

文章编号: 1003-0077(2024)02-0142-13

基于多层次不真实性传播结构的社交媒体谣言检测

高 准¹, 但志平^{1,2}, 董方敏^{1,2}, 张岩珂¹, 张洪志¹

(1. 三峡大学 计算机与信息学院, 湖北 宜昌 443002;

2. 三峡大学 水电工程智能视觉监测湖北省重点实验室, 湖北 宜昌 443002)

摘 要: 当前谣言检测工作主要研究谣言传播的方向特性, 而忽视了谣言传播的全局结构特性, 导致不能充分挖掘谣言潜在的结构特征; 此外, 现有研究忽略了谣言原始传播结构中存在的非真实关系, 从而限制了传播节点特征的学习。为此, 该文提出一种多层次的动态传播注意力网络模型 (Multi-level Dynamic Propagation Attention Networks, MDPAN) 用于检测谣言。该模型通过节点级注意力学习谣言传播图中所有连接边的贡献度, 动态地关注对识别谣言有用的传播关系, 并基于图卷积网络分别提取谣言不同层次的传播特征、扩散特征以及全局结构特征, 最后引入基于注意力机制的池化方法对这些多层次的特征进行有效融合。在公开的 Twitter15、Twitter16 和 Weibo16 数据集上的实验结果表明, 该文所提出的模型对比主流基于传播结构的 EBGCN 模型, 整体准确率分别提高了 2.1%、0.7% 和 1.7%。

关键词: 谣言检测; 传播结构; 节点级注意力; 图卷积网络

中图分类号: TP391

文献标识码: A

Rumor Detection from Social Media via Multi-Level Unreliable Propagation Structures

GAO Zhun¹, DAN Zhiping^{1,2}, DONG Fangmin^{1,2}, ZHANG Yanke¹, ZHANG Hongzhi¹

(1. College of Computer and Information Technology, China Three

Gorges University, Yichang, Hubei 443002, China;

2. Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric

Engineering, China Three Gorges University, Yichang, Hubei 443002, China)

Abstract: Current rumor detection research focuses on studying the directional characteristics of rumor propagation. To exploit the potential structural features of rumors, this paper proposes a multi-level dynamic propagation attention networks (MDPAN) to detect rumors. This method learns the contributions of all connecting edges in the propagation graph through a node-level attention, dynamically focusing on useful propagation relationships for identifying rumors. The graph convolutional networks extracts different levels of propagation features, diffusion features, and global structural features of rumors, which are fused via attention-based pooling methods. Compared with the EBGCN model on Twitter15, Twitter16 and Weibo16 datasets, the proposed method increases the overall accuracy by 2.1%, 0.7% and 1.7%, respectively.

Keywords: rumor detection; propagation structure; node-level attention; graph convolutional networks

0 引言

随着社交媒体日益盛行, 新闻消息借助网络平台可以迅速传播到许多网络用户的身边, 给人们获

取时事消息带来了便利。然而, 一些不良团队或个人为吸引大量用户赚取流量, 故意散播未经证实的消息甚至是虚假消息, 造成大量的谣言在社会面上迅速且广泛地传播, 严重破坏了社会的秩序。因此, 为尽早避免谣言在社交媒体上传播, 研究出一种高

收稿日期: 2023-03-28 定稿日期: 2023-05-22

基金项目: NSFC-新疆联合基金重点项目; 网络谣言检测与舆论引导算法研究(U1703261)

效且准确的谣言检测方法具有重要意义。

近些年,领域内学者开始深入研究社交网络中谣言的传播行为。事件源贴随着时间的推移在社会面上广泛传播,从而构成一定的社会语境,不同类型的谣言和事实在社会关系上存在明显差异^[1-3],因此利用谣言在社交网络中形成的传播结构来检测谣言是可行的。Ma 等人^[4]将谣言完整的传播过程建模成符合自然传播规律的双向树结构,并利用递归神经网络(Recursive Neural Networks,RNN)联合文本内容语义和传播线索来检测谣言。Bian 等人^[5]使用图卷积神经网络^[6](Graph Convolutional Networks,GCN)来建模双向传播图,有效聚合了邻居节点的特征。这些基于谣言传播结构的研究都取得了较为理想的效果,极大地推动了谣言检测的研究进展。

但是,社交网络中的帖子节点在响应其父节点的同时也可能被其子节点所响应,诸如 Ma^[4]和 Bian^[5]等人提出的方法只关注了谣言的传播方向,会导致丢失完整的结构信息,从而使模型不足以学习到相关节点的深层次信息。另外,目前的研究方法普遍在理想化的情况下对谣言的传播过程进行建模^[4-5,7-8],认为社交网络中正常用户和异常用户发布的源贴及其所有响应贴都是完全可信赖的。然而由于社交网络具有极强的可干预性,在许多情况下,对源贴进行响应的用户往往是被恶意操控的,如图 1 所示,异常用户 u1 发布虚假源贴 e1,并操纵正常用户 u2 对其进行支持性的回复 r1,甚至会有意删除正常用户 u3 的真实性回复 r2,这些不真实关系也会存在于事件的多级响应中,如异常用户 u5 被利用对真实评论 r3 进行虚假回复 r4。除此之外,社交网络中还存在大量与源贴内容无关的响应贴,这些不确定性因素混杂在谣言的传播中,致使谣言的传播结构存在不真实或不可靠的情况,给谣言检测带来了巨大障碍。

为了解决上述问题,本文提出了一种多层次动态传播注意力网络模型(Multi-level Dynamic Propagation Attention Networks,MDPAN)用于谣言检测。首先,在建模事件传播和扩散结构的同时,引入可建模全局传播结构的无向(Un-Directional,UD)传播图,以兼顾传播的方向和全局等多层次的结构特征,使得在特征学习阶段能获得丰富的节点信息;在此基础上,针对被恶意操纵的评论和转发以及频繁的无效评论和转发给识别谣言带来的负面影响,引入节点级注意力机制来处理这些不真实的因素,

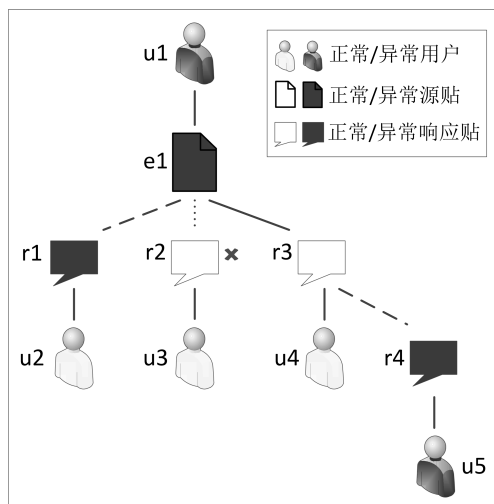


图 1 事件传播的非真实关系场景

通过度量事件中关联帖子之间关系的权重来识别关系的真实性,从而有效地降低事件原始传播结构存在的不稳定性因素给检测结果带来的影响。最后,在特征的融合阶段,使用注意力机制来捕捉不同层次的传播特征和文本内容特征之间的交互,以进一步提高模型的性能。在 Twitter15、Twitter16 以及 Weibo16 三个数据集上的对比实验结果表明,本文提出的模型能有效提高谣言检测的性能。本文的主要贡献总结如下:

(1) 为社交媒体事件构造新的全局无向传播图,并结合事件传播的方向特征,采用 GCN 从多个层次的传播行为中提取出事件的潜在传播特征。

(2) 摒弃了传统采用硬连接的方式来处理关联节点之间的联系程度,提出了一种新颖的节点级注意力机制,给事件中的每个关系动态地分配适当的权重,有效地识别节点之间关系的真实性。

(3) 为更有效融合不同传播层次的特征信息,提出了一种基于注意力机制的多特征池化方法,有效解决不同层次特征信息之间难以交互的问题。

1 相关工作

从自动检测谣言技术诞生开始,相关学者开展大量的工作,已经取得许多突破性的成果。目前的谣言检测研究根据不同的特征类型可大致分为基于内容的方法和基于传播模式的方法。

1.1 基于内容的谣言检测方法

在早期的谣言检测方法中,主要是利用手工提

取与设计的特征,结合机器学习方法对谣言进行分类^[9-10],它们依赖的特征主要是用户自身的属性(如是否为官方认证号、用户粉丝数量、用户信用度以及用户历史发帖数量等)、发布的文本内容特征以及文本的情感特征等。这类学习方法通常只能挖掘到谣言浅层的特征,而且检测效果过于依赖特征的选取。

深度学习技术出现后,研究者们开始关注网络模型结构的设计,神经网络能够自动挖掘到相关特征的更多细节,可以避免人工手动提取的烦琐过程。Ma^[11]等人研究发现谣言源贴和评论中都包含大量有价值的语义信息,提出基于循环神经网络(Recurrent Neural Networks, RNN)的谣言分类模型来捕捉新闻事件和评论中的上下文信息。由于谣言评论和谣言源贴是相互关联的,二者可相互补充作为检测谣言的重要线索,Shu 等人^[12]利用协同注意力机制对谣言句子和用户评论之间的关系联合捕捉,并为谣言检测研究提供了合理的解释。同时,还有研究者利用社交媒体事件中蕴藏的情感信号以辅助检测谣言,诸如 Zhang 等人^[13]利用情感词典获取谣言的源贴和评论双重情感特征以及双重情感的差距,葛晓义等人^[14]分别从全局和局部的角度挖掘谣言文本语义特征和评论情感特征之间的相关性。

1.2 基于传播模式的谣言检测方法

社交媒体事件具有高度动态性和复杂性,传播行为作为社交媒体事件固有的特性,有效挖掘传播特征能给谣言检测带来巨大提升。

为有效利用谣言的传播线索,一类方法将谣言的传播过程建模成双向的传播树或传播图结构。如 Ma 等人^[4]对事件的非线性传播结构建模,并使用基于自顶向下和自底向上树状结构的递归神经网络 RvNN 来检测谣言,这一方法首次将事件的文本内容语义特征和其传播特征结合起来;随后又提出了基于树形 Transformer 的谣言检测模型^[15],结合关联语境中的立场信息来增强父节点的表示。由于图卷积网络技术^[6]在处理图结构数据上的突出表现^[16-18],Bian 等人^[5]提出了一种双向的图卷积网络模型(Bi-Directional Graph Convolutional Networks, BiGCN)检测谣言,该方法首次将谣言的传播建模成自顶向下的传播图和自底向上的扩散图。一些研究在双向传播图的基础上引入辅助特征来增强谣言检测的效果,如用户影响力因素^[19]、时间因素^[20]。

另一类基于传播的方法将谣言的传播过程建模成无向交互图^[2,21-22],这类结构可同时从父节点和子节点甚至是兄弟节点中聚合相关的邻域信息,但是这些方法并没有考虑社交媒体事件固有的传播因果特性、固有的扩散结构特性以及频繁的不真实传播特性,这限制了对谣言潜在结构特征的学习。为了消除传播结构中固有的不确定性,Wei 等人^[23]提出通过观测图的先验信念来控制消息传递,以此来替换传播图中固定的边权重,然而对观测图的先验信念进行假设可能出现不确定的偏差。

虽然基于谣言传播的研究已经取得了一定进展,但目前的研究无法有效处理传播结构中的不真实关系,也没有同时联合谣言传播的方向和全局传播结构来捕捉潜在的特征。本文在谣言传播方向的基础上,引入全局无向传播图,通过多头的节点级注意力评估潜在的传播关系,进而实现对谣言有效的识别。

2 MDPAN 模型方法

本文提出的谣言检测模型 MDPAN 的总体架构如图 2 所示,该模型主要包括五个部分,分别为输入表示层、节点级注意力层、多层次图卷积层、注意力池化层以及谣言分类器。

2.1 问题描述

本文假设 $C = \{c_1, c_2, \dots, c_m\}$ 为数据集中所有事件的集合,其中, c_i 表示集合 C 中的第 i 个事件, m 表示数据集中事件的总数。对于集合 C 中的任意事件 c_i ,其定义为 $c_i = \{r_0^i, r_1^i, \dots, r_{n_i-1}^i, G_i\}$,其中, r_0^i 表示事件 c_i 的源贴, r_j^i 表示第 i 个事件 c_i 的第 j 条响应贴(评论贴或转发贴), n_i 表示事件 c_i 中的总帖子数, $G_i = \{G_i^{\text{TD}}, G_i^{\text{BU}}, G_i^{\text{UD}}\} = \{V_i, E_i\}$ 分别表示事件 c_i 自顶向下的传播图、自底向上的扩散图、全局无向传播图, $V_i = \{r_0^i, r_1^i, r_2^i, \dots, r_{n_i-1}^i\}$ 表示事件 c_i 中所有帖子节点的集合, $E_i = \{e_{st}^i \mid s, t = 0, 1, \dots, n_i - 1\}$ 为事件 c_i 中被评论或被转发帖子节点与其评论或转发帖子节点之间的边的集合。 $A_i = \{a_{st}^i\} \in \mathbb{R}^{n_i \times n_i}$ 表示事件传播图的邻接矩阵,其中, a_{st}^i 初始值的定义为:当 $e_{st}^i \in E_i$ 时, $a_{st}^i = 1$, 否则 $a_{st}^i = 0$ 。

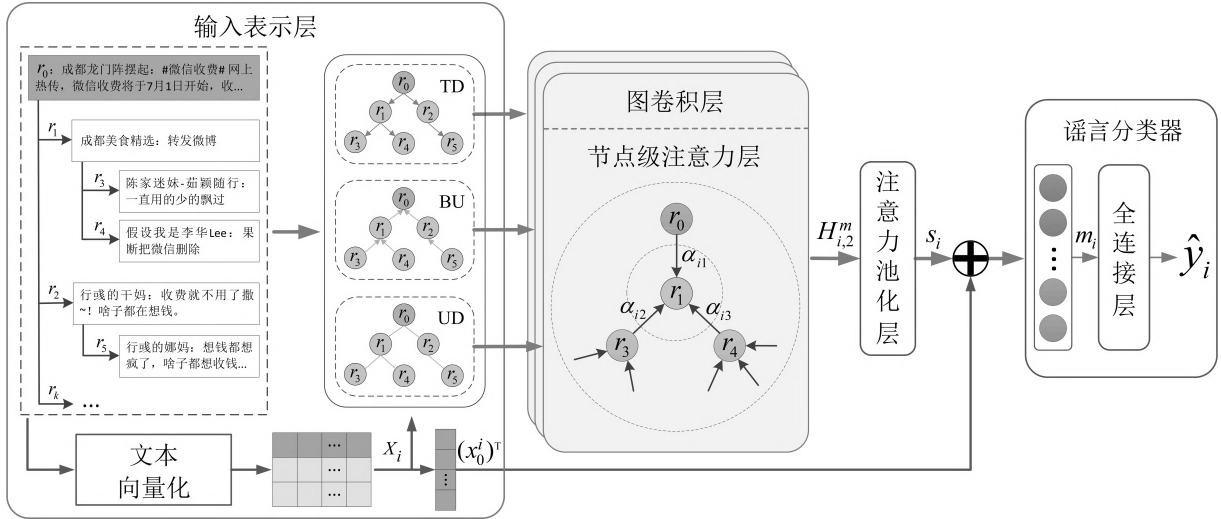


图2 MDPAN 的模型架构

在本文所使用的谣言数据集中,每个事件 c_i 都有唯一的实际类别 $y_i^* \in Y$ 与之对应,其中 Y 代表所有谣言类别的集合。谣言检测任务的目标是根据已有的数据集来训练特定的谣言分类器 $f(C|c_i; G_i; A_i; \theta) \rightarrow Y$, 其中 θ 为分类器模型的所有参数集合。

2.2 模型方法描述

2.2.1 输入表示层

根据给定的输入样本,为每个事件构造出相应的传播图,初始传播特征图的构造类似于 Bian 等人^[5]的工作,首先为每个事件 c_i 构建出两个相反方向的有向图,分别是自顶向下的传播图 G_i^{TD} 和自底向上的扩散图 G_i^{BU} ,为了利用谣言传播过程的全局结构特性,本文还引入了无向传播图 G_i^{UD} 。传播图中的每个节点代表着事件中的源贴或相关的响应贴,传播图中的连接边则表示两个节点之间有直接的评论或转发关系。使用 $X_i = [x_0^i, x_1^i, \dots, x_{n_i-1}^i] \in \mathbb{R}^{n_i \times d_i}$ 作为事件 $c_i \in C$ 的源贴以及相关响应贴的文本内容表示,其中, $x_0^i \in \mathbb{R}^{d_i}$ 为源贴的文本向量表示,其初始值提取自 TF-IDF 值的前 5 000 个词的索引词频对, d_i 为文本特征的维度,得到不同帖子的文本特征向量之后嵌入到对应的节点中,对于自顶向下的传播图、自底向上的扩散图以及全局无向传播图,它们的初始输入特征矩阵均相同,即 $X_i^{TD} = X_i^{BU} = X_i^{UD} = X_i$ 。基于事件 c_i 中对源贴 x_0^i 的一系列评论和转发关系,构造出事件 c_i 的邻接矩阵 $A_i \in \mathbb{R}^{n_i \times n_i}$,并基于此得到三种不同层次传播结构图的邻接矩阵,分别为 $A_i^{TD} = A_i$, $A_i^{BU} = A_i^T$ 以及 $A_i^{UD} = A_i + A_i^T$ 。

2.2.2 节点级注意力层

受图注意力机制^[24]思想的启发,引入节点级注意力机制来具体学习传播图中每个节点的表示,从而降低不真实的节点关系带来的负面影响。节点级注意力机制的核心是利用自注意力机制来评估当前节点的直接邻居节点和间接邻居节点的重要性,以动态调整两个关联节点之间的权重系数,在原始传播结构的基础上对传播关系的稳定性和真实性进行重新描述。具体来说,对于事件 $c_i \in C$ 的任意节点 r_j^i ,其初始特征向量为 $\mathbf{h}_j^i = \mathbf{x}_j^i \in \mathbb{R}^{d_i}$,它作为节点级注意力层的初始输入, v_k^i 为 r_j^i 的邻居节点,其初始特征向量为 \mathbf{h}_k^i ,其中, $v_k^i \in V_k^i$, V_k^i 表示邻居节点的集合,对于任意一个节点对 (r_j^i, v_k^i) ,通过获得其不同邻居节点的权重系数 α_{jk}^i ,实现以软连接方式为所有连接边重新动态地分配关注系数,如式(1)所示。

$$\alpha_{jk}^i = \text{Softmax}(\sigma(\mathbf{a}^T [\mathbf{W} \mathbf{h}_j^i \parallel \mathbf{W} \mathbf{h}_k^i])) \quad (1)$$

其中, \mathbf{W} 是可学习的参数矩阵, \parallel 表示将两个特征向量进行拼接, \mathbf{a} 为可学习的注意力权重向量, $\sigma(\cdot)$ 为非线性的激活函数 LeakyReLU。

在得到节点与邻居关系的权重之后,将邻居节点的特征向量 \mathbf{h}_k^i 与权重系数 α_{jk}^i 进行聚合,为了使模型能够融合来自不同特征子空间的信息,使用多头自注意力机制来稳定地衡量当前节点 r_j^i 与其不同邻居节点的关系,设置一个用于表示多头注意力机制头数的超参数 T ,并进行 T 次计算,得到节点 r_j^i 的嵌入表示 $\tilde{\mathbf{h}}_j^i$,如式(2)所示。

$$\tilde{\mathbf{h}}_j^i = \bigoplus_{t=1}^T \sigma(\sum_{k \in V_k^i} \alpha_{jk}^i \cdot \mathbf{h}_k^i) \quad (2)$$

传播图、扩散图以及全局无向图中的同一个节

点共享权重系数 α_{jk}^i , 且拥有相同的初始节点嵌入 $\tilde{\mathbf{h}}_j^i$ 。处理完事件 c_i 中所有节点与其邻域内其他节点之间的关系后, 得到传播图的节点特征矩阵 $\tilde{\mathbf{H}}_i = [\tilde{\mathbf{h}}_0^i, \tilde{\mathbf{h}}_1^i, \dots, \tilde{\mathbf{h}}_{n_i-1}^i]$, 它将作为多层次图卷积层的输入, 以进一步学习深层的结构特征。

2.2.3 多层次图卷积层

图卷积网络 (Graph Convolutional Networks, GCN)^[6] 能够有效处理图数据的结构信息, 因为它模拟了消息如何在节点之间进行传递, 并能够通过聚合局部邻域内所有节点的特征来更新当前节点特征, 它的消息传递结构定义如式 (3) 所示。

$$\mathbf{H}_k = M(\mathbf{A}, \mathbf{H}_{k-1}; \mathbf{W}_{k-1}) = \sigma(\hat{\mathbf{A}}\mathbf{H}_{k-1}\mathbf{W}_{k-1}) \quad (3)$$

其中, $\hat{\mathbf{A}} = \tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}}$ 是进行归一化后的邻接矩阵, $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_N$ (即为每个节点添加自连接), $\tilde{\mathbf{D}}_{ii} = \sum_j \tilde{\mathbf{A}}_{ij}$ 表示节点的度矩阵, \mathbf{H}_{k-1} 为隐藏层特征矩阵, \mathbf{W}_{k-1} 为可训练的参数矩阵, $\sigma(\cdot)$ 是激活函数。

本文分别应用了三个并行的图卷积层 (Graph Convolutional Layers, GCL) 来同时处理事件不同层次传播图的空间结构特征。具体地说, 每个 GCL 内包含两层堆叠的 GCN 块, 对于任意事件 $c_i \in C$, 为了防止其源贴信息在网络层之间过度流失, 本文沿用 Bian 等人^[5] 提出的源贴特征增强策略, 即在第 k 层 GCN 处, 将每个节点的当前隐藏层特征向量 $\tilde{\mathbf{H}}_{i,k}$ 与第 $k-1$ 层 GCN 中根源节点的隐藏层特征向量 $(\tilde{\mathbf{H}}_{i,k-1})^r$ 进行拼接, 构造出新的隐藏层特征矩阵 $\mathbf{H}_{i,k}$, 以提高源贴信息的影响力。具体计算过程如式 (4)、式 (5) 所示。

$$\tilde{\mathbf{H}}_{i,k}^m = \sigma(\hat{\mathbf{A}}_i^m \tilde{\mathbf{H}}_{i,k-1}^m \mathbf{W}_{i,k-1}^m) \quad (4)$$

$$\mathbf{H}_{i,k}^m = \text{concat}(\tilde{\mathbf{H}}_{i,k}^m, (\tilde{\mathbf{H}}_{i,k-1}^m)^r) \quad (5)$$

其中, $m \in \{\text{TD}, \text{BU}, \text{UD}\}$ 表示不同层次的传播模式, $k=1, 2$, $\tilde{\mathbf{H}}_{i,0}^m = \tilde{\mathbf{H}}_i = [\tilde{\mathbf{h}}_0^i, \tilde{\mathbf{h}}_1^i, \dots, \tilde{\mathbf{h}}_{n_i-1}^i]$ 为图卷积层的初始输入特征矩阵, $\mathbf{H}_{i,1}^m \in \mathbb{R}^{n_i \times v_{i,1}}$ 和 $\mathbf{H}_{i,2}^m \in \mathbb{R}^{n_i \times v_{i,2}}$ 分别为图卷积层中的两层隐藏层特征, $\mathbf{W}_{i,0}^m \in \mathbb{R}^{d \times v_{i,1}}$ 和 $\mathbf{W}_{i,1}^m \in \mathbb{R}^{d \times v_{i,2}}$ 是隐藏层的权重矩阵, $\sigma(\cdot)$ 为非线性的激活函数 ReLU。

对于自顶向上传播图 TD、自底向下的扩散图 BU 以及全局无向传播图 UD, 均采用上述的节点更新方式进行结构特征提取, 分别得到相应的隐藏层节点特征表示 $\mathbf{H}_{i,1}^{\text{TD}}$ 、 $\mathbf{H}_{i,2}^{\text{TD}}$ 、 $\mathbf{H}_{i,1}^{\text{BU}}$ 、 $\mathbf{H}_{i,2}^{\text{BU}}$ 和 $\mathbf{H}_{i,1}^{\text{UD}}$ 、 $\mathbf{H}_{i,2}^{\text{UD}}$, 具体的节点更新过程如表 1 所示。

表 1 MDPAN 模型的节点更新算法

MDPAN 模型的节点更新算法

输入: 事件 c_i 不同层次的传播图 $G_i = \{V_i, E_i\}$, 事件初始特征矩阵 $\mathbf{X}_i = [\mathbf{x}_0^i, \mathbf{x}_1^i, \dots, \mathbf{x}_{n_i-1}^i]$, 不同层次传播图的邻接矩阵 $\mathbf{A}_i^{\text{TD}}, \mathbf{A}_i^{\text{BU}}, \mathbf{A}_i^{\text{UD}}$
 输出: 图卷积层的隐藏层特征 $\mathbf{H}_{i,2}^{\text{TD}}, \mathbf{H}_{i,2}^{\text{BU}}, \mathbf{H}_{i,2}^{\text{UD}}$
 1. FOR $m \in \{\text{TD}, \text{BU}, \text{UD}\}$ DO
 2. FOR $r_j^i \in V_i$ DO
 3. $\mathbf{h}_j^i \leftarrow \mathbf{x}_j^i$
 4. FOR $v_k^i \in V_k$ DO
 5. $\mathbf{h}_k^i \leftarrow \mathbf{x}_k^i$
 6. 通过式 (1) 计算邻居节点关系的权重系数
 7. END
 8. 通过式 (2) 计算得到真实的节点嵌入 $\tilde{\mathbf{h}}_j^i$
 9. $\tilde{\mathbf{H}}_i \leftarrow [\tilde{\mathbf{h}}_0^i, \tilde{\mathbf{h}}_1^i, \dots, \tilde{\mathbf{h}}_{n_i-1}^i]$
 10. END
 11. 通过式 (4)、式 (5) 计算节点的隐藏层特征
 12. END

2.2.4 注意力池化层

利用 2.2.3 节得到不同层次传播图的节点级特征表示后, 为了从中聚合更多细节信息, 并且获取更具判别力的输出特征, 通常使用固定的池化方法对特征降维和抽取, 然而由于池化操作是人工指定, 而不是通过模型学习得到, 因此, 往往很难找到一种适合不同模型本身的特征汇聚方法。为了解决这个问题, 本文引入注意力机制来融合均值池化 (Mean-pooling)、最大池化 (Max-pooling) 以及随机池化 (Stochastic-pooling)^[25] 三种不同汇聚操作的结果, 以获得更具代表性的特征表示, 如图 3 所示, 这种做法综合考虑了不同池化方法的优势, 具体如式 (6)~式 (8) 所示。

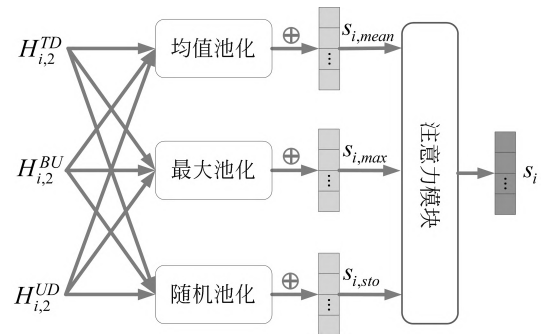


图 3 注意力池化层结构

$$s_{i,\text{mean}}^m = \text{meanpooling}(\mathbf{H}_{i,2}^m) \quad (6)$$

$$s_{i,\text{max}}^m = \text{maxpooling}(\mathbf{H}_{i,2}^m) \quad (7)$$

$$s_{i,\text{sto}}^m = \text{stopooling}(\mathbf{H}_{i,2}^m) = \sum_{j \in R} p_j a_j \quad (8)$$

其中, $m \in \{TD, BU, UD\}$ 表示不同层次的传播模式, meanpooling 为节点级向量的均值池化函数, maxpooling 为节点级向量的最大值池化函数, stopooling 为节点级向量的随机值池化函数, 它是一种基于概率加权形式的平均汇聚操作, 隐藏层特征映射 $H_{i,2}^m$ 中的区域记为 R , $p_j = \frac{a_j}{\sum_{k \in R} a_k}$ 为归一化区域 R 后计算出的每个区域的概率, 其中, $a_j \in H_{i,2}^m$ 。

然后将同一种池化函数下得到的不同层次传播模式的表示拼接起来, 以从不同特征中聚集关键信息, 具体如式(9)所示。

$$s_{i,p} = \text{concat}(s_{i,p}^m) = \text{concat}(s_{i,p}^{TD}, s_{i,p}^{BU}, s_{i,p}^{UD}) \quad (9)$$

其中, $p \in \{\text{mean}, \text{max}, \text{sto}\}$, 在训练网络模型时, 充分吸收这三种不同池化方法的优势。均值池化侧重于捕捉全局信息, 最大池化更关注局部信息, 随机池化则介于平均池化和最大池化之间, 由于其具备随机性, 能够有效防止训练模型时出现过度拟合的现象。通过引入注意力机制融合以上三种池化方法得到最终的向量表示, 具体如式(10)~式(14)所示。

$$q_i = [s_{i,\text{mean}}, s_{i,\text{max}}, s_{i,\text{sto}}] W_{q_i} \quad (10)$$

$$k_i = [s_{i,\text{mean}}, s_{i,\text{max}}, s_{i,\text{sto}}] W_{k_i} \quad (11)$$

$$v_i = [s_{i,\text{mean}}, s_{i,\text{max}}, s_{i,\text{sto}}] W_{v_i} \quad (12)$$

$$[\tilde{s}_{i,\text{mean}}, \tilde{s}_{i,\text{max}}, \tilde{s}_{i,\text{sto}}] = \text{Softmax}\left(\frac{q_i \times k_i^T}{\sqrt{d_i}}\right) v_i \quad (13)$$

$$s_i = \sum_{j \in p} \frac{\|\tilde{s}_{i,j}\|_2}{\|\tilde{s}_{i,\text{mean}}\|_2 + \|\tilde{s}_{i,\text{max}}\|_2 + \|\tilde{s}_{i,\text{sto}}\|_2} \tilde{s}_{i,j} \quad (14)$$

其中, $\tilde{s}_{i,j}$ 为自注意力机制之后获得的不同池化层向量, $\|\tilde{s}_{i,j}\|_2$ 表示向量 $\tilde{s}_{i,j}$ 的模长, s_i 则为融合不同池化方法后的最终向量表示。

2.2.5 谣言分类器

在聚合了传播特征、扩散特征以及全局传播特征后, 得到了事件 c_i 基于社会背景下的结构特征表示 s_i , 考虑到源贴的文本内容特征对识别谣言的重要性, 将事件 c_i 源贴的文本信息 $(x_0^i)^T$ 与社会背景的结构信息 s_i 进行拼接, 得到融合后的最终特征向量 m_i , 进一步增强事件源贴的文本信息对检测谣言的影响, 如式(15)所示。

$$m_i = \text{concat}(s_i, (x_0^i)^T) \quad (15)$$

然后, 定义一个全连接层来降低特征向量的维度, 再通过 Softmax 函数来获得事件 c_i 所属不同类

别标签的概率向量 \hat{y}_i , 如式(16)所示。

$$\hat{y}_i = \text{Softmax}(\text{fc}(m_i)) \quad (16)$$

其中, $\text{fc}(m_i) = W_i m_i + b_i$, W_i, b_i 分别表示全连接层中可学习的权重矩阵和偏置项。

最后, 计算预测值 \hat{y}_i 和真实值 y_i 之间的交叉熵作为模型的损失值, 通过最小化该损失值来优化 MDPAN 模型的参数 θ 。

$$\text{Loss}(\theta) = - \sum_i y_i \log(\hat{y}_i) \quad (17)$$

3 实验

为验证本文所提出模型的有效性, 分别在三个主流公开数据集上进行大量实验, 对比了 9 种基准模型和最新的谣言检测模型, 通过 4 种评测指标验证了 MDPAN 模型的有效性, 并进行消融实验以探索各个模块对谣言检测的影响。

3.1 实验数据

本文所使用的实验数据集 Weibo16^[11]、Twitter15^[2]、Twitter16^[2] 分别收集自中国和美国最大的社交媒体平台新浪微博和 Twitter。其中, Weibo16 数据集总共包含 4 464 个事件, 每个事件唯一对应于的虚假谣言(False Rumor, FR)和非谣言(Non-Rumor, NR)二分类标签, 而 Twitter15、Twitter16 数据集则分别包含 1 490、818 个事件, 每个事件对应于唯一非谣言(Non-Rumor, NR)、虚假谣言(False Rumor, FR)、真实谣言(True Rumor, TR)和未经证实的谣言(Unverified Rumor, UR)四分类标签, 三个数据集的信息如表 2 所示。

表 2 实验数据集的数据详情

数据集	Twitter15	Twitter16	Weibo16
总事件数	1 490	818	4 464
总用户数	276 663	173 487	2 746 818
总帖子数	331 612	204 820	3 805 656
非谣言	374	205	2 351
虚假谣言	370	205	2 313
真实谣言	372	205	0
未经证实的谣言	374	203	0

3.2 实验设置与评估指标

为尽可能保证评估结果的公正性, 本文将在上

述三个数据集上分别进行 10 轮迭代训练,每轮迭代分别使用 5 折交叉验证对模型进行充分训练和评估,每折交叉验证时,对数据集按 4:1 的比例重新划分训练集和测试集。在实验中,节点级注意力层的头数设置 T 为 4,多层次图卷积层中 GCN 的块数 K 设为 2,初始特征向量维度设为 5 000,隐藏层特征向量维度设置为 64,初始学习率设为 0.000 5,并在学习过程中逐步衰减。在训练过程中,通过随机梯度下降算法对模型参数进行更新,并使用 Adam 算法进行优化。每折交叉验证最大训练 200 个 Epoch, Twitter15、Twitter16 和 Weibo16 数据集上的 Batch Size 分别设置为 256、128 和 32。

本文采用多种指标来综合评测 MDPAN 模型方法的性能,其中, Twitter15、Twitter16 数据集上使用整体的准确率(Accuracy, Acc)和每个类别的 F_1 值作为评估指标, Weibo16 数据集上使用整体的准确率(Accuracy, Acc)、每个类别的精确率(Precision, P)、召回率(Recall, R)以及 F_1 值作为评估指标。这些衡量指标的具体含义和计算如式(18)~式(21)所示。

准确率:所有测试样本中分类正确的比例。

$$\text{Acc} = \frac{\text{TP} + \text{FN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (18)$$

精确率:预测结果为类别 c 中预测正确的比例。

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (19)$$

召回率:实际标签结果为类别 c 中预测正确的比例。

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (20)$$

F_1 值:精确率和召回率的调和平均值。

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (21)$$

其中, TP、TN、FP、FN 表示的含义如表 3 的混淆矩阵所示。

表 3 混淆矩阵

数据集	实际结果为类别 c	实际结果为非类别 c
预测结果为类别 c	TP	FP
预测结果为非类别 c	FN	TN

3.3 对比方法

本文在相同的数据集下将 MDPAN 模型与现

有先进的基准模型进行比较,以证明本文方法 MDPAN 的优越性。

(1) **DTC**^[9]: 基于手工提取特征获取信息的可信度,使用决策树分类器算法来识别谣言。

(2) **SVM-TS**^[10]: 利用线性的支持向量机分类器对事件社会背景的时间序列特征进行建模,以用于检测谣言。

(3) **RvNN**^[4]: 基于自顶向下和自底向上双向传播树的递归神经网络模型,利用事件树状的传播结构检测谣言。

(4) **PPC**^[26]: 一种基于事件传播路径的方法,利用 RNN 和 CNN 相结合的方式学习路径上的用户特征,从而对事件进行分类。

(5) **BiGCN**^[5]: 基于双向图卷积网络提取自顶向下的传播结构和自底向上的扩散结构用于谣言检测。

(6) **VAE-GCN**^[2]: 使用基于图卷积网络的自动编码器及其变体来学习传播结构中的潜在信息。

(7) **EBGCN**^[23]: 利用边增强贝叶斯图卷积网络来处理传播结构中的不确定性,捕捉稳健的结构特征用于谣言检测。

(8) **GSMRB**^[27]: 一种基于图的关键语义挖掘的谣言检测框架,将语义信息建模为图结构,利用门控图神经网络对谣言的配对文本信息进行建模。

(9) **DDGCN**^[28]: 基于双动态图卷积网络的谣言检测方法,对消息在传播过程中的动态以及来自知识图的背景知识的动态进行建模。

由于 GSMRB 方法和 DDGCN 方法没有公开实现代码,本文使用了其原文中公开的实验结果,其中, GSMRB 方法仅对比在 Twitter15 和 Twitter16 数据集上的结果, DDGCN 方法仅对比了在 Weibo16 数据集上的结果。

3.4 实验结果与分析

表 4、表 5 分别展示了各种对比方法在 Weibo16、Twitter15 和 Twitter16 数据集的谣言检测结果。对比实验的结果表明,首先,依赖于手工特征的机器学习方法,如决策树分类器 DTC 以及基于时间序列的支持向量机分类器 SVM-TS 的检测效果明显弱于其他深度学习检测方法,原因在于深度学习方法依靠神经网络可以有效学习到谣言深层的特征表示,更容易捕捉细节信息。其次,在有关谣言传播模式的研究中,基于树状的递归神经网络方法 RvNN 以及基于循环网络和卷积网络相结合的传播路径分类方法 PPC,它们的检测性能落后于基于图卷积网

络的谣言检测方法,这是由于图卷积网络技术更适合于处理社交网络事件的传播属性。最后,与几种较新的基于图卷积网络的谣言检测方法相比,本文提出的谣言检测模型 MDPAN 也展现出了更加优越的性能。具体地说,在 Weibo16 数据集上达到了 95.8% 的整体准确率,比现有的方法 DDGCN 高出了 1%,在 Twitter15、Twitter16 数据集上分别达到了 89.7%、91.7% 的整体准确率,在先进的谣言检测方法 GSMRB 的基础上分别提升了 0.6%、1.6%。另外,其他的评估指标也略有提升,综合证实了 MDPAN 谣言检测方法的有效性。

表 4 Weibo16 数据集上的实验结果

(单位: %)

方法	类别	Acc	P	R	F_1
DTC	NR	83.1	81.5	82.4	81.9
	FR		84.7	81.5	83.1
SVM-TS	NR	85.7	87.8	83.0	85.7
	FR		83.9	88.5	86.1

续表

方法	类别	Acc	P	R	F_1
RvNN	NR	90.8	90.4	91.8	91.1
	FR		91.2	89.7	90.5
PPC	NR	92.1	94.9	88.9	91.8
	FR		89.6	96.2	92.3
BiGCN	NR	94.0	92.5	94.3	93.3
	FR		95.1	94.3	94.7
VAE-GCN	NR	94.4	91.7	96.4	93.6
	FR		96.8	92.1	94.0
EBGCN	NR	94.1	92.2	94.5	93.2
	FR		94.6	94.0	94.3
DDGCN	NR	94.8	95.3	94.8	95.0
	FR		95.3	94.8	95.0
MDPAN	NR	95.8	95.5	95.4	95.4
	FR		96.8	93.3	95.0

表 5 Twitter15 和 Twitter16 数据集上的实验结果

(单位: %)

方法	Twitter15					Twitter16				
	Acc	NR	FR	TR	UR	Acc	NR	FR	TR	UR
		F_1	F_1	F_1	F_1		F_1	F_1	F_1	F_1
DTC	45.5	73.3	35.5	31.7	41.5	46.5	64.3	39.3	41.9	40.3
SVM-TS	54.4	79.6	47.2	40.4	48.3	57.4	75.5	42.0	57.1	52.6
RvNN	72.3	68.2	75.8	82.1	65.4	73.7	66.2	74.3	83.5	70.8
PPC	84.2	81.1	87.5	79.0	81.8	86.3	82.0	89.8	83.7	84.3
BiGCN	86.6	84.3	87.0	91.1	83.5	88.4	81.0	86.3	94.3	87.4
VAE-GCN	85.6	74.9	79.5	90.5	80.9	86.8	79.5	80.9	94.7	88.5
EBGCN	87.6	85.4	88.0	91.8	81.8	91.0	86.4	87.8	94.4	87.4
GSMRB	89.1	88.4	87.9	89.0	88.5	90.1	89.7	89.2	88.7	88.4
MDPAN	89.7	89.7	89.0	92.7	86.1	91.7	89.7	89.6	92.6	89.9

本文方法主要存在以下优势,首先,与 BiGCN、VAE-GCN、EBGCN、GSMRB 以及 DDGCN 这些基于图卷积网络的谣言检测方法相比,MDPAN 方法将谣言的传播细分为自顶向下的传播结构特征、自底向上的扩散结构特征以及全局无向结构特征。其中,自顶向下的传播结构特征主要关注当前节点与其直接传播邻居之间的传播关系,用于表达传播方向和传播路径等信息;自底向上的扩散结构特征则关注谣言在传播过程中涉及的非直接传播邻居,用

于表达谣言的扩散能力和影响力等信息;全局无向结构特征则考虑了整个谣言传播图的全局结构信息,包括节点间的相对位置和连接方式等,用于表达传播图的整体特征和网络拓扑结构。联合多层次的传播性质可挖掘到节点潜在的特征,可捕捉更全面丰富的空间结构和细节信息,这对于检测传播范围广、影响力大的谣言来说是十分关键的。其次,与基于贝叶斯方法的 EBGCN 以及基于动态知识图和动态传播结构的 DDGCN 相比,MDPAN 方法的优势

在于通过多头的自注意力机制来动态地调整两个关联帖子之间的权重系数,无须对观测图的先验信念做出假设,且在不引入外部知识的情况下,也可以稳定地描述邻居节点及其连接关系的真实性。

3.5 消融实验

为了验证 MDPAN 模型中一些关键模块的相对重要性,本节展开了一系列涉及模型不同模块的消融实验,具体如下:

(1) **MDPAN w/o UD**: 移除了全局无向传播图模块,仅考虑了事件传播的方向特性,不考虑全局特性。

(2) **MDPAN w/o NA**: 移除了节点级注意力模块,不考虑事件网络中存在的不真实关系,将原始传播结构中的所有联系视为完全可靠的关系。

(3) **MDPAN w/o AP**: 移除了注意力池化模块,在得到多层次图卷积层最终的隐藏层输出后,只进行统一的均值池化操作。

(4) **MDPAN w/o ST**: 移除了源贴文本特征模块,在注意力池化层后,不额外将源贴的文本特征与事件的传播结构特征进行拼接,只使用事件的结构特征。

不同消融实验在三个数据集上的整体准确率和整体 F_1 值如图 4 所示,从实验结果上来看,所有其他的方法都优于 MDPAN w/o NA,这表明有效处理社交网络中潜在的不真实关系对识别谣言十分重要。当不考虑谣言传播的全局结构特征时(即 MDPAN w/o UD),谣言检测性能明显下降,这表明在谣言传播的方向特性的基础上引入全局结构特性,有利于提升谣言检测的效果。同时,由于 MDPAN w/o ST 方法在注意力池化层之后没有进行源贴的文本特征增强,导致谣言检测效果受到影响,这表明了充分利用源贴文本内容信息的重要性。由 MDPAN w/o AP 的实验结果可以看出,基于注意力的多池化方法更能有效融合多层次的特征信息,从而给谣言检测的效果带来进一步提升。

3.6 参数分析

本节探索 MDPAN 模型中主要的参数对谣言检测结果的影响,其中包括节点级注意力的头数 T 以及多层次图卷积层中 GCN 的块数 K ,实验结果如图 5、图 6 所示。

从图 5 中可以观察到,当 T 设为 4 时,模型检测性能最佳,这表明在节点级别上使用 4 个不同的

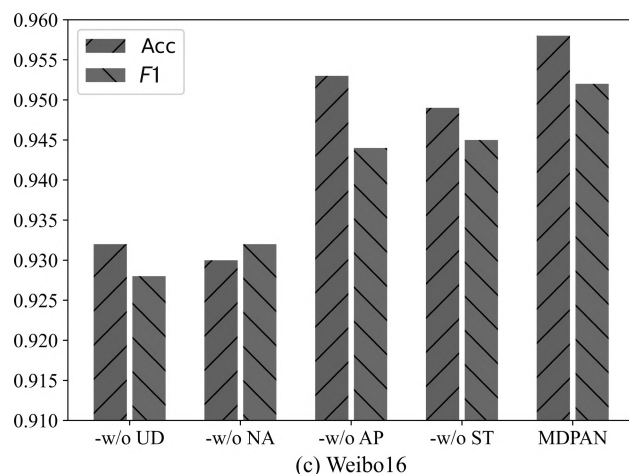
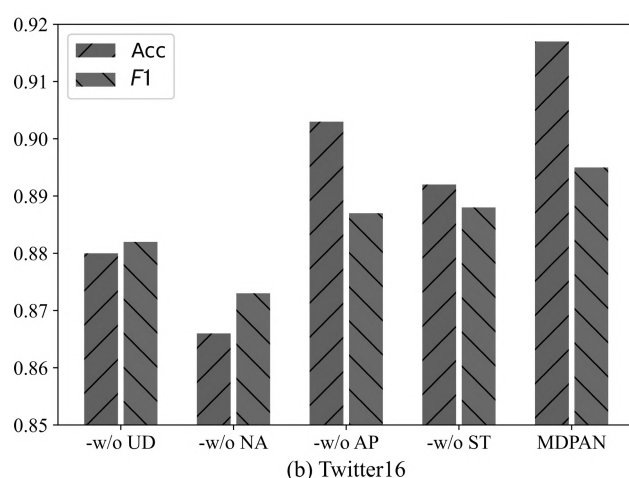
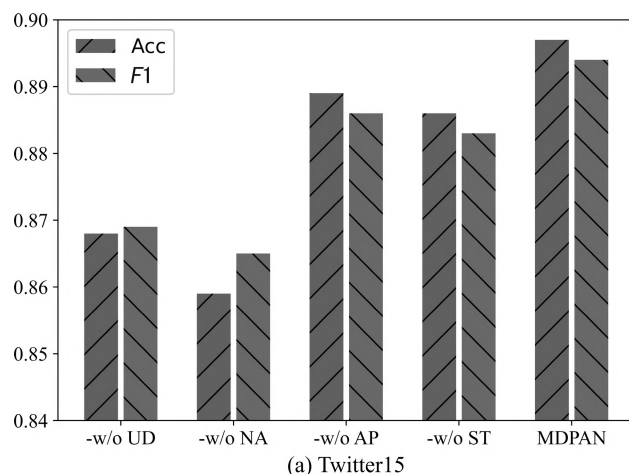


图 4 MDPAN 在三个数据集上的消融实验结果

特征子空间学习权重系数可以更好地捕捉节点之间的真实关系;而当 T 设为 8 时,性能开始出现下降的趋势,可能原因是 T 过大时,会导致模型过度关注节点关系的细节特征,而忽略了全局信息的影响。

从图 6 中可发现,当 K 设为 2 时,模型检测效果最好,而当 K 进一步增大时,模型的性能开始下降,这可能是由于过多的 GCN 层数导致模型过拟

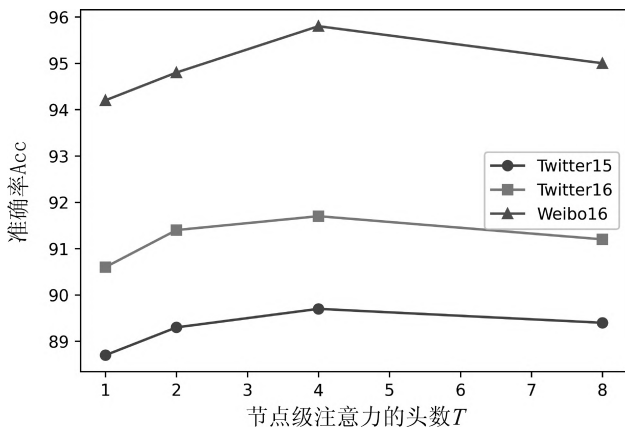


图 5 不同头数对检测结果的影响

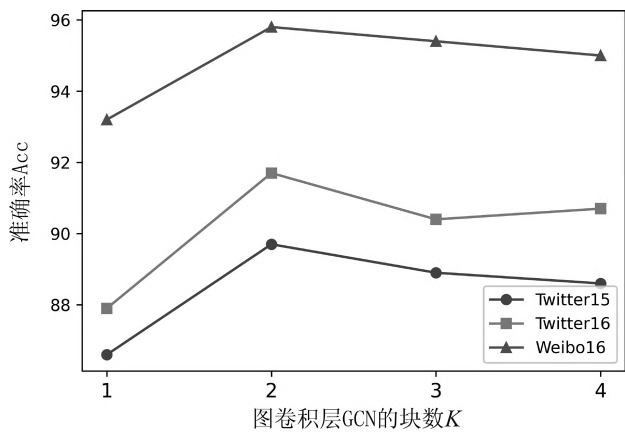


图 6 不同 GCN 块数对检测结果的影响

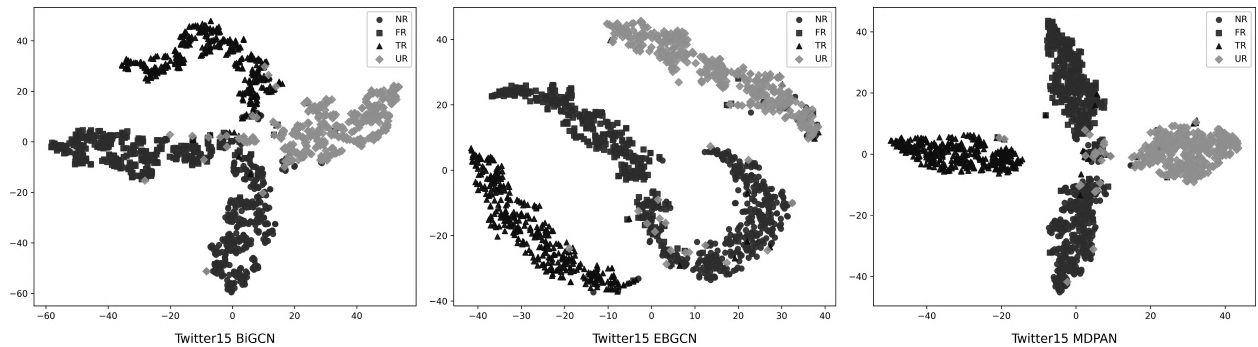


图 7 不同模型在 Twitter15 数据集上中间层特征降维后的散点分布图

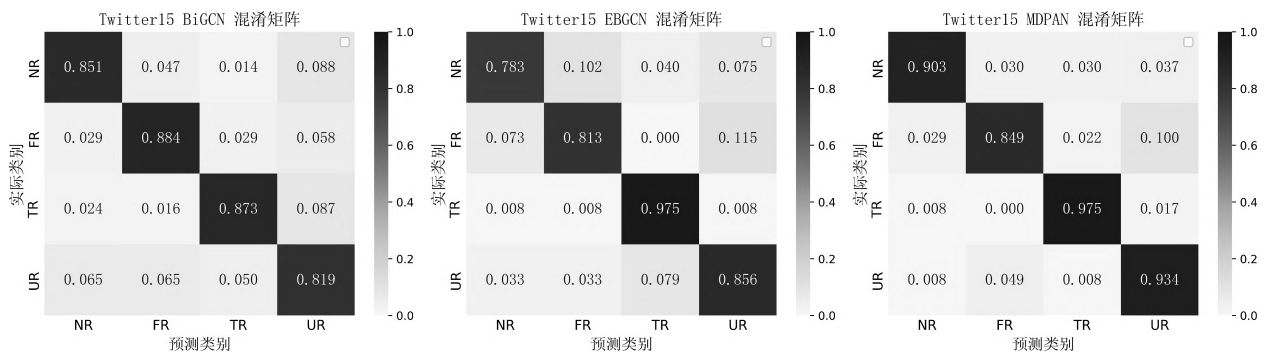


图 8 不同模型在 Twitter15 数据集上的混淆矩阵热力图

合,或者是由于 K 值过大导致信息在传播时出现了过度平滑化的情况。

3.7 可视化分析

本节对谣言数据进行可视化展示,以 Twitter15 数据集为例,使用 TSNE 技术对模型融合后的高维特征进行适当降维,生成了不同谣言类别的分布图,并进一步对比基准的双向传播图 BiGCN 模型以及基于边推理框架的 EBGCN 模型。如图 7 所示,可以观察到相比于 BiGCN 模型,MDPAN 模型在不同谣言类别之间呈现更大的间隔,这表明对谣言传播的方向和全局传播关系进行联合建模能有效挖掘到传播结构的细节信息;相比于 EBGCN 模型,MDPAN 模型得到的同一谣言类别间的聚合更为紧密,说明 MDPAN 模型学习到的节点嵌入具有更好的聚类效果,这表明 MDPAN 模型在捕捉节点之间的潜在联系方面具有更好的表现。

为了评估模型对不同谣言类别的识别能力,本节继续在测试集上可视化不同谣言类别的识别情况。由于实验中采用交叉验证的方式随机划分数据集,测试集中的数据可能存在不一致性。为了公平起见,本文固定划分训练集和测试集,重新训练模型并进行评估,然后在此基础上,按照谣言实际类别进行归一化处理,生成了混淆矩阵热力图,如图 8 所示。

可以发现,三种模型都能较准确地预测出谣言数据所属的类别,然而,与MDPAN模型相比,BiGCN和EBGCN模型在识别NR和UR类别上的表现较差,这可能是因为NR和UR在传播结构上存在相似性,而MDPAN模型引入了全局无向传播图,可以更充分地捕捉到这种相似性,这一点在图7中也有所体现。

3.8 案例分析

为了进一步验证MDPAN模型中的节点级注

意力机制对处理不真实的传播关系的有效性,本节从Weibo16数据集中随机选取了一条标签为谣言的数据,其部分传播关系和文本内容如表6所示,通过观察可以发现,转发贴r1仅仅转发了源贴r0,并没有附带文本信息,r1的评论贴r4所附带的评论内容与源贴的内容无关,但是针对r4的评论贴r7却与源贴内容密切相关,不仅如此,源贴r0的一级响应贴如r2、r3也出现了这种情况,这表明谣言原始的传播结构中存在诸多不真实的关系,为了得到稳健的传播结构,有必要对这些隐含关系进行处理。

表6 Weibo16数据集上一条谣言数据的部分传播关系

谣言源贴 (id: 3498960855601223)		评论贴或转发贴		
r0	【特供瀑布】中国最大的瀑布——黄果树瀑布,早就没有多少水了,但只要大领导一来,瀑布还是气势磅礴。原来上面有水库,可以开闸放水。这年头,竟然连风景都有特供!! @袁裕来律师	r1	转发微博	
		r4	Eww.	
			r7	避谣,如果是真的,脑子全进水了。
		r2	风景也特供。	
		r5	这个不可信。枯水期该是上游水库定时放水,丰水期没这个必要。贵州多雨,大旱之年才能看到无水瀑布。不排除领导到来,调整放水时间的可能。但说黄果树没水,绝对瞎扯。不知道不调查,乱放厥词,误导,可耻可恶。(10秒前)	
			让运动与众不同;非一般的感觉。——特步。同理。	
		r3	这个,威武的,我晕啊! 哎	

如图9所示,左侧为表6中的谣言事件的初始传播结构,现有的基于图卷积网络的模型(如BiGCN、VAE-GCN等)将节点r0与其所有邻居节点之间的边都视为是完全可信赖的,在对节点r0进行特征更新时,r0与r3之间潜在的不真实关系会使得传播结构中存在噪声,并且这种噪声会在传播时逐步累积,从而限制模型学习有用的传播特征。本文所提出的MDPAN模型,通过节点级的多头注意力机制,有效降低了不真实的节点关系带来的负面影响。如图9的右侧所示,初始的传播结构在节点级注意力的作用下,动态地生成了一组描述节点真实性关系的权重系数,得到了更为稳定可靠的传播结构。

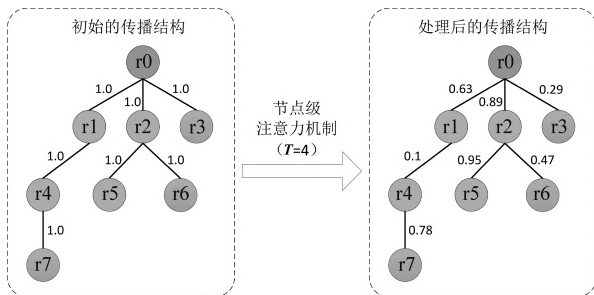


图9 不真实性传播结构的案例分析

4 总结与展望

本文所展开的谣言检测研究主要是面向谣言的传播线索,首次对谣言传统认知上的传播结构从方向上和范围上进行更加细粒度的建模,并采用图卷积网络技术对传播图、扩散图以及全局无向图等不同层次的图结构进行高阶特征学习。同时,采用了一种节点级的注意力机制来对邻居节点及其关系进行考量和评估,从而有效降低不真实节点关系的影响。最后,使用基于注意力机制的多特征池化方法,解决不同层次特征信息之间难以交互的问题。

尽管本文的方法在谣言检测任务中取得了显著效果,但仍然存在有待改进的空间。首先,本文模型对于涉及多个不同社交圈子和传播渠道的谣言可能无法有效检测,因为这些谣言往往具有更为复杂的传播结构,模型难以全面地捕捉到这些谣言的传播规律。其次,社交媒体平台事件的传播可能涉及众多用户,导致构建出的图结构数据非常庞大和复杂,本文模型在处理大规模数据时可能会受到计算资源

的限制,需进一步优化算法以提高效率。

未来的研究工作中,一方面,可以尝试引入外部知识图,从中获取更多的补充信息,以弥补当前部分帖子内容简短的不足;另一方面,可以尝试将传播网络建模成异构的消息传播图,将更多的谣言特征引入到传播结构中来,如用户影响力属性、帖子附带的图片信息、视频信息等。

参考文献

- [1] SUNSTEIN C R. On rumors: How falsehoods spread, why we believe them, and what can be done[M]. New York: Princeton University Press, 2014.
- [2] 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, 2017: 708-717.
- [3] LIN H, ZHANG X, FU X. A graph convolutional encoder and decoder model for rumor detection[C]//Proceedings of the 7th International Conference on Data Science and Advanced Analytics, IEEE, 2020: 300-306.
- [4] 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, PA: ACL, 2018: 1980-1989.
- [5] BIAN T, XIAO X, XU T, et al. Rumor detection on social media with bi-directional graph convolutional networks[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 549-556.
- [6] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[J]. arXiv preprint arXiv: 1609.02907, 2016.
- [7] 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.
- [8] SAIKIA P, GUNDALE K, JAIN A, et al. Modelling social context for fake news detection: A graph neural network based approach[C]//Proceedings of the International Joint Conference on Neural Networks, IEEE, 2022: 01-08.
- [9] CASTILLO C, MENDOZA M, POBLETE B. Information credibility on twitter[C]//Proceedings of the 20th International Conference on World Wide Web, 2011: 675-684.
- [10] 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, 2015: 1751-1754.
- [11] 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, 2016: 3818-3824.
- [12] SHU K, CUI L, WANG S, et al. Defend: Explainable fake news detection[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019: 395-405.
- [13] ZHANG X, CAO J, LI X, et al. Mining dual emotion for fake news detection[C]//Proceedings of the Web Conference, 2021: 3465-3476.
- [14] 葛晓义, 张明书, 魏彬, 等. 基于双重情感感知的可解释谣言检测[J]. 中文信息学报, 2022, 36(09): 129-138.
- [15] MA J, GAO W. Debunking rumors on twitter with tree transformer[C]//Proceedings of the 28th International Conference on Computational Linguistics, 2020: 5455-5466.
- [16] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks [C]//Proceedings of the Semantic Web: 15th International Conference, 2018: 593-607.
- [17] HAMILTON W L, YING R, LESKOVEC J. Inductive representation learning on large graphs[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 1025-1035.
- [18] GUO Z, ZHANG Y, LU W. Attention guided graph convolutional networks for relation extraction [C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019: 241-251.
- [19] CHEN L, FANG J, CHAO P, et al. Rumor detection in social network via influence based on bi-directional graph convolutional network[C]//Proceedings of the International Conference on Web Information Engineering, Springer, Cham, 2022: 274-289.
- [20] CHOI J, KO T, CHOI Y, et al. Dynamic graph convolutional networks with attention mechanism for rumor detection on social media[J]. PLoS ONE, 2021, 16(8): e0256039.
- [21] YUAN C, MA Q, ZHOU W, et al. Jointly embedding the local and global relations of heterogeneous graph for rumor detection[C]//Proceedings of the IEEE International Conference on Data Mining, 2019: 796-805.
- [22] SONG C, TENG Y, ZHU Y, et al. Dynamic graph neural network for fake news detection[J]. Neuro Computing, 2022, 505: 362-374.
- [23] WEI L, HU D, ZHOU W, et al. Towards propaga-

- tion uncertainty: Edge-enhanced Bayesian graph convolutional networks for rumor detection [C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021: 3845-3854.
- [24] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks [J]. arXiv preprint arXiv: 1710.10903, 2017.
- [25] ZEILER M D, FERGUS R. Stochastic pooling for regularization of deep convolutional neural networks [J]. arXiv preprint arXiv: 1301.3557, 2013.
- [26] LIU Y, WU Y F. Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2018: 354-361.
- [27] YAN Y, WANG Y, ZHENG P. A graph-based pivotal semantic mining framework for rumor detection [J]. Engineering Applications of Artificial Intelligence, 2023, 118: 105613.
- [28] SUN M, ZHANG X, ZHENG J, et al. Ddgc: Dual dynamic graph convolutional networks for rumor detection on social media [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36 (4): 4611-4619.



高准(1998—), 硕士研究生, 主要研究领域为自然语言处理。

E-mail: 202108540021056@ctgu.edu.cn



董方敏(1965—), 博士, 教授、博士生导师, 主要研究领域为计算机图形图像处理、智能信息处理。

E-mail: mfdong@ctgu.edu.cn



但志平(1976—), 通信作者, 博士, 教授, 主要研究领域为自然语言处理、计算机视觉、模式识别。

E-mail: zp_dan@ctgu.edu.cn

(上接第 142 页)



彭诗雅(1995—), 硕士研究生, 主要研究领域为自然语言处理。

E-mail: pengshiya_blcu@163.com



于东(1982—), 通信作者, 博士, 副教授, 主要研究领域为自然语言处理。

E-mail: yudong_blcu@126.com



刘畅(1995—), 硕士研究生, 主要研究领域为自然语言处理。

E-mail: liuchang2014@gmail.com