

2023年度大创结项答辩



# 社交媒体机器人识别

**BOT DETECTIVE**

项目成员：陈思远 梁松松 徐敬 施文韬

指导老师：李丕绩



# 研究背景及应用

# 社交媒体机器人识别

## 赵蓓：社交机器人的负面影响应引起重视



首页 关于我们 本社报刊 本社新闻 投稿 报

首页 > 中国社会科学报 > 科学与人文

### 社交机器人：风险隐忧与技术治理

2022年08月16日 09:48 来源：《中国社会科学报》2022年8月16日第2471期 作者：王璐瑶



AMHotFlash @AMHotFlash · 2019年5月13日

ATTENTION: TEXAS FARMERS: Trump is hurting us more than you know off

## 马晓晨 | 人工智能与社交网络在俄乌冲突中的应用

人工智能应用于服务、教育、医疗等行业，在许多新兴技术如计算机图形以及各类虚拟现实工具和触觉系统，也对社

### 社交机器人在舆论斗争中的应用研究

作者：龙亚蒙 周洋

是由人工智能控制的社交媒体账号，当前已经成为影响公共话语和舆论斗争的特点入手，重点分析了社交机器人在政治斗争、反恐斗争等方面提出思考与建议。

2020-03-17 15:35

### 在Twitter参与中美贸易战的社交机器人

国际新闻界 2020.05

社交机器人操纵Twitter上的中国议题研究

## 分布与互动模式：社交机器人操纵Twitter上的中国议题研究

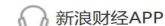
## 机器人带节奏?! 美国抹黑中国病毒溯源言论都是编的



共青团中央 | + 关注

## 被算法“操控”的美国大选

2020年10月15日 12:43 创业邦



新浪财经APP



A



A



加入书架，改变既有的

原标题：被算法“操控”的美国大选



## 张洪忠等：政治机器人在社交媒体空间的舆论干预分析

题图：英国《卫报》调查记者卡罗尔TED演讲

# 研究背景及应用

# 社交媒体机器人识别

## 问题定义

### 赵蓓：社交机器人的负面影响应引起重视

国际新闻界 2020.05

社交机器人操纵Twitter上的中国议题研究

### 分布与互动模式：社交机器人操纵Twitter上的中国议题研究

中国社会科学杂志社  
SOCIAL SCIENCES IN CHINA PRESS

首页 关于我们 本社报刊 本社新闻 投稿 报

首页 > 中国社会科学报 > 科学与人文

### 机器人带节奏？都是编的



# HUMAN

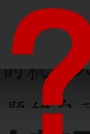


# 用户

### 社交机器人操纵舆论战

2022年08月16日 09:48 来源：《中国社会科学报》

AMHotFlash @AMHotFlash  
ATTENTION: TEXAS FARMERS: Trump is hurting us



# BOT

### 马晓晨 | 人工智能与社交网络在俄乌冲突中的应用

### 社交机器人在舆论斗争中的应用研究

作者：龙亚蒙 周

Given  $G = \{V, E, A, R\}$

是由人工智能控制的社交媒体账号，当前已经成为影响公共话语和舆论斗争的特点入手，重点分析了社交机器人在政治斗争、反恐斗争等实际提出思考与建议。

### 张洪忠等：政治机器人在社交媒体空间的舆论干预分析

在Twitter参与中美贸易战的社交机器人

题图：英国《卫报》调查记者卡罗尔TED演讲



## 前期成果：GAMHNN

前期研究成果 社交媒体机器人识别

**GENERAL ATTRIBUTED MULTIPLEX HETEROGENEOUS NETWORK**

- 将社交网络视为**多种关系、多类节点异质网络(AMHEN)**处理
- 旨在解决现有社交媒体机器人识别模型无法区分**多节点多关系**的问题
- 将阿里推荐系统中的**GATNE模型**引入社交媒体机器人识别领域

前期研究成果 社交媒体机器人识别

**实践结果：**

- 基于RandomWalk算法的模型由于遍历节点，难以并行化，**时间复杂度**高，在处理大规模数据时耗时较长
- 基于RandomWalk算法的模型存在**短视性**，无法统观大范围的社区结构
- 由于推荐系统涉及的用户-商品关系图的特殊稀疏性，基于阿里推荐系统GATNE模型构建的**GAMHNN模型**在社交媒体网络中的**泛化性差**

前期研究成果 社交媒体机器人识别

**实践反思：**

- 如何有效处理大规模的社交媒体网络数据
- 如何预处理数据使得网络节点得到更好的表示
- 如何降低模型的时间与空间复杂度

基于 BP 神经网络对社交媒体机器人的识别及对现有模型的修改

**摘要：**

本文旨在研究 BP 神经网络对社交媒体机器人的识别及对现有模型的修改。首先，为减少社交媒体网络数据的高维度和稀疏性，特提出一种改进的图神经网络方法（即 Graph2Vec），该方法通过对图结构、节点属性、边属性进行联合建模，有效提升了模型的泛化能力。

**二、模型改进**

改进一：将图神经网络与图注意力网络相结合，提升模型的泛化能力。

**三、实验结果**

本文在公开数据集上进行了实验，结果表明改进后的模型在识别准确率、召回率等方面均优于现有模型。

改进二：将图神经网络与图注意力网络相结合，提升模型的泛化能力。

**四、结论**

本文提出的改进模型在社交媒体机器人识别任务中表现出优越的性能，具有较高的泛化能力和鲁棒性。未来将进一步探索图神经网络在社交网络中的应用。

改进三：将图神经网络与图注意力网络相结合，提升模型的泛化能力。

**五、总结**

本文提出的改进模型在社交媒体机器人识别任务中表现出优越的性能，具有较高的泛化能力和鲁棒性。未来将进一步探索图神经网络在社交网络中的应用。

改进四：将图神经网络与图注意力网络相结合，提升模型的泛化能力。

**对比实验**

模型名称	准确率	召回率	F1分数
Graph2Vec	0.85	0.82	0.835
GCN	0.78	0.75	0.765
GraphSAGE	0.72	0.70	0.71

实验结果表明，改进后的模型在识别准确率、召回率等方面均优于现有模型。





## 中期成果：CDCL

**前期研究成果** **后续研究进展** **社交媒体机器人识别**

**图卷积网络的过度平滑现象**  
在图卷积网络中，信息不断从一个节点传播到其邻居节点，当进行多层的信息传播时，节点会受到越来越远的节点的影响，导致信息不断融合，同时可能导致精度消失，造成节点信息过度平滑

**大规模异质图处理**  
大规模异质图可能包含数百万甚至数十亿个节点和边，需要高效的图处理算法和存储方法，也需要设计高效的图神经网络，有效地聚合不同节点之间的语义信息

**社交网络社区化**  
在真实世界的社交网络中，群体通常聚集为小社区，这些社区中各节点之间交互频繁，且具有高度相似的特征信息，有效地利用社区中丰富的语义特征和密集连接，有助于提取社交网络的深层次信息

**模型鲁棒性及泛化能力**  
现有的社交网络数据集所包含的社交网络通常只是海量社交媒体信息中的一小部分，需要从海量的社交网络泛化到真实社交网络，要求模型具有较高的鲁棒性及泛化能力

**聚焦问题**

**前期研究成果** **后续研究进展** **社交媒体机器人识别**

**解决方案：CDCL**

**Community Detection & Link Prediction** **Graph Contrastive Learning**

**前期研究成果** **后续研究进展** **社交媒体机器人识别**

**CDCL：异质图对比学习**

**Graph Contrastive Learning**

- 利用异质图 transformer 图卷积模型 (HGT)，结合GCN的信息传播机制，以及 transformer 架构的注意力机制，深入挖掘异质子图信息，提取不同类型节点的高维特征表示
- 利用图对比学习架构，基于图卷积链接预测进行边增强，基于PageRank进行自适应节点特征增强，旨在提升模型鲁棒性，提高模型在真实社交媒体网络中的泛化能力，同时通过设计对比损失函数，使模型专注于难分类、错分类的节点

**前期研究成果** **后续研究进展** **社交媒体机器人识别**

**CDCL**

**Community Detection & Link Prediction** **Graph Contrastive Learning**

CDCL首先对整个图进行粗粒度层面的社区检测与链接预测，再对局部密子图进行细粒度的图对比学习，使得最终的节点嵌入表示包含对整个图的整体视野，同时包含局部社区的高维细节。通过理论分析，CDCL的网络结构以及训练范式能够处理大规模异质图网络，应对真实社交网络环境下社区化的挑战，缓解图卷积神经网络中的过度平滑问题，对真实社交网络具有较高的鲁棒性和泛化能力

01 粗粒度层面的社区检测与链接预测

02 细粒度层面的异质图对比学习

## 领域研究分析

## 提出解决方案

## 异质图对比学习

## 模型优势分析

**前期研究成果** **后续研究进展** **社交媒体机器人识别**

Dataset	User	Priority	Time	Edge
Twitter-20	229673	8722738	3348932	45958

Method	Accuracy	F1-score	MCC
GCN	0.650	0.618	0.653
GAT	0.624	0.603	0.676
HGT	0.627	0.654	0.695
SimpHGNN	0.608	0.643	0.693
BotGCN	0.640	0.667	0.693
RCT	0.618	0.669	0.694
ourProp CDL	0.695	0.689	0.706

**前期研究成果** **后续研究进展** **研究安排** **社交媒体机器人识别**

**CDCL**

初步实验表明，该模型能够处理大规模图数据（100万个用户节点、近9000万文本节点、1G条边），在Twitter-20数据集上表现优异

Dataset	C-15	G-17	C-17	M-18	C-S-18	C-R-19	B-F-19	Twibot-20	Twibot-22
# Human	1,950	1,394	3,474	8,092	6,174	340	380	5,237	860,657
# Bot	3,351	1,090	10,894	42,446	7,102	353	138	6,589	139,943
# User	5,301	2,484	14,368	50,538	13,276	693	518	11,826	1,000,600
# Tweet	2,827,257	0	6,677,635	0	0	0	0	33,488,192	98,217,457
# Human Tweet	2,631,730	0	2,839,361	0	0	0	0	33,488,192	81,250,102
# Bot Tweet	196,027	0	3,798,254	0	0	0	0	33,488,192	6,967,355
# Edge	7,896,134	0	6,637,635	0	0	0	0	33,716,171	170,185,937

**前期研究成果** **后续研究进展** **研究安排** **社交媒体机器人识别**

**Twibot-22 Baselines**

Model	Type	C-15	G-17	C-17	M-18	C-S-18	C-R-19	B-F-19	Twibot-20	Twibot-22
GCN	PR	0.618	0.618	0.618	0.618	0.618	0.618	0.618	0.618	0.618
GAT	PR	0.624	0.624	0.624	0.624	0.624	0.624	0.624	0.624	0.624
HGT	PR	0.627	0.627	0.627	0.627	0.627	0.627	0.627	0.627	0.627
SimpHGNN	PR	0.608	0.608	0.608	0.608	0.608	0.608	0.608	0.608	0.608
BotGCN	PR	0.640	0.640	0.640	0.640	0.640	0.640	0.640	0.640	0.640
RCT	PR	0.618	0.618	0.618	0.618	0.618	0.618	0.618	0.618	0.618
ourProp CDL	PR	0.695	0.695	0.695	0.695	0.695	0.695	0.695	0.695	0.695

**对比实验**

在Twibot-20、Twibot-22上进行对比实验，并在预训练基础上选择一些小规模的数据集进一步微调，验证模型的准确性，确保实验的充分性与完整性

**前期研究成果** **后续研究进展** **研究安排** **社交媒体机器人识别**

**消融实验**

在Twibot20中等规模数据集上开展消融实验，验证模型各组件之间的配合协作关系，对模型架构的可行性和优异性作进一步实验证明

## 模型效果对比

## 数据集选择

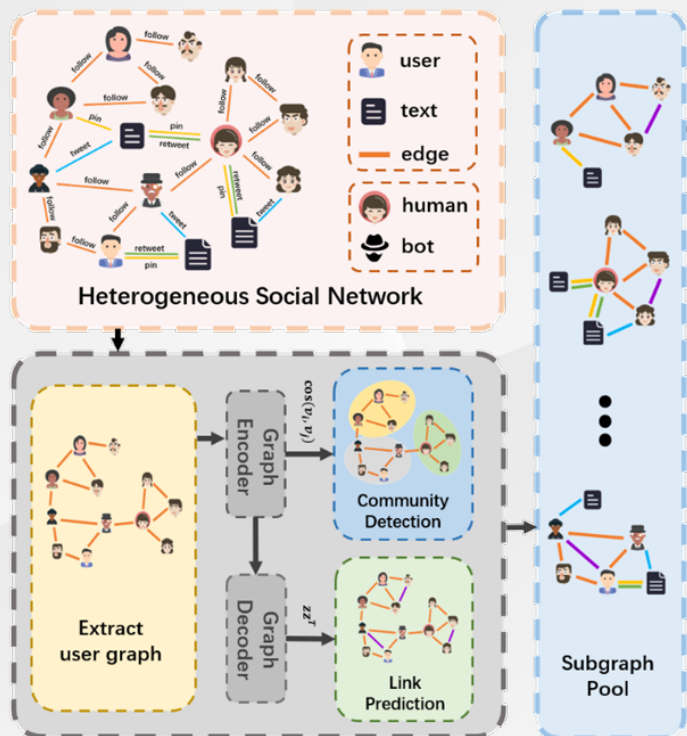
## 对比实验

## 消融实验

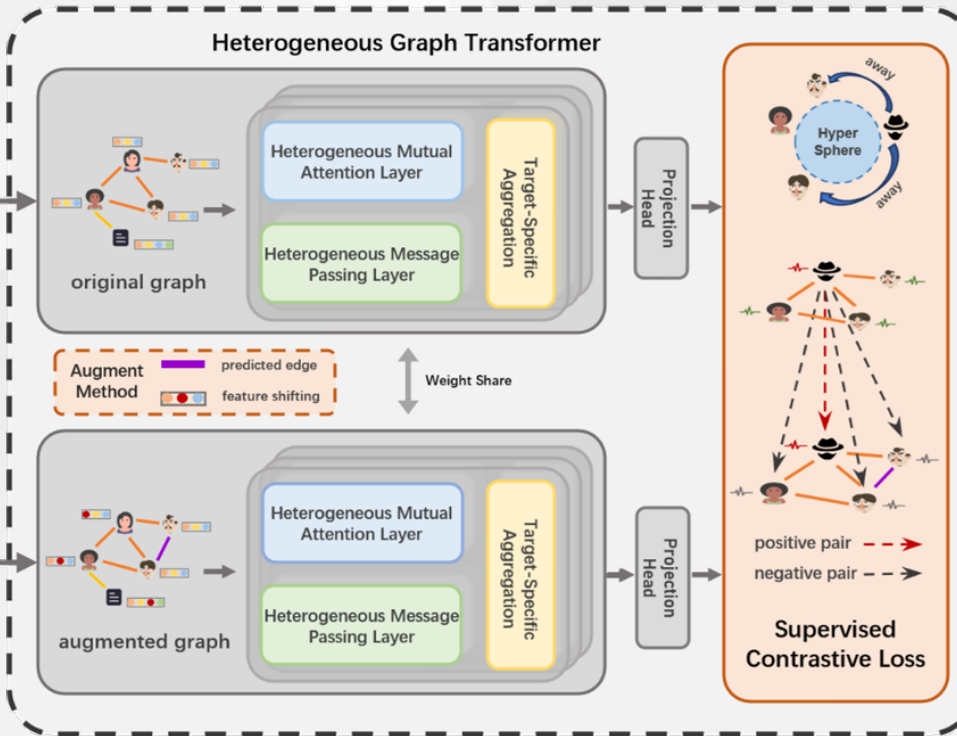


## 中期成果：CDCL

### Community Detection & Link Prediction



### Graph Contrastive Learning

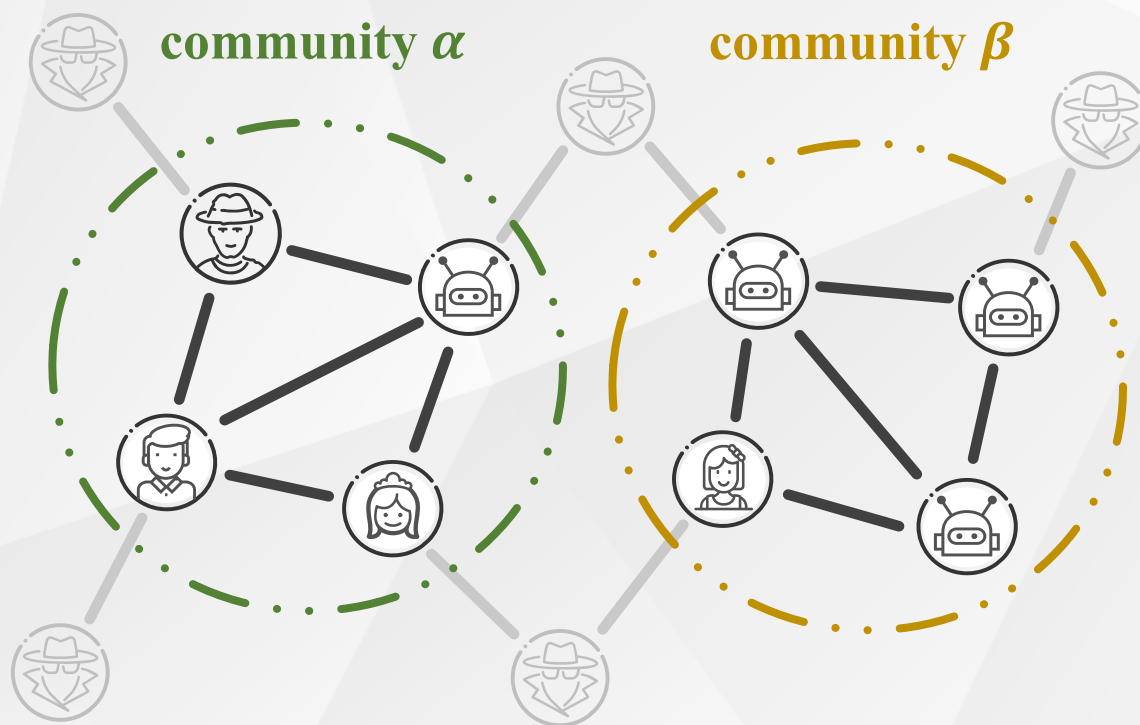


针对以往工作忽视社交网络结构特点以及GNNs过度平滑问题，提出利用社区检测与连接预测方法，在图层面识别出社交网络中的社区，在节点层面对社区子图进行对比学习，提升模型对机器人的识别能力。

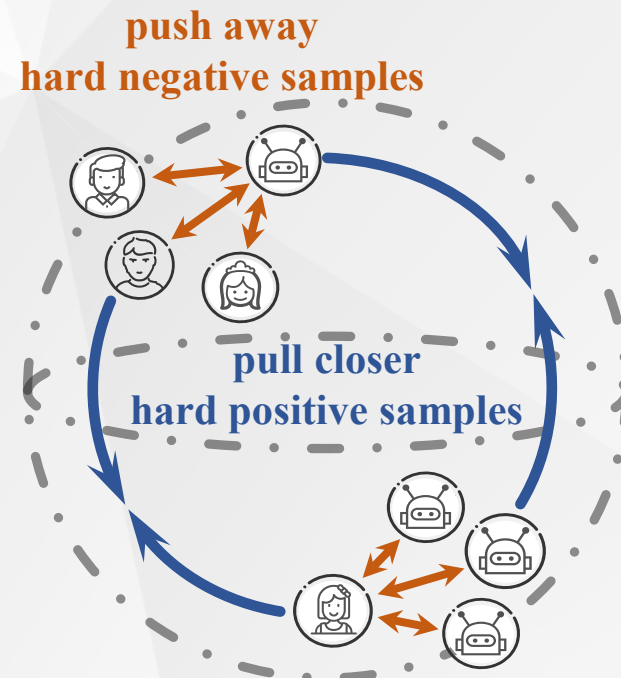




研究动机：拉近难分正样本，推远难分负样本



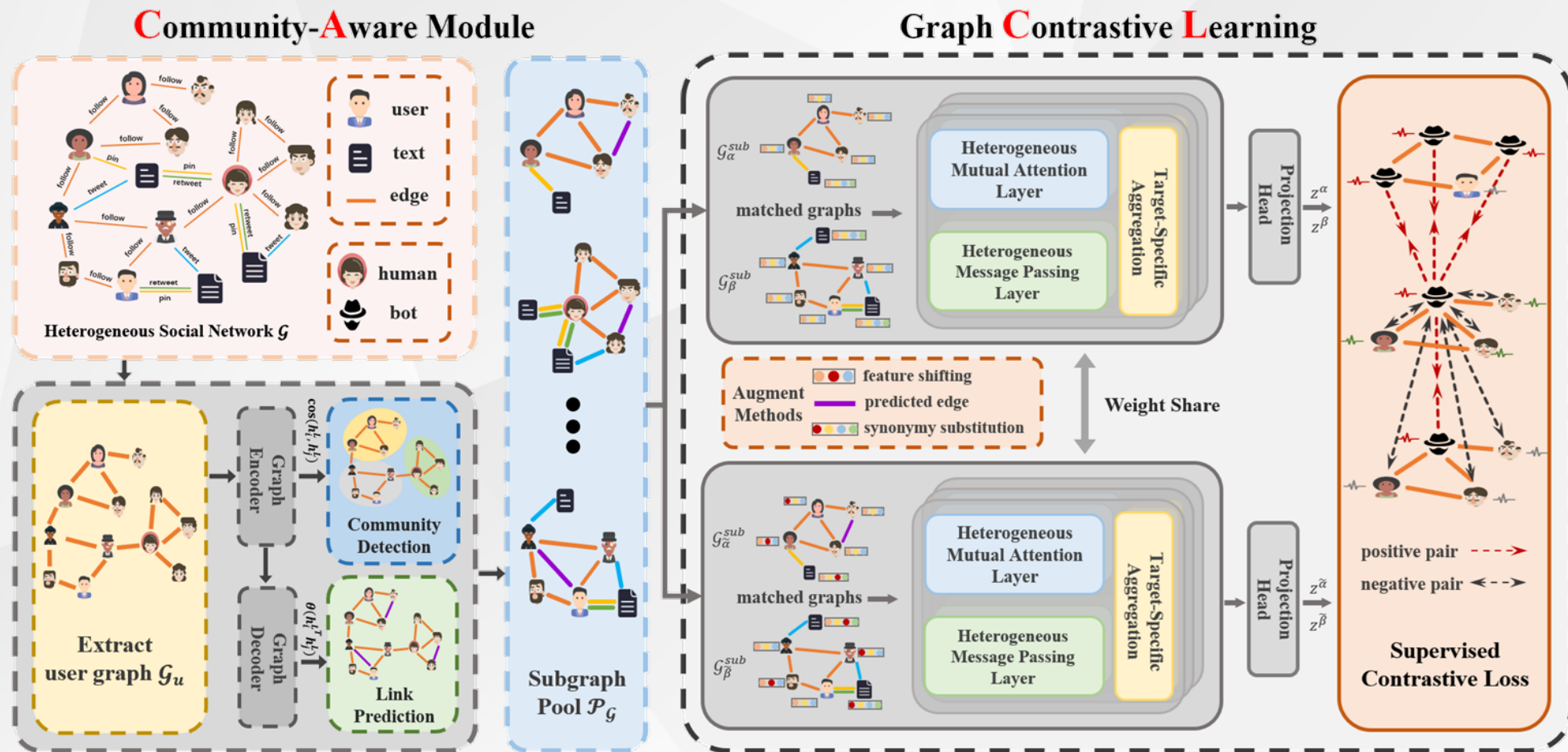
community-aware hard sample mining



contrastive learning on hard samples

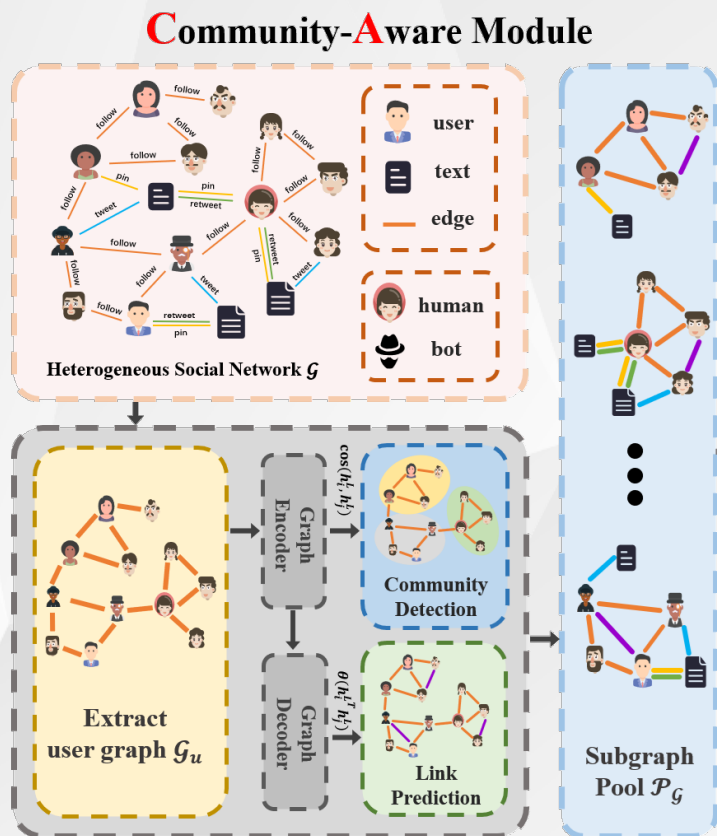


CACL: Community-Aware Heterogeneous Graph Contrastive Learning for Social Media Bot Detection



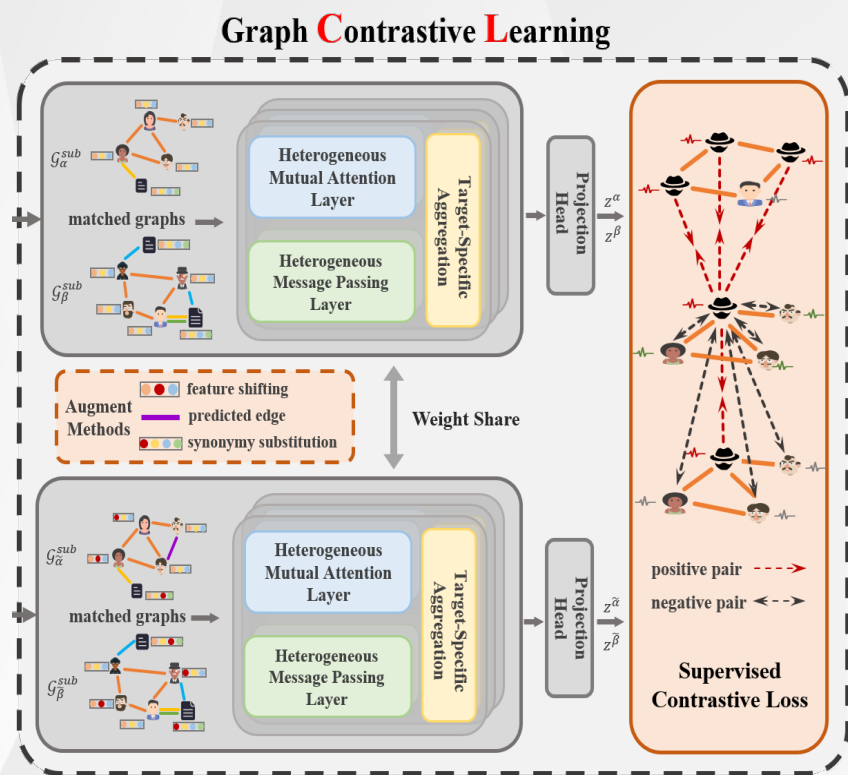
针对以往工作忽视社交网络结构特点以及GNNs过度平滑问题，提出具有社区感知能力的图对比学习框架，在图层面对社交网络进行社区检测，在节点层面对社区子图进行对比学习，提升模型对机器人的识别能力。

## CACL: 社区感知模块



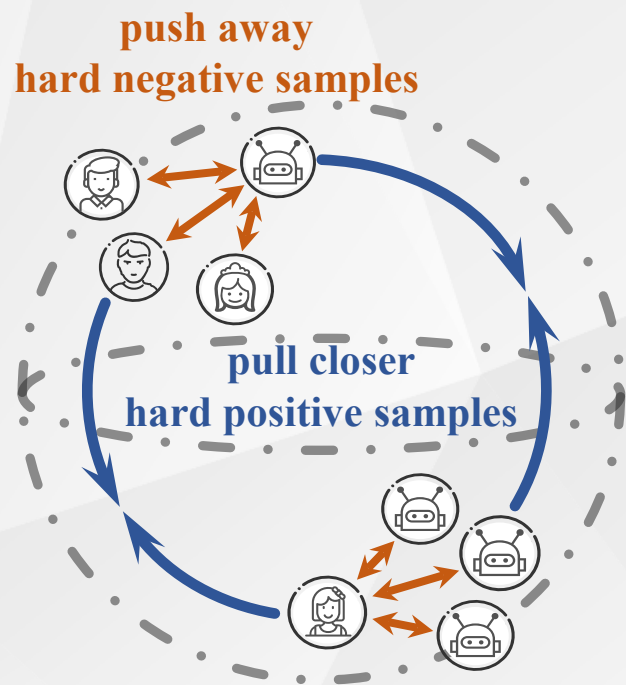
- 利用图编码器对大规模的用户图进行社区检测，在粗粒度层面识别出社区结构，将用户图分割为小规模子图，从而挖掘出难分正样本与难分负样本
- 利用图解码器重构网络空间，实现连接预测，融合图结构信息以及节点信息，为对比学习中的图增强过程提供有力支持，避免了随机图增强对图结构的破坏

## CACL: 图对比学习模块



- 利用图卷积模型（如GAT、SAGE、HGT），深入挖掘异质子图信息，提取不同类型节点的高维特征表示
- 在拓扑结构、用户属性、文本信息三个角度，基于连接预测进行边增强、基于PageRank对节点信息自适应扰动、基于WordNet进行近义词替换，实现图增强

## CACL: 对比损失函数



拉近难分正样本，推远难分负样本

$$\mathcal{L}_i^{\alpha 1} = -\frac{1}{N_{\alpha 1}} \sum_{v_i \in \mathcal{V}_{\alpha 1}} \log \frac{s_i^{self} + s_i^{posi}}{s_i^{self} + s_i^{posi} + s_i^{nega}}$$

$$s_i^{self} = e^{\cos(z_i^{\alpha 1}, z_i^{\alpha 2})}$$

$$s_i^{posi} = \frac{1}{2} \sum_{v_j \in \mathcal{V}_{\beta 1, 2}} \delta_{ij} e^{\cos(z_i^{\alpha 1}, z_j^{\beta 1, 2})}$$

$$s_i^{nega} = \frac{1}{2} \sum_{v_j \in \mathcal{V}_{\alpha 1, 2}} \hat{\delta}_{ij} e^{\cos(z_i^{\alpha 1}, z_j^{\alpha 1, 2})}$$



## CACL: 主实验

CACL框架与不同的基线模型相比呈现出显著的改进，这表明对比学习和社区感知模块有助于GNNs模型更好地捕捉节点信息。

Method	Cresci-15			Twibot-20			Twibot-22		
	Accuracy	F1-score	MCC	Accuracy	F1-score	MCC	Accuracy	F1-score	MCC
BotRGCN	94.71	95.85	89.06	<u>84.43</u>	85.79	67.17	73.90	<u>48.30</u>	31.01
DeeProBot	72.35	81.61	44.26	77.34	80.32	54.40	74.11	12.73	14.12
RGT	96.27	<u>97.04</u>	<u>92.25</u>	84.02	86.07	68.08	74.49	45.36	29.64
EvolveBot	92.18	90.07	72.77	65.83	69.75	30.79	71.09	14.09	13.38
SimpleHGN	93.13	94.68	85.91	83.93	86.33	68.40	72.53	<b>49.59</b>	30.74
RoBERTa	<u>97.01</u>	95.86	89.55	75.55	73.09	42.15	72.07	20.53	19.35
BGSRD	87.78	90.80	74.92	66.36	70.05	32.28	71.88	21.14	20.32
SATAR	92.71	94.55	85.61	84.02	85.74	67.83	/	/	/
GAT	93.52	94.99	86.78	82.24	84.93	64.95	74.17	43.75	28.12
+CACL	94.51	95.42	88.71	83.52	86.39	68.56	74.50	44.27	28.94
SAGE	93.92	95.26	87.48	83.00	85.48	66.38	74.39	46.47	30.25
+CACL	<b>97.65</b>	<b>98.12</b>	<b>95.10</b>	83.60	<u>86.53</u>	<u>68.95</u>	<b>75.38</b>	47.45	<u>32.27</u>
HGT	92.24	94.27	84.83	84.02	86.35	68.48	<u>75.07</u>	43.13	29.18
+CACL	95.88	96.74	91.45	<b>85.12</b>	<b>87.28</b>	<b>70.75</b>	<u>75.07</u>	48.27	<b>32.39</b>

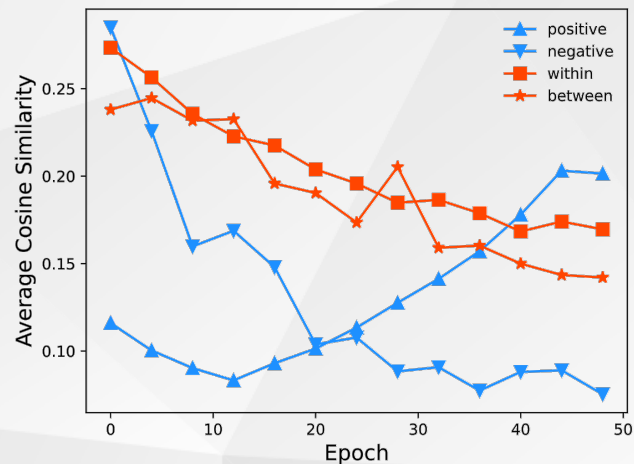
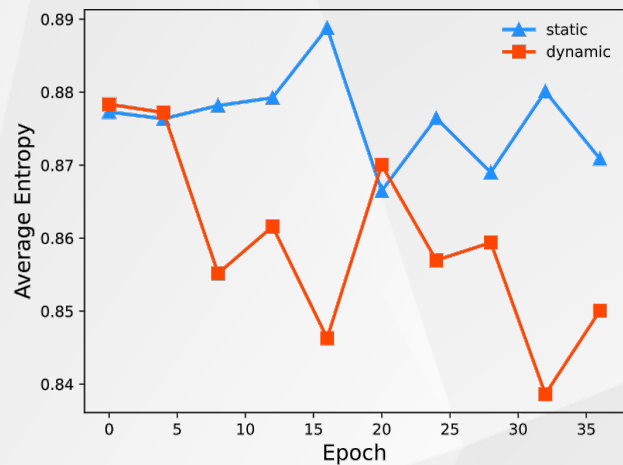
## CACL: 对比实验与消融实验

CL Loss	Accuracy	F1-score	MCC
unsuper	84.27	86.75	69.40
super	84.78 $\uparrow$ 0.51	86.95 $\uparrow$ 0.20	69.97 $\uparrow$ 0.57
static	84.45 $\uparrow$ 0.18	86.74 $\downarrow$ 0.01	69.42 $\uparrow$ 0.02
dynamic	85.12 $\uparrow$ 0.85	87.28 $\uparrow$ 0.53	70.75 $\uparrow$ 1.35

Ablation	Accuracy	F1-score	MCC
CACL	85.12	87.28	70.75
w/o SS	84.78 $\downarrow$ 0.34	87.19 $\downarrow$ 0.09	87.19 $\downarrow$ 0.24
w/o LP	84.45 $\downarrow$ 0.67	86.98 $\downarrow$ 0.30	70.01 $\downarrow$ 0.74
w/o FS	85.12 $\downarrow$ 0.00	87.09 $\downarrow$ 0.19	70.42 $\downarrow$ 0.33
w/o TX	83.34 $\downarrow$ 1.78	86.51 $\downarrow$ 0.77	69.14 $\downarrow$ 1.61
w/o MT	82.92 $\downarrow$ 2.20	86.10 $\downarrow$ 1.18	67.97 $\downarrow$ 2.78
w/o HG	82.75 $\downarrow$ 2.37	85.28 $\downarrow$ 2.00	65.87 $\downarrow$ 4.88
w/o CA	84.36 $\downarrow$ 0.76	86.44 $\downarrow$ 0.84	68.87 $\downarrow$ 1.88

- 无监督对比损失函数、有监督对比损失函数、静态与动态情况下的社区感知对比损失互相对比，验证**损失函数的有效性**
- 对三种不同的图增强方法、三种数据结构成分分别进行消融，探究**模型架构、数据组成**对模型的影响

## CACL: 内部机理探究



➤ 宗辰辰学长：试试信息熵

为了探究模型的内部机理，采用信息熵衡量社区内部节点类别纯度的变化。在静态情况下，社区内部信息熵基本稳定；动态情况下，信息熵逐步下降，证明检测出的社区内部类别趋于一致，有利于节点分类

➤ 为了证明对比损失函数的有效性，对难分正样本、难分负样本，社区内部、社区之间的余弦相似度的变化进行观察，证明节点特征按预期同类相聚、异类相斥

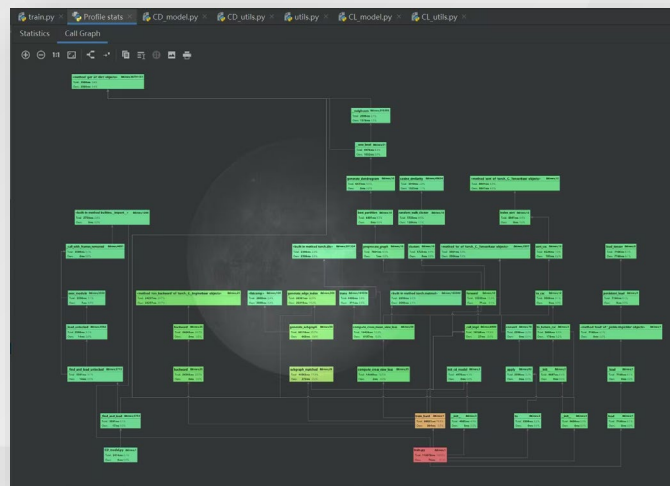
## 遇到的困难及解决办法

Issue: 手写的代码缺乏优化, 导致时间复杂度过高

Solution: 利用Pycharm图形化性能测试工具Profile查找高消耗函数并进行优化

```
Total lines in D:\SocialMediaRobots\CACL: 7400
```

```
Total lines in D:\SocialMediaRobots: 37609
```

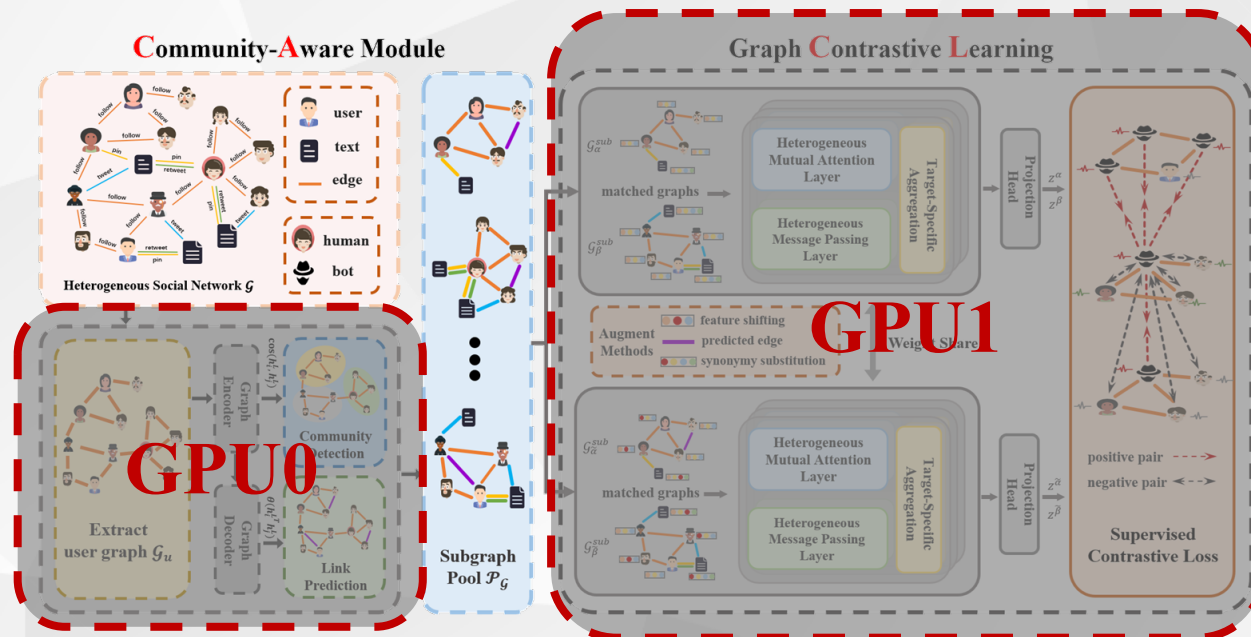




## 遇到的困难及解决办法

Issue: Twibot-22数据集过大 (100GB+), 单卡显存 (24GB) 不足

Solution: 将模型**串行**部署到两个显卡上, 动态更新图卷积网络参数



The screenshot shows the GitHub repository page for 'CACL' by SirryChen. The repository is private and has 1 branch and 0 tags. The file list includes: results, CD\_model.py, CD\_model\_old.py, CD\_utils.py, CL\_model.py, CL\_utils.py, README.md, cd\_config.json, predata\_cresci15.py, predata\_twibot20.py, predata\_twibot22.py, pretrain.py, requirements.txt, train.py, and utils.py. The repository has 18 commits and 1 star.

The screenshot shows the README file for the 'CACL' repository. The title is 'CACL: Community-Aware Heterogeneous Graph Contrastive Learning for Social Media Bot Detection'. The README includes sections for Overview, Datasets, Training details, and Citation.

## Overview

The steps to reimplement this work mainly contain:

- Data preprocessing
  - Three datasets are used and programs for preprocessing can be found as `predata_{dataset name}.py`.
  - Just download the datasets and change the file paths augments.
- Pretrain
  - Community-aware module is employed to perform community detection and link prediction, which needs to be pretrained to prevent cold-start problem.
  - Simply change the preprocessed data path augment and use `pretrain.py`.
- Train
  - After obtaining preprocessed data and pretrained community-aware module weight, use `train.py` to train the full model.
  - Hyperparameters can be adjusted in the function `super_parameter_initial()` in `utils.py`.
- Test
  - After training the model using `train.py`, the test results will be shown and saved.

## Datasets

We have used three datasets throughout the entire work. You may need to contact the author to get access to some of them.

- Cresci-15 [\[paper\]](#) [\[dataset\]](#)
- Twibot-20 [\[paper\]](#) [\[github\]](#)
- Twibot-22 [\[paper\]](#) [\[github\]](#) [\[dataset\]](#)
- Other datasets may be helpful [\[web path\]](#)

## Training details

Using dataset Cresci-15 and backbone HGT for example.

```
python3 pretrain.py --dataset cresci15 --basic_model HGT
```

Train the CACL with HGT

```
python3 pretrain.py --dataset cresci15 --basic_model HGT --max_error_times 5
```

Here are some key options of the hyperparameters

- `basic_model`: CACL framework support 3 backbones as the convolutional layer including GAT, SAGE, and HGT.
- `num_layer`: the layer number of the convolutional network, we use 2 by default.
- `lr_warmup_epochs`: during the initial training warm-up phase of the model, we increase the weight of the contrastive loss, aiming for the model to quickly find the optimal point.
- `max_error_times`: we use the validation dataset for early stopping.
- `cluster`: we implement several cluster method for community detection, we use randomwalk by default.

The details of other optional hyperparameters can be found in the function `super_parameter_initial()` in `utils.py`.

## Citation

Please consider citing the following paper when using our code for your application.

## 深圳零一学院·开放嘉年华

## 项目海报



### 社交媒体机器人识别

#### CACL: Community-Aware Heterogeneous Graph Contrastive Learning for Social Media Bot Detection

陈思远, 丰硕, 梁松松

南京航空航天大学

2024/01/26

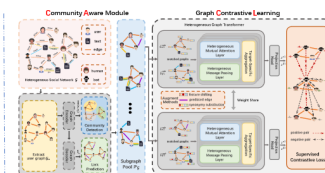


### CACL: Community-Aware Heterogeneous Graph Contrastive Learning for Social Media Bot Detection

项目成员: 陈思远 丰硕 梁松松 所在院校: 南京航空航天大学 所在专业: 长空创新班 年级: 2021级  
指导老师: 李杰斌 项目进展: 论文在投

#### 项目背景及意义

- 社交媒体机器人广泛存在于社交网络, 在X (Twitter)上9%到15%的活跃用户表现出类似机器人的行为;
- 社交媒体机器人负面应用多, 如在虚假信息传播、干扰政治活动、主张极端主义, 未来甚至可能被应用到现代化信息战中;
- 高准确率、泛化能力强的社交机器人识别模型有助于维护网络空间安全;
- 传统的机器人识别模型无法有效检测出伪装成真实用户和集体攻击的机器人;
- 基于图神经网络 (GNNs) 的方法仍然面临一些问题, 比如数据集相对较小和信息传播机制导致的模型泛化能力差;
- 大多数方法聚焦于如何改善图卷积网络的性能, 却忽视了社交网络中的社区结构。



#### 结果

Method	Twitter-17			Twitter-22		
	Accuracy	F1 score	MCC	Accuracy	F1 score	MCC
BoRoD	84.71	81.81	88.08	84.43	81.79	87.17
Neural	73.19	61.81	64.39	71.34	60.32	64.60
RD	86.22	82.26	82.22	84.02	84.07	88.08
Ensemble	82.18	80.07	75.77	80.83	80.75	71.80
RD+L	71.10	71.91	62.91	81.26	84.90	84.06
EmpyBot	81.33	84.48	82.91	83.89	84.33	88.40
GCN	80.75	80.55	80.61	80.00	81.76	87.81
GAT	81.52	84.89	86.79	82.24	84.93	84.85
CACL	85.21	85.43	88.71	85.21	86.39	89.04
CACL-SAT	81.92	81.28	87.48	83.00	82.48	86.38
RD+L+CACL	87.46	88.12	88.39	85.80	86.51	89.81
RD+L	82.54	84.27	84.83	84.02	88.33	88.48
CACL+RD+L	85.81	85.74	91.61	88.11	87.38	90.79

#### 研究思路与方法

提出了CACL, 一个能够感知社区结构的图对比学习框架。采用社区感知模块进行社区检测, 并在粗粒度层面挖掘难分样本和正样本, 利用图对比学习在细粒度层面处理这些难分样本。引入了自适应的连接预测、同义词替换和节点特征干扰, 在图结构和特征两个层面进行图增强。此外, 通过共享底层特征提取网络, 对齐了社区感知模块和对比学习模块, 进行动态、自适应的难分样本挖掘。



$$C_i^{(t)} = -\frac{1}{N_{i,t}} \sum_{j \in N_{i,t}} \log \frac{e^{g_{ij}^{(t)}} + g_{ij}^{(t)}}{e^{g_{ij}^{(t)}} + g_{ij}^{(t)} + s_{ij}^{(t)}}$$

$$s_{ij}^{(t)} = e^{-\alpha \cos(\theta_{ij}^{(t)})}$$

$$s_{ij}^{(t)} = \frac{1}{2} \sum_{k \in N_{i,t}} \delta_{k,j} \cos(\theta_{ik}^{(t)})$$

$$s_{ij}^{(t)} = \frac{1}{2} \sum_{k \in N_{i,t}} \delta_{k,i} \cos(\theta_{kj}^{(t)})$$

#### 结论与讨论, 未来计划

- 使用CACL框架提升了三种卷积网络的性能, 说明了模型框架的有效性;
- 消融实验验证了模型各组件的有效性及其相互协同性质, 证明了三种不同的图增强方法能够提升模型的鲁棒性;
- 不同对比损失函数之间的对比证明了改进的损失函数的有效性;
- 社区感知下的难分样本的挖掘影响模型效果的原则, 计划通过社区中样本类别信息的变化进行揭示。

#### 致谢

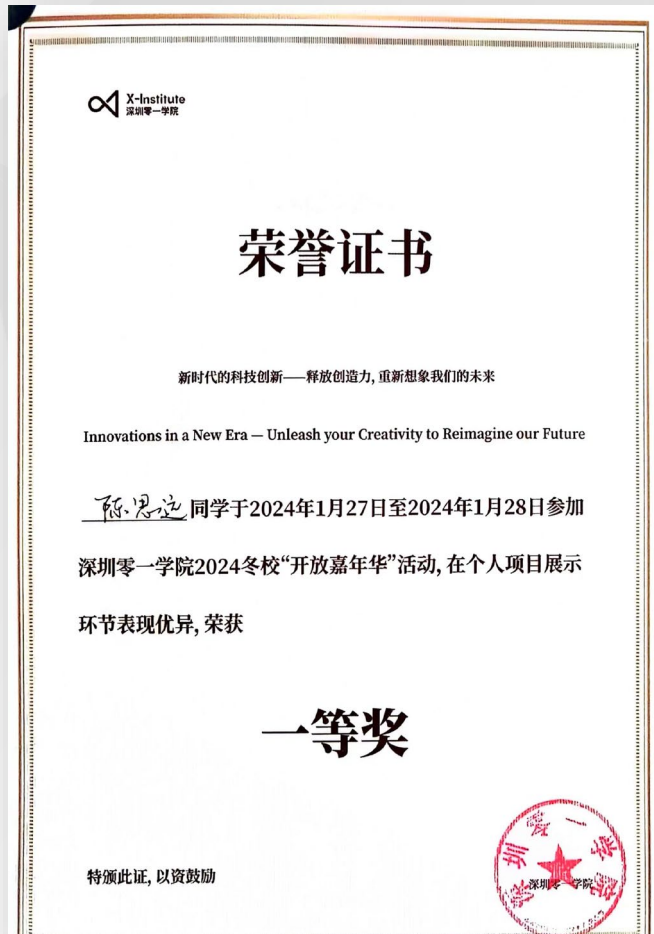
项目成员协力共进, 李杰斌老师为团队提供计算资源与指导, 南航长空学院提供资金支持。

#### 参考文献

- Cresci S, Di Pietro R, Spognardi A, et al. Demystifying Misconceptions in Social Bots Research [J]. arXiv preprint arXiv:2303.17251, 2023.
- Ferrara E. Disinformation and social bot operations in the run up to the 2017 French presidential election [J]. arXiv preprint arXiv:1707.00088, 2017.
- Berger J M, Morgan J. The ISIS Twitter Census: Defining and describing the population of ISIS supporters on Twitter [J]. 2015.
- Varol O, Ferrara E, Davis C, et al. Online human-bot interactions: Detection, estimation, and characterization [C]//Proceedings of the international AAAI conference on web and social media. 2017, 11(1): 280-289.

# 深圳零一学院·开放嘉年华

## 项目获得一等奖







## 论文：投稿至自然语言处理顶会ACL 2024

### ACL 2023 - The 61st Annual Meeting of the Association for Computational Linguistics

H5 Index : 169 TK5 Index : 304 CCF Level : A

覆盖领域 : Natural Language Proc



ACL  
2024

[Calls](#) [Program](#) [Committees](#) [Venue](#)

## The 62nd Annual Meeting of the Association for Computational Linguistics

Bangkok, Thailand  
August 11-16, 2024

### Welcome!

The 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024) will take place in **Bangkok, Thailand** from **August 11th to 16th, 2024**. More information will be announced soon.

OpenReview.net

Search OpenReview...

Notifications 1 Activity Tasks Sirry Chen

[← Back to Homepage](#)

## The 62nd Annual Meeting of the Association for Computational Linguistics

[Your Submissions](#) [Author Tasks](#)

#	Paper Summary	Reviews	Decision
1883	<a href="#">CACL: Community-Aware Heterogeneous Graph Contrastive Learning for Social Media Bot Detection</a> <a href="#">Download PDF</a> Sirry Chen, Shuo Feng, Liang Songsong, Chen-Chen Zong, Jing Li, Piji Li ACL 2024 ARR Commitment Submission <a href="#">Show details</a>	<b>0 Reviews Submitted</b> Average Rating: N/A (Min: N/A, Max: N/A) Average Confidence: N/A (Min: N/A, Max: N/A)	<b>ACL 2024 ARR Commitment Submission</b> <b>No Recommendation</b>

[About OpenReview](#)  
[Hosting a Venue](#)  
[All Venues](#)

[Contact](#)  
[Feedback](#)  
[Sponsors](#)

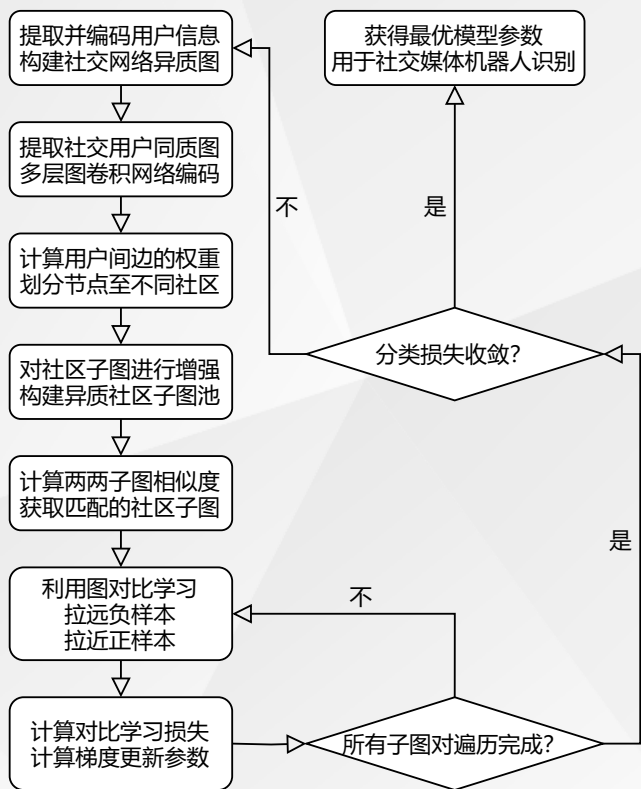
[Frequently Asked Questions](#)  
[Terms of Use](#)  
[Privacy Policy](#)





## 国家发明专利

## 受理通知书



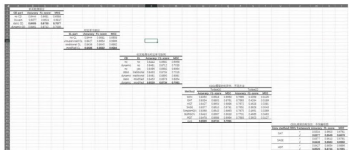
 国家知识产权局	
<b>210016</b> 南京市玄武区龙蟠路 155 号 3 幢 410 室 南京瑞弘专利商标事务所 (普通合伙) 孙晓(0252228215)	发文日: 2024 年 04 月 07 日
	
申请号: 202410407753.8	发文序号: 2024040700865650
<b>专利申请受理通知书</b>	
根据专利法第 28 条及其实施细则第 43 条、第 44 条的规定, 申请人提出的专利申请已由国家知识产权局受理。现将确定的申请号、申请日等信息通知如下: 申请号: 202410407753.8 申请日: 2024 年 04 月 07 日 申请人: 南京航空航天大学 发明人: 李杰, 陈思远, 丰硕 发明创造名称: 具有社区感知能力的图对比学习社交媒体机器人识别方法 经核实, 国家知识产权局确认收到文件如下: 权利要求书 1 份 2 页, 权利要求项数: 6 项 说明书 1 份 5 页 说明书附图 1 份 2 页 说明书摘要 1 份 1 页 专利代理委托书 1 份 2 页 发明专利申请书 1 份 4 页 实质审查请求书 文件份数: 1 份 申请方案卷号: NJ-20240379714-NC3031	
提示: 1. 申请人收到专利申请受理通知书之后, 认为其记载的内容与申请人所提交的相应内容不一致时, 可以向国家知识产权局请求更正。 2. 申请人收到专利申请受理通知书之后, 再到国家知识产权局办理各种手续时, 均应当准确、清晰地写明申请号。	
审查员: 自动受理 联系电话: 010-62356655	
201011 纸质申请, 回函请寄: 100088 北京市海定区中关村西土城路 6 号 国家知识产权局专利局受理处收 2023.03 电子申请, 应当通过专利业务办理系统以电子文件形式提交相关文件。除另有规定外, 以纸质等其他形式提交的文件视为未提交。	



2023-12-5:

不同对比损失函数导致的纯净度的变化  
不同社区同类样本拉近, 同一社区异类样本拉近

2023-11-27:



2023-11-9:

复现记录:

1. SimpleHGN: 0.8368
2. BotRGCN: 0.8410
3. RGT: 0.8478
4. HGT: 0.8427

2023-10-29:

搭建完了模型, 在 twiBot20 预处理后的同质图数据上 acc: 0.857 (sota: 0.87), 发现自己的 Doc2Vec 编码推文效果确实不如用 transformer 预训练好的模型, 另外数据预处理似乎有些地方跟后面的论文不太一样 (之前参考的较早的一篇论文)

2023-10-01:

BotFinder 论文阅读: 主体是半监督学习, 但是实现得很奇怪

对某节点  $i$ :

1. 分别对节点特征进行提取, 得到特征向量  $V_i$
2. 使用 Node2vec 学习一个表征, 再结合标签信息, 与其他节点计算相似度, 分别得到节点  $i$  与机器人、真人节点的最大、平均相似度, 拼接得到特征向量  $U_i$
3. 聚合  $V_i$  和  $U_i$ , 得到一个节点的特征表示, 然后使用 LightGBM (由决策树弱分类器构成) 判断是否是机器人, 获得暂时的分类结果
4. 使用 Louvain 算法进行无监督社区检测 (这个算法只利用节点的出度和入度信息), 将每个节

点划分入不同的社区

5. 对第 3 步得出的结果进行修正: 在同一社区内的节点分类必须相同, 分类不同的节点不进行预测 (我觉得就是这个结果不参与 loss, acc 等计算), 最后得到部分节点的分类结果

6. 进行 loss 计算和模型更新

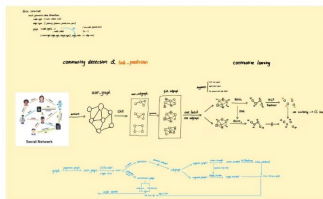
Now, after communities is divided, we will label communities with rules as follows:

- 1) All users in the community are supposed to be of the same label if the users with labels belongs to the same community.
- 2) If the users in a community do not have any label, or if the users are of different labels, we will not make prediction.

However, prediction may not cover all users. So, performance in this rule is limited. But the result of this rule is more accurate than LightGBM. Through combining the above two steps, performance can be further improved.

2023-09-26:

用 community detection 切割子图从而提取社区化 batch, 利用 contrastive learning 处理异质图, 学习用户节点信息, loss 部分目前采用余弦相似度, 还在考虑优化。



2023-08-28:

基于 GGATNE 模型的代码已经完善, 在跑, 很慢, 看起来不太行:

最近在看对比学习, 主要想从 BYOL 改, 有两个困难: 1. 如何处理异质图 2. 子图分割。在找一些论文解决这两个问题。

2023-07-04:

初步编写好了基于 GATNE 模型的社交媒体机器人识别模型代码, 发现数据预处理后数据量太大, 正在寻找文本特征提取模型进行轻量化数据预处理: 完成数据预处理后再对模型主体代码进行测试。

2023-05-05:

发现了 HAN 模型的缺点, 提出利用 GATNE 模型处理 Attributed Multiplex Heterogeneous Network (or AMHEN) 网络, 初步搭建了模型整体框架

2023-04-28:

阅读并复现了两篇论文:

BotRGCN: 利用 RGCN 处理图异质网络

Heterogeneity-Aware: 利用加入 transformer 的 CNN (HAN) 处理图异质网络

2023-04-15:

4-9: 整理了 SATAR 的代码逻辑, 发现代码里有 neighbors 子网络与论文里有阐述的不一样 (原文代码太乱就没用, 后面 pipeline 还是用数据集论文里的代码比较方便)

2023-04-07:

4-7: 跑完了代码, acc 略低于论文里的, 应该是因为 pretrain 里 epoch 设得比较小

4-5: 看完了 SATAR 代码, 对照论文大概理清了整体思路, 准备把代码跑完看看;

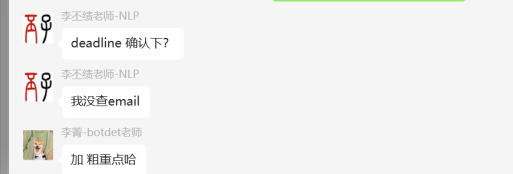
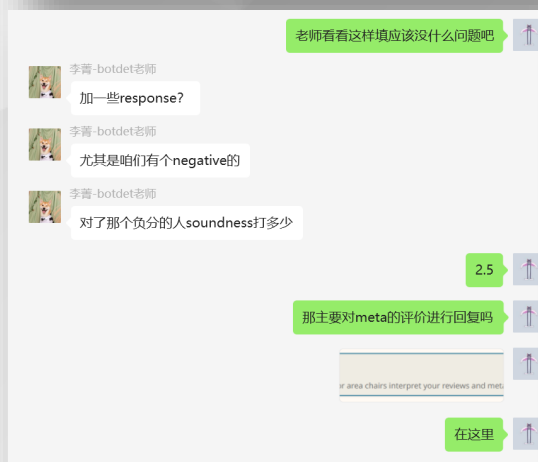
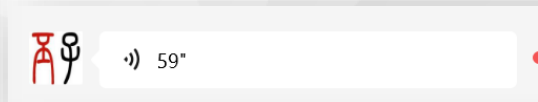
4-4: 找到一篇采用利用异质图和 transformer 来做 botdet 的论文, 准备研究一下;

2023-03-31:

3-28 到 4 月 2 日: 看完了 pretrain 部分代码, 还有一些细节问题留着解决;

3-27: 阅读完了 SATAR 数据预处理部分的代码, 基本理清了数据文件架构及内部数据格式;

3-25: 完成了 xshell, xftp, pycharm 环境配置, 学习了 linux 系统的基本命令, 尝试在服务器上跑了代码;



## 论文撰写、审稿 全流程老师指导

## 研究进度记录 (学长查看)

2023年度大创结项答辩



社交媒体机器人识别

请各位老师批评指正

指导老师：李丕绩

项目成员：陈思远 梁松松 徐敬 施文韬

