

# 基于自监督异质子图注意力网络的虚假新闻检测<sup>①</sup>



李铭伟, 陈浩鹏, 李风环, 陈 宸

(广东工业大学 计算机学院, 广州 510006)

通信作者: 李风环, E-mail: fhli20180910@gdut.edu.cn

**摘 要:** 由于虚假新闻检测任务的现有工作往往忽略了新闻文本的语义稀疏性及丰富信息之间的潜在联系, 限制了模型对虚假新闻的理解和识别能力, 本文提出了一种基于异质子图注意力网络的虚假新闻检测方法. 针对新闻样本的文本、所属党派、主题等多种信息, 构建了异质图, 以建模虚假新闻的丰富特征. 在特征层采用异质图注意力网络捕获不同类型信息之间的关系, 在样本层引入子图注意力网络挖掘新闻样本间的交互. 同时基于自监督对比学习的互信息机制关注全局图结构中的判别性子图表征, 以捕获新闻样本的特异性. 实验结果表明, 本文提出的方法在 Liar 数据集上相比现有方法在准确率与  $F1$  值分别取得了约 9% 和 12% 的提升, 显著提升了虚假新闻检测的性能.

**关键词:** 虚假新闻检测; 异质图; 子图; 图注意力网络; 自监督对比学习

引用格式: 李铭伟,陈浩鹏,李风环,陈宸.基于自监督异质子图注意力网络的虚假新闻检测.计算机系统应用,2025,34(2):237–245. <http://www.c-s-a.org.cn/1003-3254/9766.html>

## Fake News Detection Based on Self-supervised Heterogeneous Subgraph Attention Network

LI Ming-Wei, CHEN Hao-Peng, LI Feng-Huan, CHEN Chen

(School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China)

**Abstract:** Since existing work on the task of fake news detection frequently ignores the semantic sparsity of news text and the potential relationships between rich information, which limits the model's capacity to understand and recognize fake news, this study proposes a fake news detection method based on heterogeneous subgraph attention networks. Heterogeneous graphs are constructed to model the abundant features of fake news, such as text, party affiliation, and topic of news samples. The heterogeneous graph attention network is constructed at the feature layer to capture the correlations between different types of information, and a subgraph attention network is constructed at the sample layer to mine the interactions between news samples. Moreover, the mutual information mechanism based on self-supervised contrastive learning focuses on discriminative subgraph representations within the global graph structure to capture the specificity of news samples. Experimental results demonstrate that the method proposed in this study achieves about 9% and 12% improvement in accuracy and  $F1$  score, respectively, compared with existing methods on the Liar dataset, which significantly improves the performance of fake news detection.

**Key words:** fake news detection; heterogeneous graph; subgraph; graph attention networks; self-supervised contrastive learning

网络平台与社交媒体已成为新闻传播的主要渠道, 信息获取与分享更加便捷的同时<sup>[1]</sup>, 也出现了虚假新闻

泛滥的现象. 针对这一问题, 学界已开展了大量研究, 主要基于新闻内容、用户评论和新闻传播路径等方面

① 基金项目: 广东省自然科学基金 (2021A1515012290); 广州市黄埔区国际科技合作项目 (2022GH01)

收稿时间: 2024-07-19; 修改时间: 2024-08-13; 采用时间: 2024-08-27; csa 在线出版时间: 2024-12-19

CNKI 网络首发时间: 2024-12-20

进行检测.然而,基于内容的虚假新闻检测往往存在以下几个挑战<sup>[2]</sup>.

(1) 新闻内容短文本通常语义稀疏,缺乏充分的上下文信息,使得准确识别变得困难<sup>[3]</sup>.

(2) 标记的训练数据有限,极大地限制了监督学习方法的性能.如何在有限的标记数据下,提升模型的分类型性能,成为短文本领域亟待解决的关键问题<sup>[4]</sup>.

(3) 需要有效捕捉不同类别信息的重要性,以实现更准确的分类结果.

为了探索短文本虚假新闻检测,本文对虚假新闻数据集 Liar<sup>[5]</sup>进行数据分析.发现政治立场、新闻主题都会对新闻真实性产生显著影响.例如,表1中某些政党发布的虚假新闻比例和表2中特定主题下的虚假新闻概率明显偏高.这些发现为虚假新闻检测研究提供了重要线索.

表1 不同党派的可信度分布情况表(%)

| 党派    | 真实    | 几乎真实  | 半真    | 勉强为真  | 假的    | 明显的谎言 |
|-------|-------|-------|-------|-------|-------|-------|
| 无党派人士 | 19.70 | 35.00 | 18.30 | 12.90 | 12.20 | 2.00  |
| 共产党人士 | 10.90 | 11.00 | 26.90 | 20.50 | 22.80 | 7.70  |

表2 不同主题的可信度分布情况表(%)

| 主题   | 真实    | 几乎真实  | 半真    | 勉强为真  | 假的    | 明显的谎言 |
|------|-------|-------|-------|-------|-------|-------|
| 历史主题 | 13.40 | 9.00  | 8.90  | 20.90 | 22.40 | 25.40 |
| 教育主题 | 15.80 | 24.00 | 22.50 | 15.80 | 14.20 | 7.20  |

基于以上分析,本文提出了一种基于异质子图注意力网络的虚假新闻检测方法.该方法构建了一个能够捕获新闻数据复杂特征的异质图模型.在特征层采用异质图注意力网络捕获不同类型信息之间的关系,将新闻样本建模为子图,在样本层引入子图注意力网络挖掘新闻样本间的交互.并基于自监督对比学习的互信息机制关注全局图结构中的判别性子图表征,以捕获新闻样本的特异性.实验结果表明,本文提出的方法在 Liar 数据集上取得了优异的结果,显著提升了虚假新闻检测的性能.

## 1 相关工作

### 1.1 基于内容的虚假新闻检测

基于内容的虚假新闻检测方法主要依赖于文本内容的分析.这些方法通过提取文本中的关键词、主题和情感<sup>[6]</sup>等特征,结合机器学习或深度学习算法,判别新闻的真实性.传统的基于内容的方法在处理复杂的语言结构和语义关系时存在局限性.基于深度学习的

方法通常利用循环神经网络(RNN)、卷积神经网络(CNN)等模型捕捉文本的语义信息和上下文关系.除此之外,Transformer 模型如 BERT<sup>[7]</sup>等通过自注意力机制,能够更好地捕捉文本的全局依赖关系.然而,基于内容的虚假新闻检测方法在面对短文本数据集时通常表现不佳,缺乏足够的信息使得数据存在语义稀疏性.

### 1.2 异质图神经网络

异质图神经网络是一种能够处理具有多种类型节点和边的图结构数据的神经网络<sup>[8]</sup>.异质图神经网络在虚假新闻检测中的应用具有显著优势.例如 Xie 等人<sup>[9]</sup>提出了一种通过多元信息进行虚假新闻检测的新型异质图神经网络;以及 Kang 等人<sup>[10]</sup>创新性地利用时间和来源等方面的相关性来检测虚假新闻.然而,这些方法专注于单个文本信息的挖掘,忽视了新闻样本之间的联系.

### 1.3 子图神经网络

子图神经网络的核心思想是从原始图中采样出较小的子图,减少整个图结构的计算复杂度和内存消耗,同时保持对原始图结构的近似表示<sup>[11]</sup>.在虚假新闻检测中,可以通过构建包含关键节点和边的子图来捕捉与虚假新闻相关的重要信息.例如, Yang 等人<sup>[12]</sup>提出了一种增强的子图生成方法,通过去除与任务无关的信息来提高模型的泛化和辨别能力; Zhang 等人<sup>[13]</sup>设计了一种子图生成策略,通过限制这些子图的删除,以有效地生成与原图不同的子图,作为谣言检测的反事实证据.然而所有的方法都集中于利用子图来缓解全图结构的复杂性,没有通过子图和全图的关联捕获样本的重要信息和样本之间的交互.

## 2 本文方法

### 2.1 方法概述

本文方法由两大核心模块组成,即异质图注意力网络与子图注意力网络,如图1所示.首先,将和新闻样本相关的特征数据建模为异质图,利用异质图注意力网络捕获丰富的特征及其复杂的关系.其次,将新闻样本的特征数据建模为子图,将子图视为超节点,并构建子图注意力网络建模新闻样本之间的交互,例如新闻样本之间的相关性、相似性等.为了进一步增强子图在全局图中的判别性表征能力,采用了自监督互信息机制,以考虑子图与全局图之间的交互,以捕获新闻样本表示的特异性.

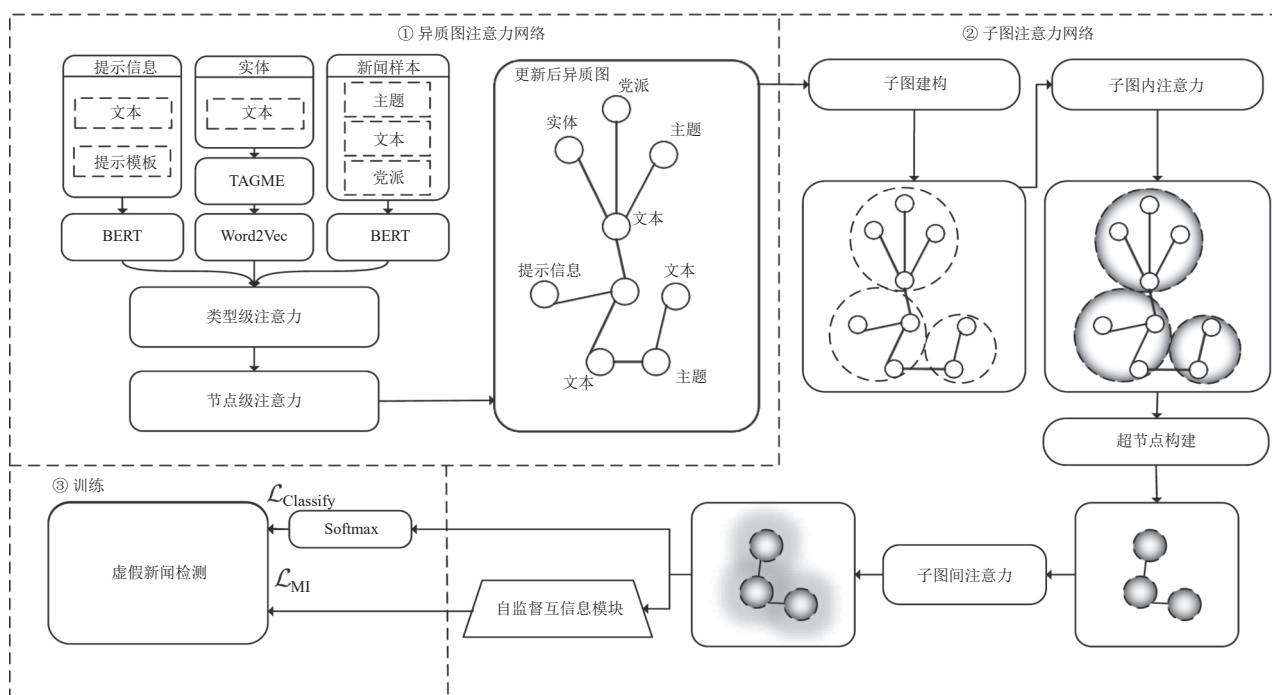


图1 模型整体框架图

## 2.2 异质图注意力网络

首先, 本文利用新闻文本与特征数据之间的复杂关系, 构建异质图。其次, 本文采用半监督学习的方法, 对模型进行训练。在这一过程中, 本文还利用有限的标注数据来指导模型的学习, 并通过图传播机制, 将学习到的知识有效地传递给未标注数据, 以学习到分类的内在规律并进一步扩展分类能力, 从而提高分类的准确性。

### 2.2.1 异质图构建

为了有效应对短文本的稀疏性问题, 即短文本中包含的信息和语义相对较少, 缺乏丰富的上下文信息的问题, 本文构建了一个灵活的异质图模型, 专门用于对短文本进行建模。

对新闻文本预处理后, 本文利用 BERT 模型提取文本特征向量。针对主题信息和党派信息, 借助 BERT 模型, 进一步提取主题和党派特征向量, 以捕捉其深层次的语义信息, 并为每个文本与其所属的主题和党派之间建立一条边, 从而构建出完整的异质图模型。

在处理实体信息时, 利用 TAGME 实体链接工具, 将识别出的实体映射到维基百科中相应的条目, 并将文本与识别出的实体之间建立边。本文视每个实体为一个完整的单词, 并利用基于维基百科语料库的 Word2Vec 模型来学习实体的嵌入表示。

为了进一步增强短文本的语义丰富度并促进信息的有效传播, 本文特别关注实体和主题之间的关系。若两个实体和主题之间的相似度得分 (采用余弦相似度作为度量标准) 高于预设的阈值  $\delta$  时, 才会在它们之间建立一条边, 以捕捉它们之间的潜在联系。

在提示学习方面<sup>[14]</sup>, 本文设计了一个提示模板, 设计形式为: “新闻文本+Here is a piece of news with [mask] information.”, 其中 [mask] 代表掩码标记。结合 BERT 模型的文本预测能力, 本文将原始的文本分类任务转化为文本类别预测任务, 即预测特定的 [mask] 位置为 False 和 True 的概率。最后, 将文本节点和提示信息节点之间建立边。

本文将图表示为  $G = (V, \mathcal{E})$ , 其中  $D = d_1, \dots, d_n$  表示短文本节点集合,  $T = t_1, \dots, t_k$  表示主题节点集合,  $E = e_1, \dots, e_m$  表示实体节点集合,  $P = p_1, \dots, p_l$  表示党派节点集合,  $F = f_1, \dots, f_x$  表示提示学习所得到的提示节点集合, 因此  $V = D \cup T \cup E \cup P \cup F$ , 而  $\mathcal{E}$  为所有节点之间的边集, 并且本文引入异质图  $G$  的邻接矩阵  $A$ , 并在此基础上加上自连接矩阵  $I$ , 即  $A' = A + I$ 。

综上所述, 本文考虑了 4 种类型的附加信息: 主题、党派、实体和提示信息。全图的实例如图 2(a) 所示。通过整合文本与所有这些附加信息之间的关系, 极大地丰富了短文本的语义内容, 从而为后续的分类任

务提供了充足信息. 例如, 如表 3 所示, 文本中提及的“明尼苏达州”和“共和党”通过与实体的关联得到了语义上的丰富, 同时结合主题“经济”和“历史”的信息. 因

此, 考虑到这些因素以及与其相似的新闻较高的虚假新闻概率, 该文本被正确分类为虚假新闻的可能性大大增加.

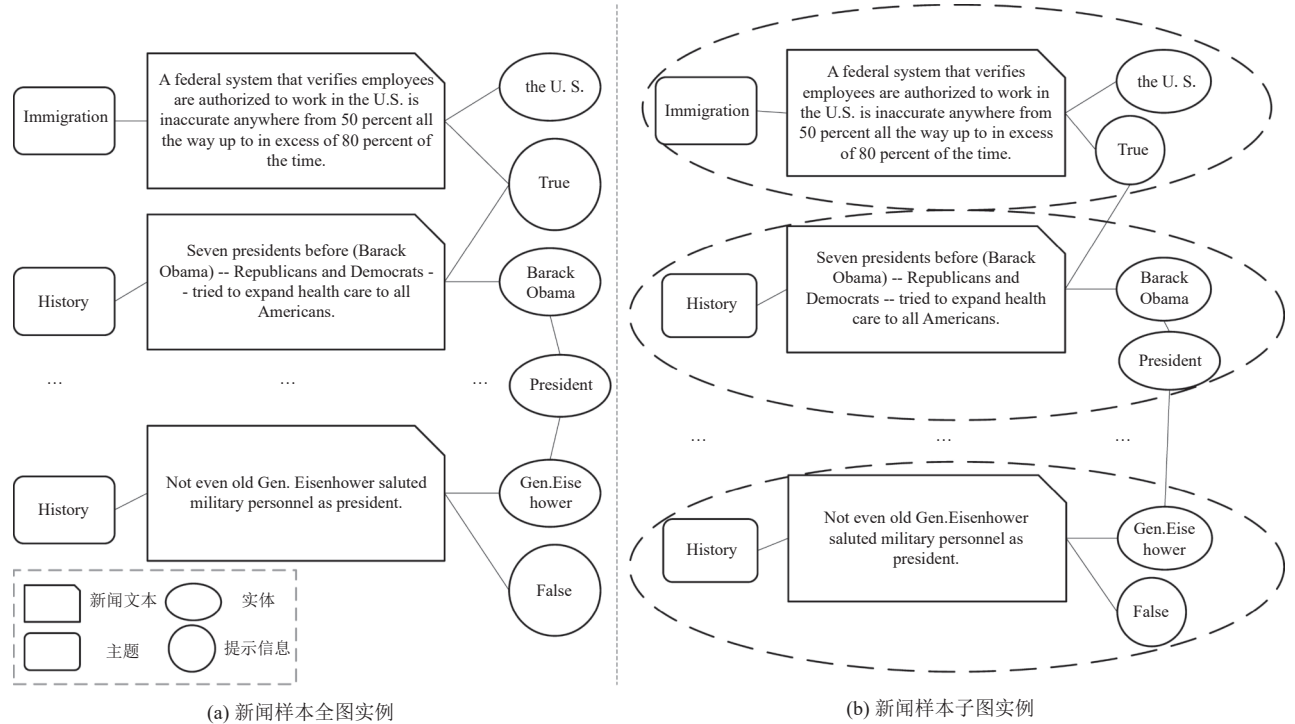


图2 全图与子图的实例

表3 Liar 中的一条随机数据

| 属性            | 内容                                                                                                                                                                                                                                                      |
|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Text          | Minnesotas advantage in economic growth has been having Republicans in charge of at least one part of government for all but two years, while Wisconsin Democrats for many years before 2011 controlled both legislative chambers and the governorship. |
| Subject       | economy, history                                                                                                                                                                                                                                        |
| Party         | republican                                                                                                                                                                                                                                              |
| Dataset label | False                                                                                                                                                                                                                                                   |

2.2.2 双层注意力机制

各节点受到不同类型相邻节点的影响概率是不尽相同的. 为了准确地捕捉这种节点级和类型级的重要性差异, 本文使用了一种双层注意力机制, 以精确地考虑不同节点和类型在信息传播过程中的作用.

类型级注意力: 类型级注意力机制能够捕捉相邻节点的不同类型对当前节点的重要性, 融合了多个相同类别节点所贡献的综合影响. 本文首先选定特定的节点 $v$ . 接着, 根据其邻居节点的类型信息, 计算类别向量, 公式如下:

$$h_{\tau} = \sum_{v'} A'_{vv'} h_{v'} \quad (1)$$

其中, 节点 $v'$ 为全图中所属类型 $\tau$ 的 $v$ 的邻居节点,  $h_{v'}$ 为节点 $v'$ 的特征向量. 该向量不仅反映了节点自身的特征, 还融入了与该节点类型相同的相邻节点对其的影响. 因此, 本文可以计算出基于当前节点特征向量 $h_v$ 和类型向量 $h_{\tau}$ 来计算类型级注意力得分:

$$a_{\tau} = \sigma(\mu_{\tau}^T \cdot [h_v || h_{\tau}]) \quad (2)$$

其中,  $\mu_{\tau}^T$ 是类型 $\tau$ 的可训练注意力向量,  $||$ 表示向量拼接,  $\sigma(\cdot)$ 表示激活函数. 最后, 本文使用 Softmax 函数对所有类型的注意力得分归一化以获得类型级别的注意力权重:

$$\alpha_{\tau} = \frac{\exp(a_{\tau})}{\sum_{\tau' \in T} \exp(a_{\tau'})} \quad (3)$$

节点级注意力: 节点级注意力能够捕捉相邻节点对当前节点的重要性, 并降低噪声节点的权重, 提升信息提取的准确性. 具体而言, 本文假设特定节点 $v$ , 其类别为 $\tau$ 及其相邻不同类型节点 $v'$ , 类型为 $\tau'$ . 接着, 结合

节点的特征向量 $h_v$ 和 $h_{v'}$ 以及 $\tau'$ 类型的注意力权重 $\alpha_{\tau'}$ , 计算节点级注意力系数, 公式如下:

$$b_{vv'} = \sigma(W^T \cdot \alpha_{\tau'} [h_v \| h_{v'}]) \quad (4)$$

其中,  $W^T$ 是注意力向量, 该系数能够准确反映不同节点在整体网络结构中的重要性. 然后, 本文使用 Softmax 函数对节点级注意力得分进行归一化, 其中  $N_v$  是节点  $v$  的邻居节点集合.

$$\beta_{vv'} = \frac{\exp(b_{vv'})}{\sum_{i \in N_v} \exp(b_{vi})} \quad (5)$$

本文通过双层注意力机制, 得到了更为合理且精准的节点表示, 具体的逐层传播规则如下:

$$H^{(l+1)} = \alpha \left( \sum_{\tau \in \Gamma} B_{\tau} \cdot H_{\tau}^{(l)} \cdot W_{\tau}^{(l)} \right) \quad (6)$$

其中,  $B_{\tau}$ 表示注意力矩阵, 第 $v$ 行第 $v'$ 列的元素即是 $\beta_{vv'}$ , 代表了特定的注意力权重,  $\Gamma$ 是节点类型的集合.

### 2.3 子图注意力网络

子图注意力网络从3个层次进行考量: 节点、子图内部以及子图之间. 首先, 从原始图中构建子图及子图表示, 旨在充分提取出能够代表新闻样本特征的特征. 其次, 本文利用子图内注意力机制将子图转换为超节点, 利用子图间注意力机制建模新闻样本间的关系, 并利用自监督互信息机制, 进一步挖掘样本子图和样本集合全图的交互, 以获得判别性的样本表示.

#### 2.3.1 子图内注意力

本文首先从原始图中提取每个新闻样本的子图. 实例如图2(b)所示. 针对选定的新闻文本中心节点 $v$ , 选取与中心节点相关联的附加信息节点构建子图, 从而提取出一个子图 $G_v = \{v_1, v_2, \dots, v_s\}$ , 其中 $s$ 为新闻样本的特征节点个数. 重复即可得到全部子图 $P = \{G_1, G_2, \dots, G_n\}$ , 其中 $n$ 为新闻样本的个数.

为了将节点表示统一整合成为子图表示, 本文引入了子图内注意力机制. 注意力系数 $c_j^{(i)}$ 表示节点 $v_j$ 对于子图 $G_i$ 重要性, 计算公式如下:

$$c_j^{(i)} = \sigma(a_{\text{intra}}^T W_{\text{intra}} h_j^{(i)}) \quad (7)$$

其中,  $W_{\text{intra}} \in R^{d_1 \times d_1}$ 是一个权重矩阵,  $a_{\text{intra}} \in R^{d_1}$ 是一个权重向量,  $h_j^{(i)}$ 为子图 $G_i$ 的节点 $v_j$ 的特征表示, 之后本文通过 Softmax 函数归一化子图中节点之间的注意力系数. 最后, 计算表示子图 $G_i$ 的特征向量 $z_i$ , 公式如下:

$$z_i = \sum_{v_j \in G_i} c_j^{(i)} h_j^{(i)} \quad (8)$$

#### 2.3.2 子图间注意力

本文将子图 $G_i$ 视为新闻样本的超节点 $g_i$ , 当两个超节点之间的公共节点数量超过预设的阈值 $b_{\text{com}}$ 时, 本文就在它们之间建立一条边 $e_{i,j}$ . 最终, 通过这种方式, 本文将原始图重构为了一个超图 $G' = (V', E')$ , 具体公式如下:

$$V' = \{g_i\}; E' = \{e_{i,j}\}, \forall |V(g_i) \cap V(g_j)| > b_{\text{com}} \quad (9)$$

引入子图间注意机制学习子图间的影响. 通过多头注意力机制计算注意力系数 $\alpha_{ij}$ , 子图特征向量 $z_i'$ 的计算公式为:

$$z_i' = \frac{1}{M} \sum_{m=1}^M \sum_{e_{ij} \in E'} \alpha_{ij}^m W_{\text{inter}}^m z_i \quad (10)$$

其中,  $W_{\text{inter}} \in R^{d_2 \times d_1}$ 是权重矩阵, 以及 $M$ 是多头注意力的数量. 获得子图特征向量后, 本文进一步通过 Softmax 函数转化为标签的预测值.

#### 2.3.3 自监督对比学习

本文利用互信息 (MI) 来评估所获取的子图和全图的交互关系, 以自监督对比学习的方式, 增强子图在全图中的判别性表征.

首先利用 READOUT 函数将各个子图向量汇总为固定长度的向量, 即全图表示 $r$ .

结合互信息判别器 $D$ , 判断一个子图/图表示是否来自同一个图, 公式为:

$$D(z_i', r) = \sigma(z_i' W_{\text{MI}} r) \quad (11)$$

其中,  $W_{\text{MI}}$ 是一个评分矩阵, 并且 $\sigma(\cdot)$ 是 Sigmoid 函数.

在自监督互信息机制 (self-supervised mutual information, SMI) 中, 对全图 $G$ 进行扰动, 获得全图 $\tilde{G}$ , 负样本是通过将来自全图 $\tilde{G}$ 中的子图表示 $z'$ 与来自图 $G$ 的表示进行配对来实现的. 最后, 自监督学习的目标可以定义为标准二进制交叉熵 (BCE) 损失.

$$L_{\text{MI}}^G = \frac{1}{n + n_{\text{neg}}} \left( \sum_{g_i \in G} E_{\text{pos}} [\log(D(z_i', r))] + \sum_{g_j \in \tilde{G}} E_{\text{neg}} [\log(1 - D(z_j', r))] \right) \quad (12)$$

其中,  $n$ 表示正样本的数量,  $n_{\text{neg}}$ 表示负样本的数量,  $L_{\text{MI}}^G$ 相当于最大化 $z'$ 和 $r$ 之间的互信息.

### 2.4 训练

对于分类损失 $L_{\text{Classify}}$ , 本文使用交叉熵损失函数

计算,其公式为:

$$L_{\text{Classify}} = -\frac{1}{n} \sum_{i=1}^n (y_i \cdot \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) \quad (13)$$

其中,  $y$  是真实值,  $\hat{y}$  是预测值,  $n$  为样本个数.

本文将分类损失  $L_{\text{Classify}}$  和 MI 损失  $L_{\text{MI}}^G$  结合起来. 因此, 本文可以得到 SANSMI (subgraph attention network with self-supervised mutual information) 的损失  $L$  定义如下:

$$L = L_{\text{Classify}} + \beta L_{\text{MI}}^G + \lambda \|\Theta\|^2 \quad (14)$$

其中,  $\beta$  是自监督 MI 的系数,  $\lambda$  是  $\Theta$  上  $L2$  正则化的系数,  $\Theta$  是该框架中的一组可训练参数.

### 3 实验与分析

#### 3.1 训练集

在这项研究中, 本文使用了 Liar 数据集. 表 4 展示了它的详细统计数据.

表 4 Liar 数据集统计

| 参数名            | 值     |
|----------------|-------|
| 训练集规模          | 10269 |
| 验证集规模          | 1284  |
| 测试集规模          | 1283  |
| 主题种类数          | 144   |
| 党派种类数          | 25    |
| 平均新闻长度 (token) | 17.9  |

Liar 数据集的数据根据真实性被赋予了 6 个不同的评级标签, 为简化新闻的二元分类任务, 本文将其归纳为两个主要类别: False 和 True. 经过转换后, 数据集包含了 56% 的真实陈述和 44% 的虚假陈述.

#### 3.2 实验设置

为了构建 Liar 数据集的训练集和测试集, 从每个主题中分别选取了 80% 的数据作为训练集, 剩余的 20% 则作为测试集.

参数的调整将对结果有较大影响, 实验参数设置如表 5 所示.

#### 3.3 基线方法

在这项研究中, 本文复现并评估了 5 个用于虚假新闻检测的深度学习模型<sup>[15]</sup>, 下面描述了这些模型.

(1) HAN (2016): HAN 是一种分层注意力网络模型, 通过单词级别与句子级别的编码来表示文本.

(2) RoBERTa (2019): RoBERTa 是一种基于 BERT

的预训练语言模型, 在多种文本任务上取得出色的性能.

(3) ELMo (2018): ELMo 是一种基于 LSTM 的预训练语言模型, 能捕捉丰富的文本语义信息.

(4) CompareNet (2021)<sup>[16]</sup>: 基于异质图注意力网络的模型, 将新闻文本与外部知识通过实体进行比较, 以捕捉更丰富新闻特征.

(5) HGNNR4FD (2023)<sup>[9]</sup>: 基于知识图谱的异质图注意力网络模型, 将异质图与知识图谱结合, 进一步获取虚假新闻特征.

表 5 各个模块的实验参数

| 模块          | 参数名                     | 值     |
|-------------|-------------------------|-------|
| BERT        | attention_dropout_prob  | 0.1   |
|             | hidden_act              | GELU  |
|             | hidden_dropout_prob     | 0.1   |
|             | hidden_size             | 768   |
|             | initializer_range       | 0.02  |
|             | intermediate_size       | 3072  |
|             | max_position_embeddings | 512   |
|             | num_attention_heads     | 12    |
|             | num_hidden_layers       | 12    |
|             | type_vocab_size         | 2     |
| 异质图注意力网络    | vocab_size              | 30522 |
|             | hidden_size             | 512   |
|             | input_size              | 512   |
|             | output_size             | 512   |
|             | learn_rate              | 0.005 |
|             | dropout                 | 0.8   |
|             | regularization factor   | 0.5   |
|             | num_info                | 1     |
|             | lr                      | 0.01  |
|             | max_pool                | 0.06  |
| 子图<br>注意力网络 | momentum                | 0.8   |
|             | num_epoch               | 1000  |
|             | batch_size              | 64    |
|             | sg_encoder              | GAT   |
|             | MI_loss                 | 0.8   |
|             | start_k                 | 0.8   |

#### 3.4 性能比较

在表 6 中, 本文详细比较了不同方法在 Liar 数据集上的性能. 与所有基线方法相比, 本文的方法展现出了显著的优越性. 在 Liar 数据集上, 本文的方法实现了 71% 的准确率, 这一显著的效果主要由于本文模型中的异质图注意力网络和子图注意力网络. 为了进一步分析本文方法训练过程, 由图 3 可以观察到, 本文的优化策略能够有效训练模型, 最终得到较高的性能. 相较于基于 LSTM 的基线模型, 如 ELMo, 图注意机制使

本文的模型能够更深入地理解文本中的复杂结构,并在处理文本中的噪声信息时更为有效,从而保留关键特征.与预训练模型相比,本文的模型在 Liar 数据集上的平均准确率提升了 10%.这一提升可能在于相比较预训练模型,本文方法可以利用子图注意力网络,更有利于捕捉图数据中细粒度语义信息,并利用自监督机制,更好的捕获样本的判别性表征,从而在分类任务中表现出更高的准确性.与异质图注意力网络模型比,如 HGNNR4FD、CompareNet,本文模型在 Liar 数据集上性能分别提升了约 9% 与 12%,这一提升可能在于本文模型在异质图注意力网络基础上,利用子图注意力网络,融合子图局部表征,进一步捕获特征,并突出子图间的关联.

表 6 所有模型在 Liar 数据集上的结果

| 方法         | Acc         | P           | R           | F1          |
|------------|-------------|-------------|-------------|-------------|
| HAN        | 0.58        | 0.55        | 0.49        | 0.52        |
| RoBERTa    | 0.62        | 0.58        | 0.63        | 0.60        |
| ELMo       | 0.58        | 0.51        | 0.63        | 0.56        |
| CompareNet | 0.59        | 0.56        | 0.57        | 0.56        |
| HGNNR4FD   | 0.62        | 0.61        | 0.57        | 0.59        |
| Ours       | <b>0.71</b> | <b>0.69</b> | <b>0.75</b> | <b>0.72</b> |

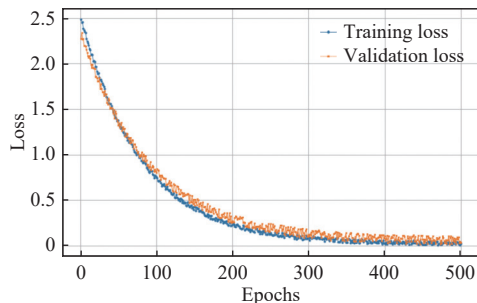


图 3 损失曲线图

### 3.5 消融实验

#### 3.5.1 模块

为了探究预训练模型 BERT、异质图注意力网络 (HGAT) 和子图注意力网络对实验结果的具体影响,设计了如表 7 所示的实验.通过观察实验结果分析,完整的模型在性能上超越了所有其他变种.这充分说明了,在缺少 BERT (w/o BERT) 的情况下,模型无法在处理新闻文本时获得更准确地文本的含义和意图;而当缺少异质图注意力网络 (w/o HGAT) 时,模型无法充分利用与之相关的特征数据,难以捕捉文本中的主题变化以及实体间的关系;而当缺少子图注意力网络 (w/o SANSMI) 时,模型没有考虑文本之间的复杂结构;而

当同时缺少了异质图注意力网络与子图注意力网络 (w/o HGAT + SANSMI) 时,模型无法考虑新闻文本及附属信息之间的潜在关联和新闻样本间的复杂联系.因此实验结果表明这些信息对于分类新闻内容具有重要意义和在模型中的不可或缺性.

为了验证自监督 MI 机制的有效性,本文将完整模型 (ours) 与除去自监督 MI 机制的子图注意力网络 (w/o SMI) 进行了对比实验.在子图注意力网络中,移除了自监督 MI 机制模块,以观察这一变化对模型性能的影响.实验结果如表 7 所示,可以看到在保留自监督 MI 机制的情况下表现更为出色,这充分证明了自监督 MI 机制在提升模型性能方面的有效性.

表 7 模型在去除部分模块的效果

| 模型              | Acc         | P           | R           | F1          |
|-----------------|-------------|-------------|-------------|-------------|
| w/o HGAT+SANSMI | 0.58        | 0.57        | 0.60        | 0.58        |
| w/o BERT        | 0.61        | 0.58        | 0.62        | 0.60        |
| w/o SANSMI      | 0.63        | 0.58        | 0.57        | 0.57        |
| w/o HGAT        | 0.65        | 0.64        | 0.63        | 0.63        |
| w/o SMI         | 0.65        | 0.64        | 0.65        | 0.64        |
| Ours            | <b>0.71</b> | <b>0.69</b> | <b>0.75</b> | <b>0.72</b> |

#### 3.5.2 特征

为了分析不同新闻特征信息对实验结果的影响,本文采用了 4 种不同的附加信息组合,包括主题+党派、主题、党派以及无.通过分析如图 4 所示的实验结果,本文发现党派数据对实验结果的提升效果最为显著,这表明了党派数据在虚假新闻预测中的重要性超过其他新闻附属数据.

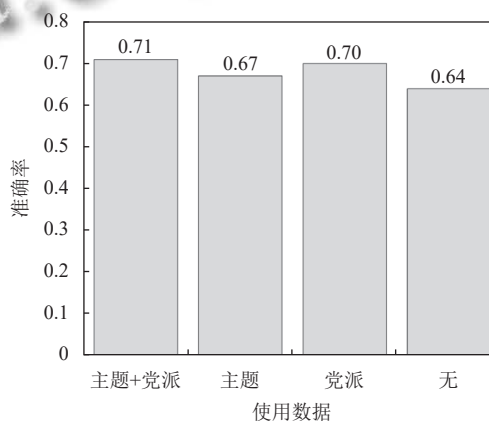


图 4 模型采用不同附加数据的准确率

### 3.6 参数分析

在自监督 MI 模块中尝试使用了不同的损失函数系数,该系数表明自监督损失函数在总损失函数中的

比重,在训练过程中起到平衡的作用.如图5所示,发现当MI损失函数系数设置为0.8时,自监督MI模块能够有效增强特征表示,性能表现最为出色;当系数为0.2时,自监督MI模块明显影响不足,无法显著提升性能;当系数为1时,自监督MI模块明显影响过大,导致性能略微下降.

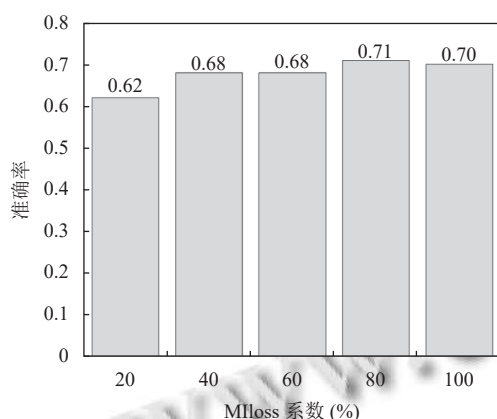


图5 不同MI损失函数系数的准确率

### 3.7 可视化

为了探讨异质图注意力网络在融合附加信息方面的能力,本文展示了异质图注意力网络中新闻文本节点及其相邻节点的注意力系数分布,如图6所示.这些节点按照编号依次为:文本、党派、主题、实体以及提示信息.通过观察这些节点的注意力系数,可以发现,在附加信息中,党派和主题节点的权重明显高于实体和提示信息节点的权重,这意味着党派和主题信息对最终的文本特征向量产生了较大的影响.

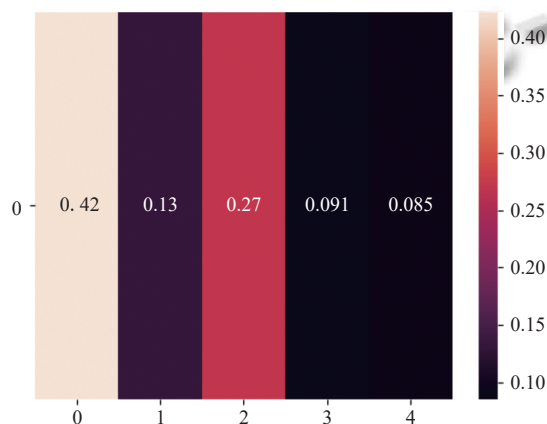


图6 异质图注意力网络注意力系数

为了探讨样本间的交互作用,展示了子图注意力网络中各子图之间的注意力系数分布,如图7所示,表

明横坐标子图对纵坐标子图的影响.子图注意力系数代表着信息传递的影响力,以及子图间的相关性和重要性,表明样本之间的交互.通过观察可以发现,子图间的影响是不同的,相关性越高,子图注意力系数也越高,子图间消息传递的影响也越大,对最终分类结果的影响也越大.

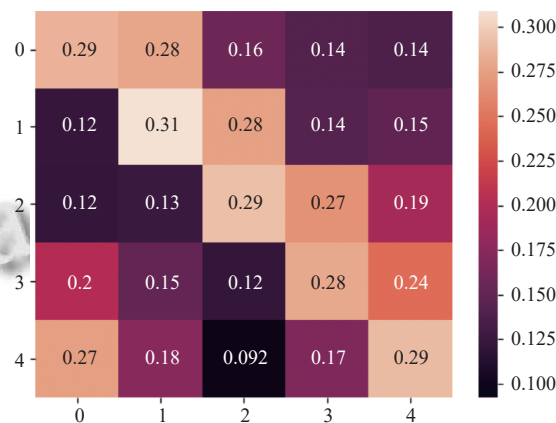


图7 子图注意力网络注意力系数

## 4 结论与展望

本文通过构建能够全面反映新闻数据复杂性的异质图模型,并引入子图注意力网络与自监督互信息机制,有效提升了模型对新闻间复杂关系的理解和分析能力,从而取得了较佳的短文本虚假新闻检测性能.该方法揭示了虚假新闻检测中需要同时考虑内容、传播和关联数据等多方面因素的复杂性.在有限标注数据情况下,利用自监督学习增强模型对潜在语义关系的理解,是提升短文本虚假新闻检测性能的有效策略.此外,子图注意力网络为灵活捕捉不同粒度信息重要性提供了新的框架.然而,随着数据规模和数据类型的不断丰富,子图注意力网络部分的计算复杂度也随之增加,这需要进一步优化算法.未来可尝试轻量级子图表示学习方法,或探索基于强化学习的自适应子图构建策略,以提高模型的泛化能力和实用性.

### 参考文献

- 1 Kim SJ. The role of social media news usage and platforms in civic and political engagement: Focusing on types of usage and platforms. *Computers in Human Behavior*, 2023, 138: 107475. [doi: [10.1016/j.chb.2022.107475](https://doi.org/10.1016/j.chb.2022.107475)]
- 2 Liao Q, Chai HY, Han H, et al. An integrated multi-task

- model for fake news detection. *IEEE Transactions on Knowledge and Data Engineering*, 2022, 34(11): 5154–5165. [doi: [10.1109/TKDE.2021.3054993](https://doi.org/10.1109/TKDE.2021.3054993)]
- 3 Liu YH, Huang L, Giunchiglia F, *et al.* Improved graph contrastive learning for short text classification. *Proceedings of the 2024 AAAI Conference on Artificial Intelligence*. Vancouver: AAAI, 2024. 18716–18724.
- 4 Zhang BJ, He Q, Zhang DM. Heterogeneous graph neural network for short text classification. *Applied Sciences*, 2022, 12(17): 8711. [doi: [10.3390/app12178711](https://doi.org/10.3390/app12178711)]
- 5 Wang WY. “Liar, liar pants on fire”: A new benchmark dataset for fake news detection. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. Vancouver: ACL, 2017. 422–426.
- 6 Kolev V, Weiss G, Spanakis G. FOREAL: RoBERTA model for fake news detection based on emotions. *Proceedings of the 14th International Conference on Agents and Artificial Intelligence*. Setúbal: Scitepress—Science and Technology Publications, 2022. 429–440.
- 7 Zeng Z, Ye LY, Liu RG, *et al.* Fake news detection by using common latent semantics matching method. *Proceedings of the 33rd IEEE International Conference on Tools with Artificial Intelligence*. Washington: IEEE, 2021. 1059–1066.
- 8 Yang TC, Hu LM, Shi C, *et al.* HGAT: Heterogeneous graph attention networks for semi-supervised short text classification. *ACM Transactions on Information Systems*, 2021, 39(3): 32. [doi: [10.1145/3450352](https://doi.org/10.1145/3450352)]
- 9 Xie BB, Ma XX, Wu J, *et al.* Heterogeneous graph neural network via knowledge relations for fake news detection. *Proceedings of the 35th International Conference on Scientific and Statistical Database Management*. Los Angeles: ACM, 2023. 1–11.
- 10 Kang ZZ, Cao YN, Shang YM, *et al.* Fake news detection with heterogenous deep graph convolutional network. *Proceedings of the 25th Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Cham: Springer, 2021. 408–420.
- 11 Sun QY, Li JX, Peng H, *et al.* SUGAR: Subgraph neural network with reinforcement pooling and self-supervised mutual information mechanism. *Proceedings of the 2021 Web Conference*. New York: ACM, 2021. 2081–2091.
- 12 Yang RC, Wang XT, Jin YQ, *et al.* Reinforcement subgraph reasoning for fake news detection. *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2022. 2253–2262.
- 13 Zhang KW, Yu JC, Shi HC, *et al.* Rumor detection with diverse counterfactual evidence. *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2023. 3321–3331.
- 14 Jiang GY, Liu S, Zhao Y, *et al.* Fake news detection via knowledgeable prompt learning. *Information Processing & Management*, 2022, 59(5): 103029. [doi: [10.1016/j.ipm.2022.103029](https://doi.org/10.1016/j.ipm.2022.103029)]
- 15 Khan JY, Khondaker MTI, Afroz S, *et al.* A benchmark study of machine learning models for online fake news detection. *Machine Learning with Applications*, 2021, 4: 100032. [doi: [10.1016/j.mlwa.2021.100032](https://doi.org/10.1016/j.mlwa.2021.100032)]
- 16 Hu LM, Yang TC, Zhang LH, *et al.* Compare to the knowledge: graph neural fake news detection with external knowledge. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*. ACL, 2021. 754–763.

(校对责编: 张重毅)