则改进了Bi-GCN中传播结构的不确定性，针对传播图中不同类型的边，动态分配边权值．

3 问题定义

针对谣言检测问题，本文给出以下问题定义和表示：

设$C=\{c_1,c_2,\cdots,c_n\}$表示谣言检测数据集，其中c_i表示数据集中的第i个谣言样本，n表示数据集中总样本数．每个谣言样本$c_i=\{x_{i0},x_{i1},\cdots,x_{im_i},u_{i0},u_{i1},\cdots,u_{is_i},G,G'\}$，其中$x_{i0}$表示源帖的文本内容，$x_{ij(j\neq0)}$表示第$j$个评论源帖的文本内容，$m_i$表示$c_i$样本中评论文本的总数．$u_{i0}$表示发布源帖的用户，$u_{ij(j\neq0)}$表示第$j$个转发或者评论源帖的用户，$s_i$表示$c_i$样本中转发评论源帖的用户总数．$G$表示基于文本特征构建的传播树，$G'$表示基于用户特征构建的传播树．

每条谣言事件c_i根据数据集可被标注为$y_i\in Y$，其中Y表示数据集中类别标签集合．谣言检测任务可以描述为一个学习函数f，记为

$$f: f(c_i) \to Y \tag{1}$$

其中，$Y=\{0,1,2,3\}$（Twitter 数据集）或者 $Y=\{0,1\}$（Weibo数据集）.

4　基于传播树的多特征谣言检测方法

本文提出一种基于传播树的多特征谣言检测方法——MPT（a rumor detection approach based on Multi-feature Propagation Tree），该方法主要包括基于文本的传播树表示、基于用户的传播树表示和谣言分类器3个部分，总体框架如图1所示.

在文本传播树模块，利用图卷积网络提取了父子传播特征，为了充分挖掘结构特征，还探究了同层节点之间的兄弟传播关系. 同时，源帖文本作为传播树的根节点，拥有相比其他节点更丰富的特征并且会影响传播树的扩散，因此对源帖还进行了额外的特征提取.

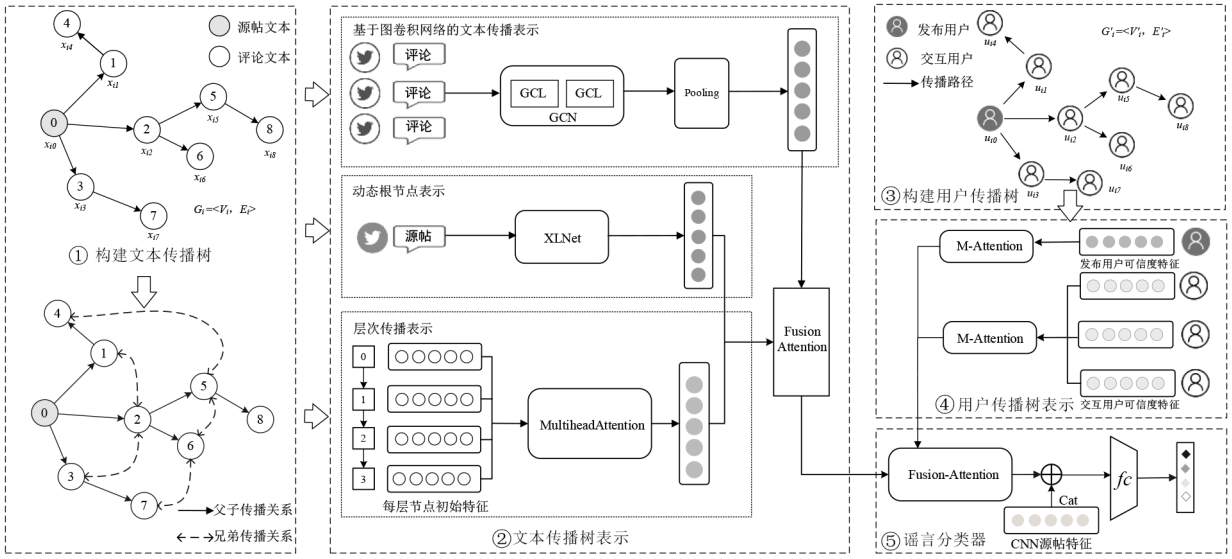

图1　基于传播树的多特征谣言检测方法框架

在用户传播树中，受“回音室效应”的启发构建了用户可信度，探究用户可信度特征和谣言之间的关系. 最后，基于注意力机制动态为每种特征分配权值，使用融合后的特征进行谣言检测.

4.1　基于文本的传播树表示

基于文本的传播树表示包括：构建文本传播树、基于图卷积网络的文本传播表示、动态根节点表示、层次传播表示和文本特征融合5个部分.

4.1.1　构建文本传播树

根据源帖发布后产生的转发评论关系，构建以源帖为起点的传播结构树. 具体的，对每条谣言样本 c_i 生成传播树 $\boldsymbol{G}_i=<\boldsymbol{V}_i,\boldsymbol{E}_i>$?（如图1模块①所示），其中，$\boldsymbol{V}_i=\{x_{i0},x_{i1},\cdots,x_{im_i}\}$ 表示树中的节点集合，节点 x_{i0} 表示源帖的文本信息，节点 $x_{i1},\cdots,x_{im_i}$ 表示评论产生的文本信息. $E_i=\{e_{st}^i|s,t=0,\cdots,m_i\}$，表示传播树的交互关系，使用 $\boldsymbol{A}^i$ 表示传播树中节点的依赖关系：

$$A_{ts}^i=\begin{cases}1, & \text{如果 } e_{st}^i\in E_i\\ 0, & \text{其他}\end{cases} \tag{2}$$

4.1.2　基于图卷积网络的文本传播表示

图卷积网络能够不断聚合邻接节点特征，因此本文使用图卷积网络构建文本传播表示，探究传播树中的相邻节点的父子传播特征. 但在使用图卷积网络更新节点信息的过程中，会存在过拟合的问题. 为了缓解该问题，采用Rong等人提出的DropEdge[23]方法，以一定的概率随机删除传播树中的边，增加输入数据的随机性和多样性. 经过DropEdge后，表示传播树依赖关系的邻接矩阵变为

$$\boldsymbol{A}=\boldsymbol{A}-\boldsymbol{A}^{\text{drop}} \tag{3}$$

其中，$\boldsymbol{A}^{\text{drop}}$ 是由原邻接矩阵随机选择（$\eta\times m_i$）条边生成的新矩阵，η 为删除概率，m_i 为传播树的总边数.

对于数据集中所有节点产生的文本信息，根据词频提取出现频率较高的前5 000个词，使用TF-IDF构造每个节点的初始特征 $\boldsymbol{X}_i$，$\boldsymbol{X}_i=[\boldsymbol{X}_{i0},\boldsymbol{X}_{i1},\cdots,\boldsymbol{X}_{im_i}]$. 其中，$\boldsymbol{X}_{i0}\in\mathbb{R}^{d_0}$ 表示源帖 x_{i0} 的文本特征向量，$\boldsymbol{X}_{ij}\in\mathbb{R}^{d_0}$ 表示源帖发布后第 j 个评论的文本特征向量.

根据构建好的邻接矩阵 $\boldsymbol{A}$ 和节点初始特征 $\boldsymbol{X}$，使用两层图卷积网络挖掘全局传播特征，各隐藏层特征由下列公式求出：

$$\boldsymbol{H}_1=\sigma(\tilde{\boldsymbol{A}}\boldsymbol{X}\boldsymbol{W}_0) \tag{4}$$

$$\boldsymbol{H}_2=\sigma(\tilde{\boldsymbol{A}}\boldsymbol{H}_1\boldsymbol{W}_1) \tag{5}$$

$$\tilde{\boldsymbol{A}}=\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\boldsymbol{A}\tilde{\boldsymbol{D}}^{-\frac{1}{2}} \tag{6}$$

$$\tilde{\boldsymbol{D}}_{ii}=\sum_j \boldsymbol{A}_{ij} \tag{7}$$

其中，$\tilde{\boldsymbol{A}}$是标准化邻接矩阵，$\tilde{\boldsymbol{D}}_{ii}$表示第$i$个节点的度，$\boldsymbol{W}_{k-1}\in\mathbb{R}^{v_{k-1}\times v_k}$是可训练参数，$\sigma(\cdot)$是激活函数. 然后，将得到的隐层特征使用平均池化操作进行聚合，得到基于图卷积网络的文本传播表示：

$$\tilde{\boldsymbol{X}}=\mathrm{MEAN}(\sigma(\tilde{\boldsymbol{A}}\boldsymbol{H}_1\boldsymbol{W}_1)) \tag{8}$$

4.1.3 动态根节点表示

使用图卷积网络对文本传播表示建模时，采用TF-IDF对文本内容进行编码，得到静态文本特征. 这种静态编码的方法有向量维度较高，不能根据上下文准确表示文本语义的缺点. 此外，在基于传播树构建文本特征时，并没有区分节点类型，而是将树中所有节点视为同等重要，但相比其他节点而言，根节点中因存在丰富的源帖信息而更为重要.

因此，为了弥补TF-IDF无法获取上下文语义的缺点，充分探究源帖中的语义特征，使用基于自回归语言和双向编码的XLNet[7]模型对源帖进行动态编码，生成源帖的动态词嵌入表示$\boldsymbol{D}_i$

$$\boldsymbol{D}_i=\mathrm{XLNet}(\boldsymbol{x}_{i0}) \tag{9}$$

4.1.4 层次传播表示

基于图卷积网络的文本传播表示挖掘了邻接节点间的父子传播关系，通过不断聚合邻接节点信息更新当前节点特征. 如图2左边所示，实线表示父子传播关系，如节点3通过聚合其父节点0、子节点5和子节点6更新自己的节点特征.

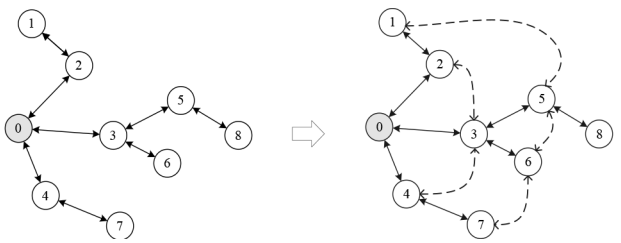

图2 父子传播关系和兄弟传播关系示意图

但是，这种方法却忽略了层内节点间的兄弟传播关系. 一些评论有时能够煽动他人改变对源帖的看法，为了考虑不同帖子之间的相互影响关系，探究了层次节点即同层次帖子之间的隐式依赖关系. 因此，本文在捕获父子传播关系的基础上，通过构建传播树的层次表示，捕获层内节点的兄弟传播关系. 如图2右边所示，虚线表示节点间的兄弟传播关系，节点3除了可以获取父子节点0、5、6的特征外，还可以捕获兄弟节点2和4的特征.

基于以上分析，将每条谣言样本建立的文本传播树转换为层次传播结构，层次节点的初始特征仍由TF-IDF方法构建. 然后，使用多头注意力模型挖掘每层节点更深层次的依赖关系：

$$\mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{softmax}\left(\frac{\boldsymbol{QK}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{10}$$

$$\mathbf{head}_j=\mathrm{Attention}(\boldsymbol{QW}_j^Q,\boldsymbol{KW}_j^K,\boldsymbol{VW}_j^V) \tag{11}$$

$$\mathbf{MultiAttention}_i=\mathrm{Concat}(\mathbf{head}_1,\mathbf{head}_2,\cdots,\mathbf{head}_h)\boldsymbol{W}^o \tag{12}$$

其中，$\boldsymbol{Q}\in\mathbb{R}^{\mathrm{ln}\times d_0}$是层次特征矩阵，$\boldsymbol{K}\in\mathbb{R}^{\mathrm{ln}\times d_0}$，$\boldsymbol{V}\in\mathbb{R}^{\mathrm{ln}\times d_0}$通过随机初始化得到$\boldsymbol{W}_j^Q\in\mathbb{R}^{\mathrm{ln}\times d_0/h}$、$\boldsymbol{W}_j^K\in\mathbb{R}^{\mathrm{ln}\times d_0/h}$、$\boldsymbol{W}_j^V\in\mathbb{R}^{\mathrm{ln}\times d_0/h}$、$\boldsymbol{W}^o\in\mathbb{R}^{h\mathrm{ln}\times d_0}$，是转换矩阵.

在构建层次传播表示时，传播树中每层的节点个数往往并不相同，认为拥有更多节点的层，能够提供更丰富的层内依赖关系. 因此按照每层节点数所占总节点数的比值，为层特征赋予不同的权重，加权融合后得到最终的层次传播表示：

$$\boldsymbol{L}_i=\sum_{j=1}^{\mathrm{ln}}\mathbf{MultiAttention}_{ij}\times\frac{\mathrm{len}_j}{\mathrm{len}_i} \tag{13}$$

其中，ln是当前传播树的总层数，len_i是当前传播树的总节点数，len_j是第j层的节点个数，完整的计算流程如算法1所示.

算法1 获取层次传播表示

输入：

1. 传播树的节点特征矩阵$\boldsymbol{X}\in\mathbb{R}^{n\times d_0}$;
2. 传播树的层次关系矩阵$\mathbf{layer}\in\mathbb{R}^{\mathrm{ln}\times d_0}$;
3. 权重矩阵$\boldsymbol{W}^o\in\mathbb{R}^{h\mathrm{ln}\times d_0}$和多头注意力模块的多头数$h\in[1,2,\cdots,H]$.

输出：层次传播表示$\boldsymbol{L}$

①for $i\in\{1,2,\cdots,n\}$ do

② for $j\in\{1,2,\cdots,\mathrm{ln}\}$ do

$\boldsymbol{N}_{ij}=\mathrm{Lookup}(\boldsymbol{X}_i,\mathbf{layer}_{ij})$;/* Lookup(.)函数将传播树转换为层次传播结构，并根据节点特征矩阵$\boldsymbol{X}$和层次关系矩阵$\mathbf{layer}$提取每层节点的层初始特征$\boldsymbol{N}_{ij}$;*/

③ end for

④ $\boldsymbol{N}_i=[\boldsymbol{N}_{i1},\boldsymbol{N}_{i2},\cdots,\boldsymbol{N}_{i\mathrm{ln}}]$;/*第$i$条谣言样本的层次特征*/

⑤ for $h\in[1,2,\cdots,H]$

$\mathrm{head}_h=\mathrm{Attention}(N_iW_h^N,KW_h^K,VW_h^V)$;/*使用多头注意力机制挖掘深层特征*/

⑥ end for

⑦ $\mathbf{MultiAttention}_i=\mathrm{Concat}(\mathbf{head}_1,\mathbf{head}_2,\cdots,\mathbf{head}_h)\boldsymbol{W}^o$;

⑧ $\boldsymbol{L}_i=\mathrm{fuse}(\mathbf{MultiAttention}_i)$;/*使用式(13)加权融合，得到最终层次传播特征*/

⑨ end for

⑩ $\boldsymbol{L}=[\boldsymbol{L}_1,\boldsymbol{L}_2,\cdots,\boldsymbol{L}_i]$;

return $\boldsymbol{L}$

4.1.5 文本特征融合

根据构建好的文本传播树，能够得到谣言样本c_i基于图卷积网络的文本传播表示$\tilde{\boldsymbol{X}}_i\in\mathbb{R}^d$、层次传播表示

$\boldsymbol{L}_i \in \mathbb{R}^{\ln \times d}$和动态根节点特征$\boldsymbol{D}_i \in \mathbb{R}^{l \times d}$. 为了更好地衡量不同文本特征对模型的重要程度,本文使用一种基于注意力机制的特征融合方法——Fusion-Attention融合3种文本特征. 该模块首先将层次传播表示$\boldsymbol{L}_i$和动态根节点特征$\boldsymbol{D}_i$按以下方法进行融合:

$$\boldsymbol{L}_i' = \sum_{j=1}^{ln} \boldsymbol{\varepsilon}_j \boldsymbol{L}_{ij} \tag{14}$$

$$\boldsymbol{\varepsilon} = \mathrm{softmax}(\boldsymbol{D}_i \boldsymbol{L}_i^{\mathrm{T}}) \tag{15}$$

然后,将基于图卷积网络的文本传播表示$\tilde{\boldsymbol{X}}_i$和上述融合好的特征$\boldsymbol{L}_i'$进行融合,得到最终的文本传播树表示$\tilde{\boldsymbol{f}}_i$:

$$\tilde{\boldsymbol{f}}_i = [\tilde{\boldsymbol{X}}_i, \boldsymbol{L}_i']\boldsymbol{W}_t + b_t \tag{16}$$

其中,$\boldsymbol{W}_t \in \mathbb{R}^{4d \times d}$是变换矩阵,$b_t \in \mathbb{R}^d$是偏置项.

4.2　基于用户的传播树表示

基于用户的传播树表示主要由构建用户传播树和用户可信度表示组成.

4.2.1　构建用户传播树

在日常生活中,人们更愿意相信人民日报、央视新闻、地方政府等权威用户发布的消息,这是因为具有良好信誉的用户一般不会发布不实信息. 同时,Monti等人[24]研究了用户在社交媒体上的互动情况,发现用户更喜欢与具有相同可信度的用户进行互动,“回音室效应”也证明同种类型的用户更容易聚在一起. 受上述启发,本文根据Ma[25]所提供的数据集,将用户分为不可信、不确定和可信三类. 首先对帖子进行分类,将帖子标签为假谣言和非谣言的标记为真,标签为真谣言和不确定的谣言的标记为假. 然后,遍历数据集中的用户,统计每个用户交互的帖子类型. 最后,将只与真帖子交互的用户标记为可信用户;将只与假帖子交互的用户标记为不可信用户;将同时与真帖子和假帖子交互的用户标记为不确定用户.

根据用户与谣言的交互关系,对每条谣言样本c_i构建用户传播树$\boldsymbol{G}_i' = <V_i', E_i'>$(如图1模块③所示). 其中$\boldsymbol{V}_i' = \{u_{i0}, u_{i1}, \ldots, u_{is_i}\}$表示传播树中的用户集合,$u_{i0}$表示发布源帖的用户;$u_{i1}, \ldots, u_{is_i}$表示按照顺序与源帖进行转发评论的用户;节点集合$\boldsymbol{E}_i' = \{e_{st}^i | s, t = 0, \ldots, s_i\}$表示从源帖到相应转发或评论的一组定向边,记录了树的传播方向.

4.2.2　用户可信度表示

本文将用户划分为交互用户和发布用户两种类型. 交互用户指对源帖进行转发评论的用户,发布用户指发布源帖的用户.

对于参与转发评论的交互用户,根据传播树构建用户传播序列,使用传播序列中每个用户的可信度值构建初始可信度特征$\boldsymbol{I}_i \in \mathbb{R}^{k \times \mathrm{ud}}$,再将$\boldsymbol{I}_i$转换为嵌入向量$\boldsymbol{IE}_i$,使用M-Attention模块学习用户可信度的深层次表示. 该模块叠加了M个注意力模块,并将每次得到的结果进行拼接作为最终的特征,由此交互用户可信度特征$\tilde{\boldsymbol{U}}_i$的计算方法如下:

$$\mathbf{att}_i = \mathrm{Attention}(\mathbf{IE}_i, \boldsymbol{I}_i, \boldsymbol{I}_i) \tag{17}$$

$$\boldsymbol{U}_j = \mathrm{ELU}[(\mathbf{att}_1, \mathbf{att}_2, \cdots, \mathbf{att}_M)\boldsymbol{W}_u] + \mathbf{IE}_i \tag{18}$$

$$\tilde{\boldsymbol{U}}_i = [\boldsymbol{U}_1, \boldsymbol{U}_2, \cdots, \boldsymbol{U}_n] \tag{19}$$

同样对于发布源帖的发布用户,根据用户可信度构建初始特征$\boldsymbol{I}'_i \in \mathbb{R}^{\mathrm{ud}}$,然后将初始特征转换为嵌入向量$\mathbf{IE}'_i$,使用M-Attention模块获取发布者可信度特征$\boldsymbol{P}_i$.

4.3　谣言分类器

根据用户传播树和谣言样本c_i得到交互用户可信度特征$\tilde{\boldsymbol{U}}_i \in \mathbb{R}^{k \times \mathrm{ud}}$、发布者可信度特征$\boldsymbol{P}_i \in \mathbb{R}^{\mathrm{ud}}$和基于文本的传播树表示$\tilde{\boldsymbol{f}}_i \in \mathbb{R}^{l \times \mathrm{ud}}$,使用第4.1.5节中的Fusion-Attention模块对3种特征进行融合. 首先,将交互用户可信度特征$\tilde{\boldsymbol{U}}_i$和基于文本的传播树特征$\tilde{\boldsymbol{f}}_i$进行融合:

$$\boldsymbol{U}_i' = \sum_{k=1}^{K} \boldsymbol{\varepsilon}_k' \tilde{\boldsymbol{U}}_{ik} \tag{20}$$

$$\boldsymbol{\varepsilon}' = \mathrm{softmax}(\tilde{\boldsymbol{f}}_i \boldsymbol{U}_i^{\mathrm{T}}) \tag{21}$$

然后,将融合好的特征$\boldsymbol{U}_i'$和发布者可信度特征$\boldsymbol{P}_i$用以下公式进行融合,得到最终的传播树特征$\tilde{\boldsymbol{g}}$:

$$\tilde{\boldsymbol{g}}_i = [\boldsymbol{P}_i, \boldsymbol{U}_i']\boldsymbol{W}_m + b_m \tag{22}$$

其中,$\boldsymbol{W}_m \in \mathbb{R}^{4\mathrm{ud} \times \mathrm{ud}}$是变换矩阵,$b_m \in \mathbb{R}^{\mathrm{ud}}$是一个偏置项. 为了更充分挖掘谣言的源帖文本特征,我们使用卷积神经网络对每条谣言样本c_i提取了源帖特征$\boldsymbol{T}_i$. 将$\boldsymbol{T}$和$\tilde{\boldsymbol{g}}$进行拼接作为最终特征,应用全连接层将最终特征投影到目标类别:

$$\hat{y} = \mathrm{softmax}([\boldsymbol{T}, \tilde{\boldsymbol{g}}]\boldsymbol{W} + b) \tag{23}$$

其中,$\boldsymbol{W} \in \mathbb{R}^{4\mathrm{ud} \times \mathrm{ud}}$是变换矩阵,uc是用户可信度的类别数,$b \in \mathbb{R}$是一个偏置项. 最后,使用交叉熵损失函数优化目标函数:

$$\mathcal{L} = -\sum_{j=1}^{|N|} y_j \log \hat{y} + \frac{\lambda}{2} \| \theta \|_2^2 \tag{24}$$

其中,y_j是新闻j的真实标签,λ是惩罚因子,θ表示模型需要训练的参数.

5　实验

本节通过实验来验证模型是否能够提高谣言的检测性能,模型的每个模块对于整体检测任务是否有一定的影响,以及与现有的模型相比该模型能否在谣言发布早期达到良好的识别效果.

5.1　数据集

本文在Twitter15、Twitter16[25]和Weibo[8]3个数据集上评估本文提出的模型. Twitter15和Twitter16数据

集包含四类标签NR、FR、TR、UR(非谣言,假谣言,真谣言,未经证实的谣言),Weibo数据集包含两类标签TR和FR(真谣言和假谣言). 每个事件的标签都是根据一些谣言验证网站依据事件真实性进行标注的,各数据集的信息如表1所示.

表1 谣言检测数据集统计信息

	Twitter15	Twitter16	Weibo
样本数量	1 490	818	4 664
帖子总数量	331 612	204 820	3 805 656
用户总数量	276 663	173 487	2 746 818
真谣言数	374	205	2 351
假谣言数	370	205	2 313
未验证谣言数	374	203	0
非谣言数	372	205	0
最多帖子数	1 768	2 765	59 318
最少帖子数	55	81	10
平均帖子数	223	251	816
最多交互用户数	2 604	2 749	55 356
最少交互用户数	38	73	6
平均交互用户	185	212	588

5.2 对比模型

实验将本文提出的MPT模型与其他一些谣言检测模型进行对比,对比模型包括:DTC[2]:使用基于人工设计的文本特征,通过决策树分类器进行谣言分类. SVM-RBF[26]:一种带有径向基函数的支持向量机分类器,以从文本中提取的特征作为输入特征. SVM-TS[27]:利用时间序列构建特征,通过线性支持向量机对谣言进行分类. DTR[28]:使用聚类的方法构建决策树检测器,检测带有质疑的帖子. RFC[29]:通过检验谣言传播的时间,结构和语言来确定谣言的特性. RvNN[30]:构造了一种基于RNN(Recurrent Neural Net)的自上而下和自下而上的双向传播树模型. PPC[31]:将谣言的传播信息与评论信息看作一个时间序列,对该序列使用RNN和CNN(Convolutional Neural Network)建模. Bi-GCN[21]:构建基于图卷积网络的双向传播结构,融合双向传播特征对谣言进行检测. EBGCN[22]:改进了Bi-GCN模型中传播结构的不确定性,替换传播树中的固定边权值. GLAN[32]:提出一种新的全局局部注意力网络,构建了全局和局部语义信息. HGAT[33]:根据谣言传播的过程构建异构图,使用图注意力网络挖掘传播结构特征. Rumor2vec[34]:基于CNN在联合图中捕获文本内容特征和传播结构特征. UMLARD[35]:提出了一个构建用户多视图特征的模型,挖掘用户与谣言之间的潜在关系.

5.3 参数设置与评估方法

MPT模型的参数情况设置如下:Twitter数据集的batchsize为16,Weibo数据集的batchsize为32;为了防止过拟合,使用了dropout机制,初始比率设为0.3,学习率设置为0.000 1. 在文本传播表示模块,图卷积网络删除边的比率η=0.2,节点特征维度d_0=5 000. 获取层次节点特征时,多头注意力的个数h=5,动态根节点特征和层次节点特征向量维度d=128. 在用户传播表示模块,用户特征向量维度ud=100,对于Twitter15、Twitter16和Weibo数据集设置M-Attention模块的M分别为10、8和7.

本文使用准确率和F_1分数作为评价指标,分别评估了每个类别的准确率和F_1分数. 此外,本文将数据集随机划分成五部分,进行5折交叉验证确保模型的泛化能力,避免模型出现过拟合.

5.4 实验结果和分析

通过实验得到模型的总体准确率和各类别的F_1分数验证所提出模型的检测性能,加粗的数据是所有模型中最优的实验结果. 由表2~表4可知,基于人工设计特征的传统检测方法(DTR(Decision Tree Regression)、DTC(Decision Tree Classifier)、RFC(Random Forest Classifier)、SVM-RBF(Support Vector Machine with Radial Basis Function kernel)、SVM-TS(Support Vector Machine with Time Series))性能比较差,基于深度学习的方法(RvNN(Rumor Detection Approach based on Recursive Neural Networks)、PPC(Propagation Path Classification)、Bi-GCN(Bi-directional Graph Convolutional Networks)、EBGCN(Edge-enhanced Bayesian Graph Convolutional Networks)、GLAN(Global-local Attention Network)、HGAT(Heterogeneous Graph Attention Networks)、Rumor2vec(A rumor detection framework with joint text and propagation structure representation learning)、UMLARD(User-aspect Multi-view Learning with Attention for Ru-

表2 Twitter15数据集上的实验结果

模型	ACC	F_1			
		NR	FR	TR	UR
DTC	0.454	0.733	0.355	0.317	0.415
SVM-RBF	0.318	0.225	0.282	0.455	0.218
SVM-TS	0.544	0.796	0.472	0.404	0.483
DTR	0.409	0.501	0.311	0.364	0.473
RFC	0.565	0.810	0.422	0.401	0.543
RvNN	0.723	0.682	0.758	0.821	0.654
PPC	0.842	0.811	0.875	0.818	0.790
Bi-GCN	0.845	0.804	0.853	0.899	0.813
EBGCN	0.848	0.809	0.850	**0.908**	0.814
GLAN	0.850	0.842	0.857	0.881	0.818
HGAT	0.869	0.862	0.889	0.886	0.838
Rumor2vec	0.796	**0.883**	0.746	0.836	0.723
UMLARD	0.857	0.835	0.786	0.887	0.837
MPT	**0.893**	0.880	**0.908**	0.863	**0.919**

mor Detection)、MPT)明显优于传统的检测方法.

表3　Twitter16数据集上的实验结果

模型	ACC	F_1			
		NR	FR	TR	UR
DTC	0.465	0.643	0.393	0.419	0.403
SVM-RBF	0.321	0.423	0.085	0.419	0.037
SVM-TS	0.574	0.755	0.420	0.571	0.526
DTR	0.414	0.394	0.273	0.630	0.344
RFC	0.585	0.752	0.415	0.547	0.563
RvNN	0.737	0.662	0.743	0.835	0.708
PPC	0.863	0.820	0.898	0.843	0.837
Bi-GCN	0.875	0.809	0.880	0.941	0.866
EBGCN	0.847	0.794	0.847	0.896	0.826
GLAN	0.858	0.812	0.857	0.903	0.859
HGAT	0.880	0.854	0.860	0.926	0.879
Rumor2vec	0.852	0.857	0.769	0.927	0.850
UMLARD	0.901	0.822	**0.965**	**0.960**	0.855
MPT	**0.917**	**0.892**	0.908	0.925	**0.946**

表4　Weibo数据集上的实验结果

模型	ACC	TR			FR		
		Prec	Rec	F_1	Prec	Rec	F_1
DTC	0.831	0.815	0.847	0.830	0.847	0.815	0.831
SVM-RBF	0.818	0.815	0.824	0.819	0.822	0.812	0.817
SVM-TS	0.857	0.878	0.830	0.857	0.839	0.885	0.861
DTR	0.732	0.726	0.749	0.737	0.738	0.715	0.726
RFC	0.849	0.947	0.739	0.830	0.786	0.959	0.864
RvNN	0.908	0.904	0.918	0.911	0.912	0.897	0.905
PPC	0.921	**0.949**	0.889	0.918	0.896	**0.962**	0.923
Bi-GCN	0.930	0.928	0.939	0.929	0.940	0.930	0.931
GLAN	0.932	0.925	0.943	0.934	0.924	0.932	0.928
Rumor2vec	0.951	0.945	0.956	0.950	0.958	0.948	0.953
UMLARD	0.928	0.942	0.965	0.924	0.913	0.899	0.901
MPT	**0.964**	0.942	**0.991**	**0.966**	**0.990**	0.937	**0.962**

相比最优的传播树模型Bi-GCN,本文提出的MPT性能更好.Bi-GCN仅仅关注到了传播树中的父子传播关系,而忽略兄弟节点间的信息传播,同时过于依赖文本信息.而MPT则考虑了除文本之外的用户特征,同时探究了层次节点之间的传播关系,进而获得了更好的检测效果.

对于检测效果较好的模型HGAT而言,HGAT更关注源帖和词语之间的联系,而忽略了评论帖子中的信息.本文则挖掘了评论帖子之间的多种传播特征,有效弥补了HGAT仅关注源帖的局限性.而对于构建用户特征的UMLARD模型而言,本文的模型既充分地挖掘了文本特征,也探究了用户可信度和谣言之间的关系.

本文提出的MPT模型在Twitter15数据集上获得了89.3%的准确率,分别优于HGAT模型和UMLARD模型2.4%和3.6%,在两个类别的F_1分数都高于其他对比模型;在Twitter16数据集上,MPT的准确率达到91.7%,分别优于HGAT模型和UMLARD模型3.7%和1.6%,在两个类别的F_1分数都高于其他对比模型;在Weibo数据集上,MPT的准确率为96.4%,分别优于GLAN模型和UMLARD模型3.2%和3.6%.综合得出MPT在Twitter15、Twitter16和Weibo数据集上均优于基线方法,具有较好的检测效果.

为了评估本文模型在计算时间上的消耗,选取近年效果较好的方法(Bi-GCN、EBGCN、GLAN、HGAT)作为对比模型.记录模型在每个测试集预测一条样本标签的平均推理时间(以ms为单位),得到的实验结果如表5所示.

观察表5可得,GLAN所消耗的计算时间是最少的,这是因为GLAN模型没有使用图网络,只是通过注意力机制获取文本特征.虽然GLAN模型推理速度快,但是模型的检测效果并不是很好.HGAT构建了异构图,虽然能够取得很好的检测效果,但是复杂多样的边关系也导致计算时间大大增加.

表5　模型计算时间的对比实验结果

数据集	模型	推理时间/ms	ACC
Twitter15	Bi-GCN	21.170	0.845
	EBGCN	24.145	0.848
	GLAN	**17.729**	0.850
	HGAT	24.795	0.869
	MPT	23.392	**0.893**
Twitter16	Bi-GCN	21.748	0.875
	EBGCN	34.085	0.847
	GLAN	**19.267**	0.858
	HGAT	27.130	0.880
	MPT	24.753	**0.917**
Weibo	Bi-GCN	213.542	0.930
	GLAN	**154.004**	0.932
	MPT	237.313	**0.964**

MPT模型与耗时和检测效果都相对较好的Bi-GCN相比,在Twitter15、Twitter16和Weibo数据集上的推理时间虽然分别增加了2.222 ms、3.005 ms和23.771 ms,但是准确率分别提高了4.8%、4.2%和3.4%.与检测效果较优的HGAT模型相比,在Twitter15和Twitter16数据集上的推理时间减少了1.403 ms和2.377 ms,准确率却增加了2.4%和3.7%,无论计算耗时还是检测准确率都优于HGAT.与推理速度最快的GLAN模型相比,虽然耗时增加了,但在Twitter15、Twitter16和Weibo数据集上的准确率也分别优于GLAN模型4.3%、5.9%和3.2%.综合得出,MPT在计算时间的消耗是值得的.

5.5 消融实验

为了探究不同模块对模型的影响程度，对模型进行了一系列的消融实验．主要包括以下4部分．

（1）w/o TGCN．移除基于文本的传播树表示模块，其余部分不变．用于验证在传播过程中文本传播树表示是否具有一定的有效性．

（2）w/o UG．移除基于用户的传播树表示模块，用于验证用户传播特征是否具有一定的有效性．

（3）w/o Text．移除与模型最后进行拼接的文本特征，研究源帖特征对模型的重要程度．

（4）w/o Fusion-Attention．移除模型中的Fusion-Attention模块，用于验证基于注意力机制的特征融合方法能否获取更准确的特征表示．

由图3消融实验结果可以看出，移除文本传播树模块会影响模型的检测效果，在Twitter15、Twitter16和Weibo数据集上的结果分别下降了1.8%、1.6%和0.9%．移除用户传播树模块后，在Twitter15、Twitter16和Weibo数据集上的结果分别下降了4.9%、4%和3.9%．去除CNN文本表示模块，模型在Twitter15、Twitter16和Weibo数据集上的结果分别下降了2.8%、3.6%和2.8%．去除特征融合模块后，结果表明在Twitter15、Twitter16和Weibo数据集上的检测准确率分别下降了0.8%、1.1%和1.9%．

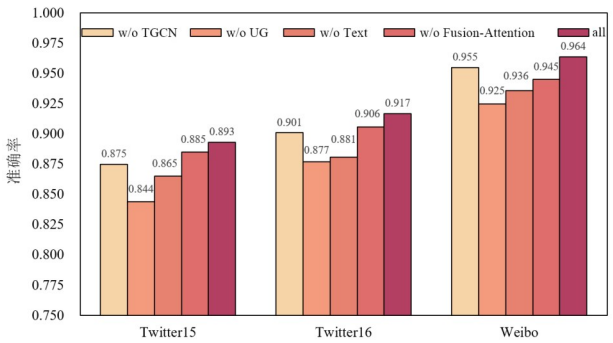

图3 模型在3个数据集上的消融实验结果

以上实验结果证明：用户传播表示相比文本传播表示能够提供更丰富的传播特征，两种传播表示对于整体模型的效果有重要的影响．同时，源帖中存在丰富有效的信息，能够体现谣言的潜在特征．此外，基于注意力机制的融合方法能够一定程度上获取更全面、更有效、更丰富的特征向量．

5.6 早期谣言检测任务

在谣言检测任务中，如何尽快地检测出谣言并进行干预是至关重要的．为了检测模型在谣言出现后不同时间段的检测效果，本文在Twitter15和Twitter16数据集上设计了早期谣言检测实验，通过设置一系列检测截止时间（4 h、8 h、12 h、16 h、20 h），评估不同时间节点产生的数据在模型中的检测效果，得到如图4和图5所示的实验结果．

当谣言样本的发布时间为0 h时，源帖刚发布还未产生转发评论信息，此时MPT在Twitter15和Twitter16数据集上分别取得了86.3%和88%的检测准确率，虽然此时只有源帖相关信息，但是对源帖文本和用户信息的充分建模，使得本文的模型有更好的检测性能．

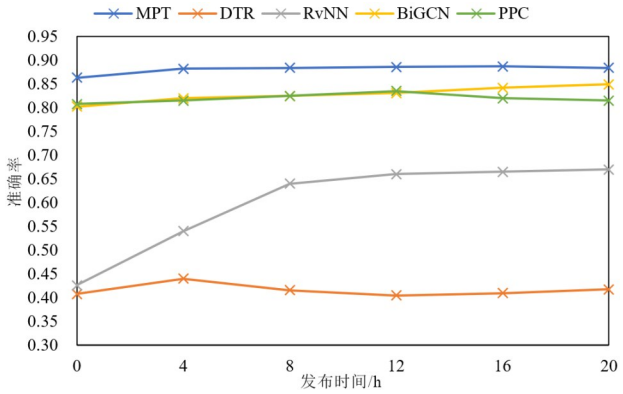

图4 twitter15数据集早期检测实验结果

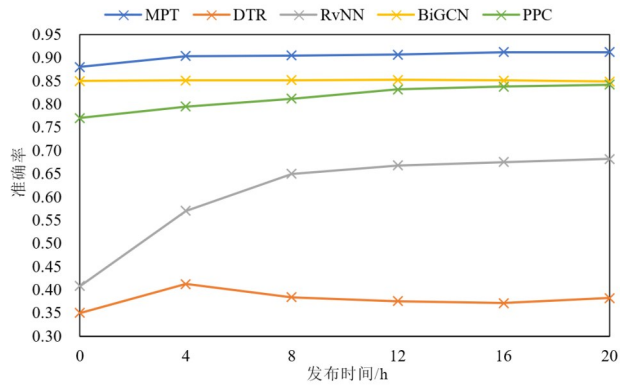

图5 twitter16数据集早期检测实验结果

在发布时间为0~4 h阶段，可以看出本文的模型和Bi-GCN都有较好的效果，这表明图卷积获取的传播结构特征有助于在早期发现谣言．在4~8h阶段，可以看出PPC、Bi-GCN、MPT的检测性能远远高于DTR和RvNN，这是因为DTR和RvNN只是基于文本内容构建特征，而PPC、Bi-GCN和MPT在构建文本特征的同时，还关注到了传播过程中的信息，因此模型的检测性能更好．在发布时间为8~20 h阶段，随着时间的不断推移，更多的转发评论关系使得图结构更加完整，MPT模型的性能也呈上升趋势．总体来看，MPT模型的早期检测性能优于其他对比模型．

6 总结

本文研究了基于传播结构的谣言检测任务，提出一种基于传播树的多特征谣言检测模型．该方法在对

文本传播特征建模时,使用图卷积网络聚合邻接节点信息捕获父子传播关系,通过多头注意力模块挖掘兄弟传播关系;同时采用动态嵌入的方式对源帖进行编码,增强根节点信息的影响力,丰富文本传播表示. 在对用户传播特征建模时,根据用户传播树中用户的历史交互数据和传播路径,构建用户可信度序列,叠加多个注意力模块提取用户传播特征. 最后,使用一种基于注意力机制的特征融合方法,融合文本传播特征和用户传播特征进行谣言检测. 在真实数据集上的实验结果表明,本文提出的方法能够挖掘谣言传播过程中的深层次规律,有效检测社交媒体中的谣言信息. 在未来的研究中将结合动态建模方法(如Transformer、Bert),构建动态节点表示,更好地表示文本中的语义信息,进一步挖掘动态传播树中的潜在特征.

参考文献

[1] 郭家炜, 朱云霞. 重大突发事件中谣言传播与舆论引导研究——以新冠肺炎疫情期间部分典型谣言为例[J]. 科技传播, 2021, 13(17): 6.
GUO J W, ZHU Y X. A study on the spread of rumors and the guidance of public opinion in major emergencies—Taking some typical rumors during the COVID-19 epidemic as an example[J]. Public Communication of Science & Technology, 2021, 13(17): 10-15. (in Chinese)

[2] CASTILLO C, MENDOZA M, POBLETE B. Information credibility on twitter[C]//Proceedings of the 20th International Conference on World Wide Web. New York: ACM, 2011: 675-684.

[3] SHU K, SLIVA A, WANG S, et al. Fake news detection on social media: A data mining perspective[J]. ACM SIGKDD Explorations Newsletter, 2017, 19(1): 22-36.

[4] HORNE B, ADALI S. This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news[J]. Proceedings of the International AAAI Conference on Web and Social Media, 2017, 11(1): 759-766.

[5] POTTHAST M, KIESEL J, REINARTZ K, et al. A stylometric inquiry into hyperpartisan and fake news[EB/OL]. (2017)[2022]. https://arxiv.org/abs/1702.05638.

[6] GIUDICE K D. Crowdsourcing credibility: The impact of audience feedback on Web page credibility[J]. Proceedings of the American Society for Information Science and Technology, 2010, 47(1): 1-9.

[7] YANG Z, DAI Z, YANG Y, et al. XLNet: Generalized autoregressive pretraining for language understanding[EB/OL]. (2019)[2022]. https://arxiv.org/abs/1906.08237.

[8] MA J, GAO W, MITRA P, et al. Detecting rumors from microblogs with recurrent neural networks[C]//IJCAI International Joint Conference on Artificial Intelligence. New York: ACM, 2016: 3818-3824.

[9] MA J, GAO W, WONG K F. Detect rumor and stance jointly by neural multi-task learning[C]//Companion Proceedings of the Web Conference 2018. New York: ACM, 2018: 585-593.

[10] CHEN T, LI X, YIN H, 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. Springer: Cham, 2018: 40-52.

[11] 夏鑫林, 许亮. 基于注意力机制的谣言检测算法研究[J]. 现代计算机, 2020, 37(8): 47-51.
XIA X L, XU L. Research on rumor detection algorithm based on attention mechanism modern computer[J]. Modern Computer, 2020, 37(8): 47-51. (in Chinese)

[12] DOU Y, SHU K, XIA C, et al. User preference-aware fake news detection[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2021: 2051-2055.

[13] LU Y J, LI C T. GCAN: Graph-aware co-attention networks for explainable fake news detection on social media [EB/OL]. (2020)[2022]. https://arxiv.org/abs/2004.11648.

[14] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[EB/OL]. (2016)[2022]. https://arxiv.org/abs/1609.02907.

[15] YUAN C, MA Q, ZHOU W, et al. Early detection of fake news by utilizing the credibility of news, publishers, and users based on weakly supervised learning[EB/OL]. (2020)[2022]. https://arxiv.org/abs/2012.04233.

[16] KWON S, CHA M, JUNG K, et al. Prominent features of rumor propagation in online social media[C]//2013 IEEE 13th International Conference on Data Mining. Piscataway: IEEE, 2013: 1103-1108.

[17] WU K, YANG S, ZHU K Q. False rumors detection on sina weibo by propagation structures[C]//2015 IEEE 31st International Conference on Data Engineering. Piscataway: IEEE, 2015: 651-662.

[18] HUANG Q, YU J, WU J, et al. Heterogeneous graph attention networks for early detection of rumors on Twitter [C]//2020 International Joint Conference on Neural Networks (IJCNN). Piscataway: IEEE, 2020: 1-8.

[19] SHU K, WANG S, LIU H. Beyond news contents: The

role of social context for fake news detection[C]//Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. New York: ACM, 2019: 312-320.

[20] KANG Z, CAO Y, SHANG Y, et al. Fake news detection with heterogenous deep graph convolutional network[C]// Pacific-Asia Conference on Knowledge Discovery and Data Mining. Cham: Springer, 2021: 408-420.

[21] BIAN T, XIAO X, XU T Y, et al. Rumor detection on social media with Bi-directional graph convolutional networks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 549-556.

[22] WEI L, HU D, ZHOU W, et al. Towards propagation uncertainty: Edge-enhanced bayesian graph convolutional networks for rumor detection[EB/OL]. (2021)[2022]. https://arxiv.org/abs/2107.11934.

[23] RONG Y, HUANG W B, et al. DropEdge: Towards deep graph convolutional networks on node classification [EB/OL]. (2019)[2022]. https://arxiv.org/abs/1907.10903.

[24] MONTI F, FRASCA F, EYNARD D, et al. Fake news detection on social media using geometric deep learning[EB/OL]. (2019)[2022]. https://arxiv.org/abs/1902.06673.

[25] 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: Association for Computational Linguistics, 2018: 1-12.

[26] YANG F, LIU Y, YU X, et al. Automatic detection of rumor on sina weibo[C]//Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics. New York: ACM, 2012: 1-7.

[27] 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.

[28] 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. Republic and Canton of Geneva: International World Wide Web Conferences Steering Committee, 2015: 1395-1405.

[29] KWON S, CHA M, JUNG K. Rumor detection over varying time windows[J]. PloS One, 2017, 12(1): e0168344.

[30] 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: Association for Computational Linguistics, 2018: 1-10.

[31] LIU Y, WU Y F B. Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks[C]//32nd AAAI Conference on Artificial Intelligence. New York: AAAI Press, 2018: 354-361.

[32] YUAN C, MA Q, ZHOU W, et al. Jointly embedding the local and global relations of heterogeneous graph for rumor detection[C]//2019 IEEE International Conference on Data Mining (ICDM). Piscataway: IEEE, 2019: 796-805.

[33] HUANG Q, YU J S, WU J, et al. Heterogeneous graph attention networks for early detection of rumors on twitter [EB/OL]. (2020)[2022]. https://arxiv.org/abs/2006.05866.

[34] 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.

[35] 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.

作者简介

张鑫昕 女，1998年出生，甘肃白银人．现为宁波大学信息科学与工程学院硕士研究生．主要研究方向为深度学习、自然语言处理、谣言检测等．

E-mail: 2011082349@nbu.edu.cn

潘善亮 男，1970年出生，浙江台州人．现为宁波大学教授、硕士生导师．主要研究方向为服务计算、自然语言处理、数据挖掘等．

E-mail: panshanliang@nbu.edu.cn

茅琴娇 女，1983年出生，浙江天台人．博士，现为宁波工程学院网络空间安全学院(计算机学院)讲师．主要研究方向为机器学习，智能信息处理．

E-mail: maoqinjiao@163.com