BotRGCN: Twitter Bot Detection with Relational Graph Convolutional Networks

Herun Wan

LUD Lab, Xi'an Jiaotong University

wanherun@stu.xjtu.edu.cn

November 11, 2021

Table of Contents

What are Twitter bots?

Definition (Wikipedia)

A **Twitter bot** is a type of bot software that controls a Twitter account via the Twitter API.

Definition (Ours)

Besides used by genuine users, Twitter is home to an ample amount of automated programs, which are also known as **Twitter bots**.

`medium.com`

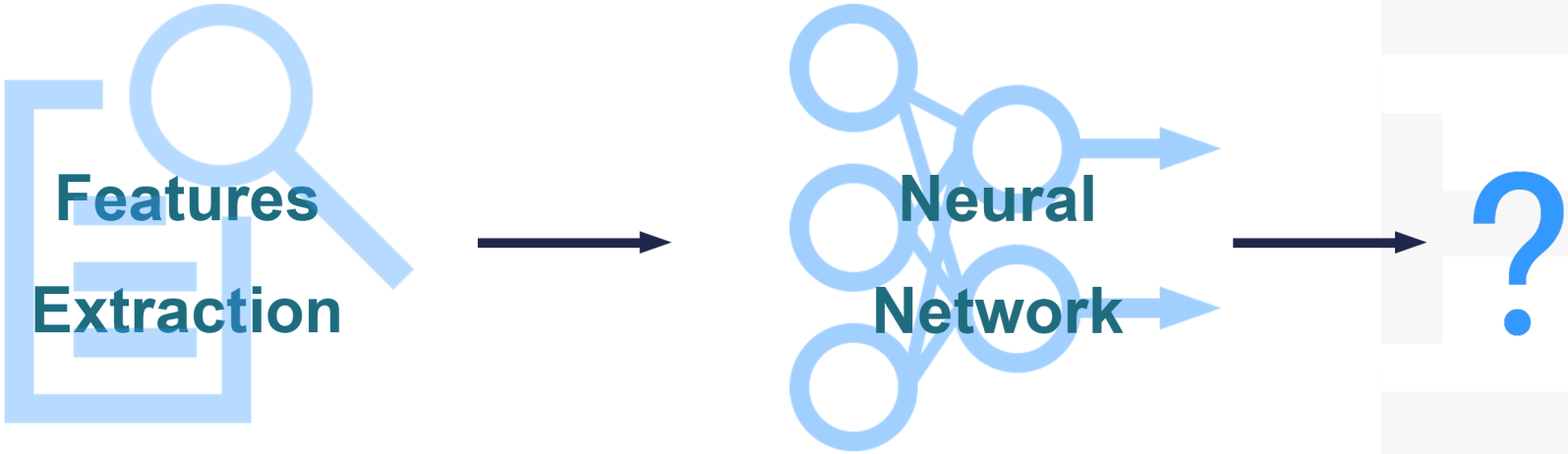
Why should we detect Twitter bots?

Twitter bots are responsible for:

- election interference
  - "Are 'bots' manipulating the 2020 conversation? Here's what's changed since 2016." Washington Post.
  - Deb et al., WWW 2019.
- misinformation campaign
  - "Twitter Bots Poised to Spread Disinformation Before Election." New York Times.
  - "Researchers: Nearly Half Of Accounts Tweeting About Coronavirus Are Likely Bots." NPR.
- extreme ideology propaganda
  - Berger et al., The Brookings project on US relations with the Islamic world.

We focus on identifying **malicious bots** on Twitter.

Existing Bot Detectors



- User profile, ICMSM 2011
- Social networks, ISONAM 2017
- Timeline of accounts, IEEE Intelligent Systems, 2016
- Redirection of URLs, TDSC 2013
- Classification of websites, S&P 2011

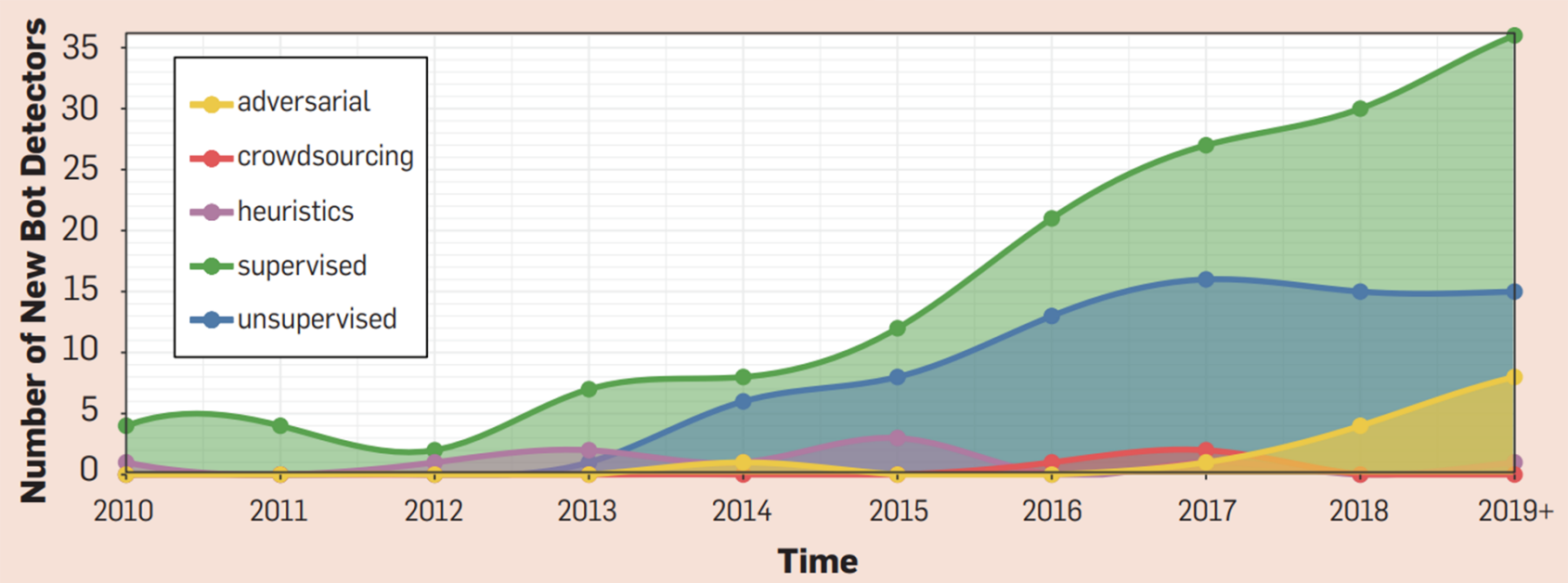
...

- RNN, TPS-ISA 2019
- RNN + Property, Information Sciences, 2018
- GAN, IJCAI 2019
- GCN, WWW 2019

...

Existing Bot Detectors

Existing approaches are generally supervised or unsupervised, while real-world data calls for **semi-supervised** approaches.



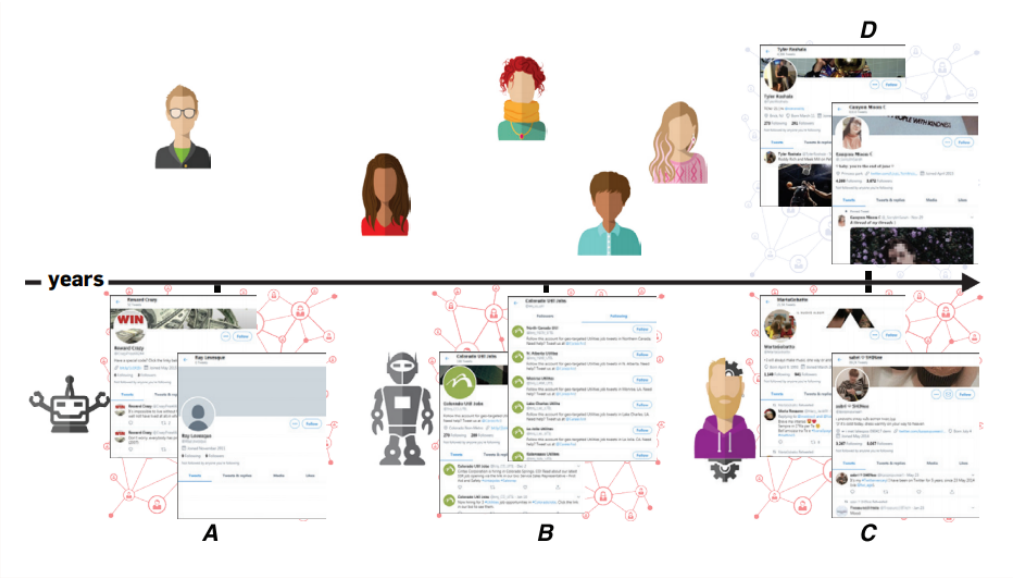
Cresci et al., Communications of the ACM 2020

New Challenges

Challenge of Disguise

Novel Twitter bots try to disguise as genuine users by imitating their behavior.

[Cresci et al. 2020] spotted bots that use stolen names and profile pictures and intersperse few malicious messages with neutral ones.



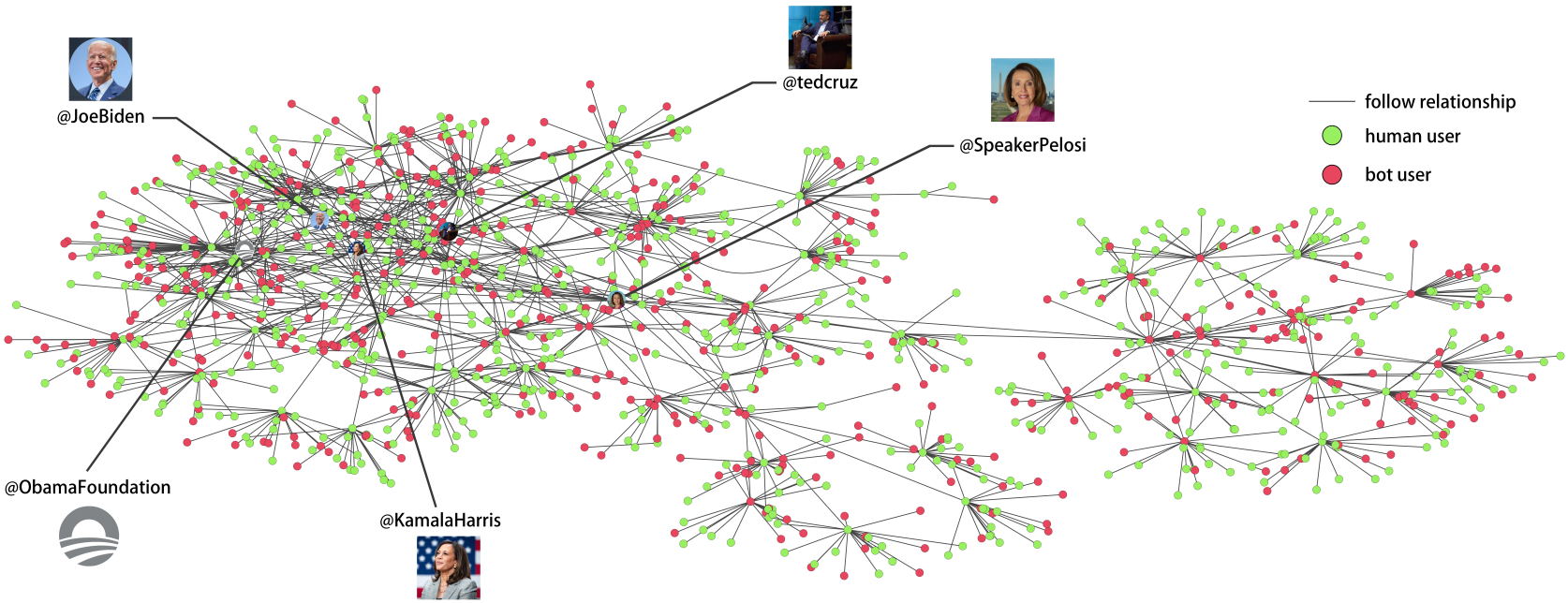
Cresci et al., Communications of the ACM 2020

New Challenges

Challenge of Community

Novel bot communities attack in groups and seem genuine alone, while existing measures analyze each user independently.

[Cresci et al. 2017] identified a bot group and their collective action towards influencing the mayoral election of Rome in 2014.



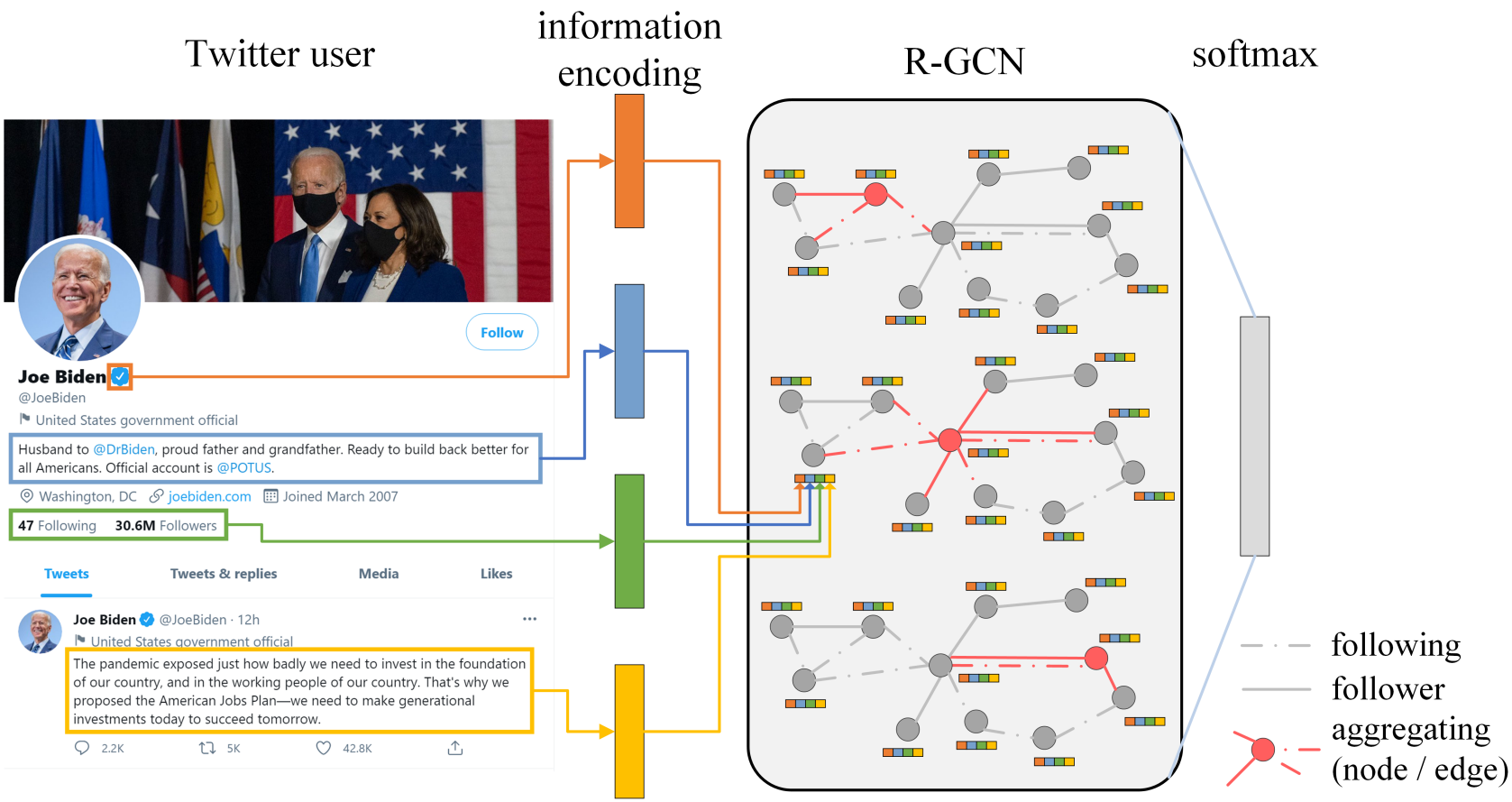
Feng et al. CIKM 2021

That being said

We need a bot detector that is

- **community-aware**
- **disguise-revealing**
- adapt to the **semi-supervised** nature of real-world Twitter.

BotRGCN Overview



Feature Encoding

We jointly leverage four aspects of user information:

- categorical user metadata, e.g. verified
  - one-hot encoding
- numerical user metadata, e.g. # followers
  - z-score normalization
- user description
  - pre-trained language models (RoBERTa)
- user tweets
  - average RoBERTa of 200 most recent tweets

Graph Construction

We construct a heterogeneous information network (HIN) to represent the topological structure of Twitter:

- nodes: each Twitter user
- edges: follow relationship between users

We differentiate follower and following as two different relations, which allow information to flow in the network in two different ways.

Bot Detection

We adopt relational graph convolutional networks (R-GCN) to learn node representations:

$$x_i^{(l+1)} = \Theta_{self} \cdot x_i^{(l)} + \sum_{r \in R} \sum_{j \in N_r(i)} \frac{1}{|N_r(i)|} \Theta_r \cdot x_j^{(l)} \tag{1}$$

BotRGCN then uses a softmax layer on node vectors to classify them into bot or not.

How does BotRGCN addresses challenges?

BotRGCN is **community-aware**, **disguise revealing**, and **semi-supervised**.

- **community-aware**: BotRGCN puts individual users into their social context, identifying bot groups and clusters in the process.
- **disguise-revealing**: BotRGCN jointly leverages semantic, property and neighborhood user information, leaving bot operators with no place to hide.
- **semi-supervised**: BotRGCN operates with a Twitter HIN with 5% annotations, successfully leveraging mass unlabeled data.

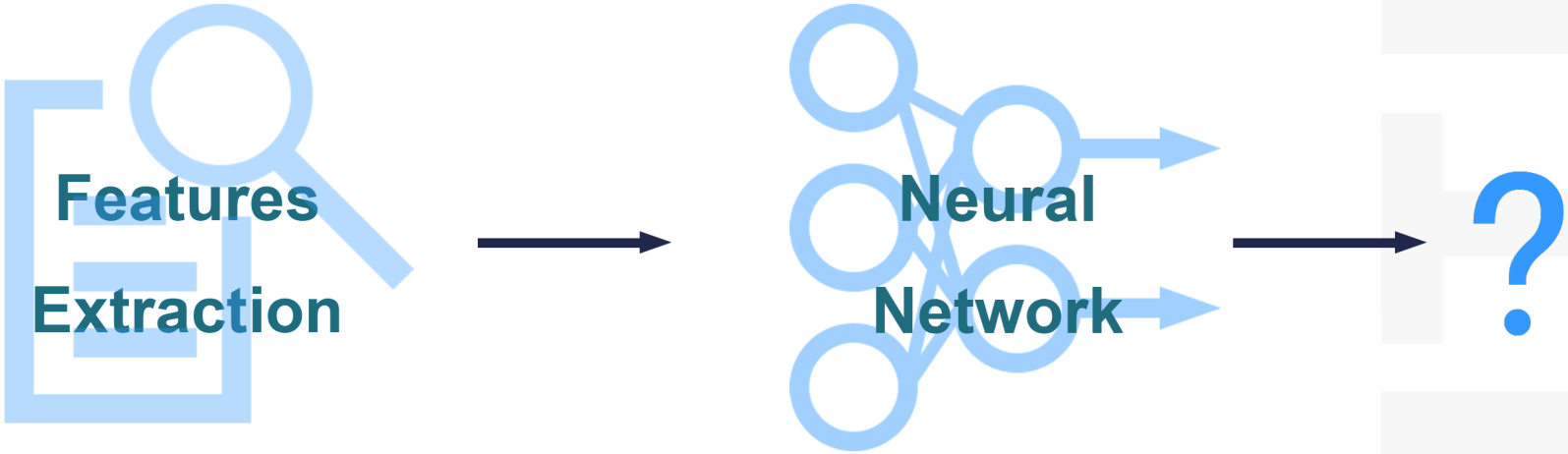
Dataset

TwiBot-20 Benchmark

"TwiBot-20: A Comprehensive Twitter Bot Detection Benchmark."
In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (CIKM).

The first publicly available bot detection benchmark with user follow relations to support graph based approaches. (to the best of our knowledge)
TwiBot-20 contains 229,573 users, 33,488,192 tweets, 8,723,736 user property items and 455,958 follow relationships

Baselines



- User profile, ICMSM 2011
- Social networks, ISONAM 2017
- Timeline of accounts, IEEE Intelligent Systems, 2016
- Redirection of URLs, TDSC 2013
- Classification of websites, S&P 2011

...

- RNN, TPS-ISA 2019
- RNN + Property, Information Sciences, 2018
- GAN, IJCAI 2019
- GCN, WWW 2019

...

BotRGCN Performance

TABLE III
BOT DETECTION PERFORMANCE ON TWIBOT-20 BENCHMARK.

Method	Accuracy	F1-score	MCC [14]
Lee *et al.* [15]	0.7456	0.7823	0.4879
Yang *et al.* [16]	0.8191	0.8546	0.6643
Kudugunta *et al.* [8]	0.8174	0.7517	0.6710
Wei *et al.* [7]	0.7126	0.7533	0.4193
Miller *et al.* [4]	0.4801	0.6266	-0.1372
Cresci *et al.* [17]	0.4793	0.1072	0.0839
Botometer [18]	0.5584	0.4892	0.1558
Alhosseini *et al.* [19]	0.6813	0.7318	0.3543
SATAR [20]	0.8412	0.8642	0.6863
BotRGCN	**0.8462**	**0.8707**	**0.7021**

BotRGCN Feature Study

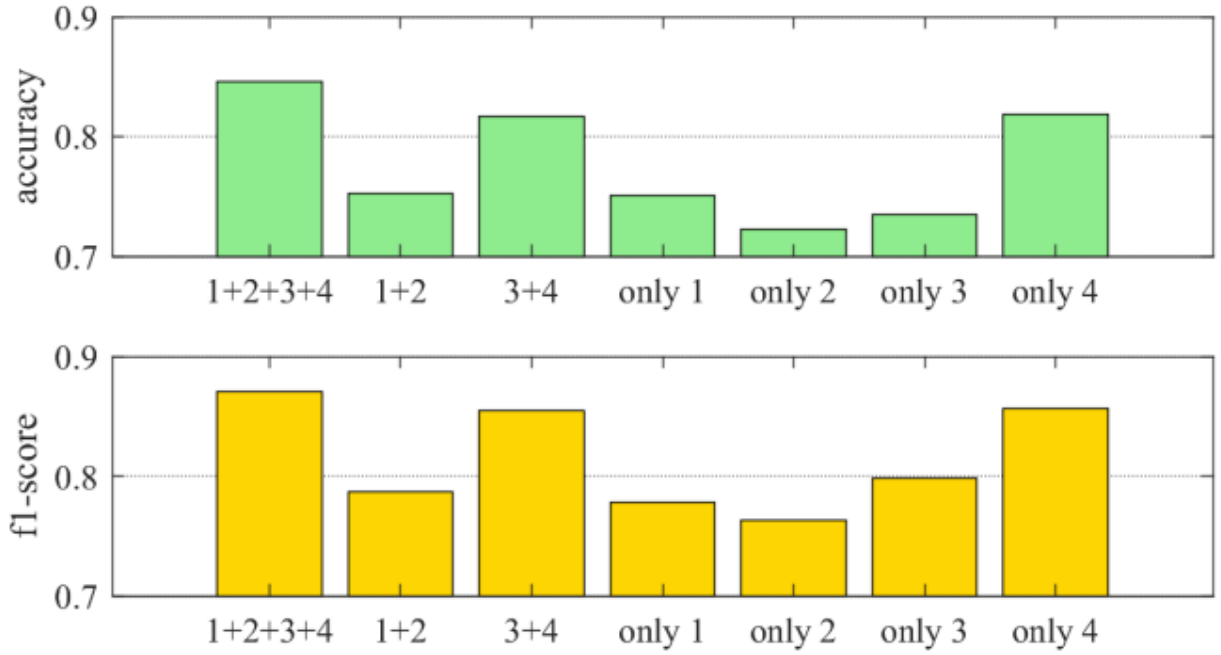


Fig. 2. BotRGCN performance with different user feature sets.

1: description 2: tweets 3: numerical metadata 4: categorical metadata

BotRGCN Graph Study

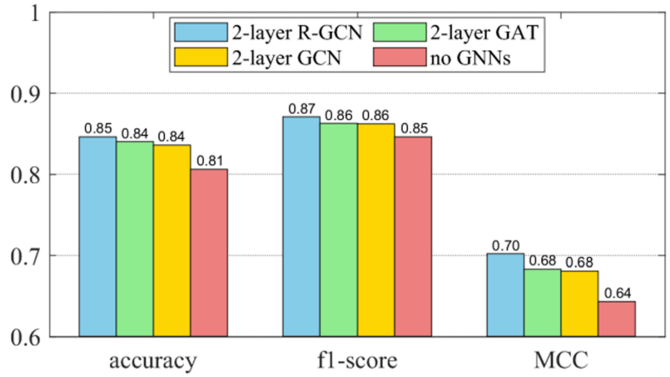


Fig. 3. BotRGCN performance with different types of graph neural networks.

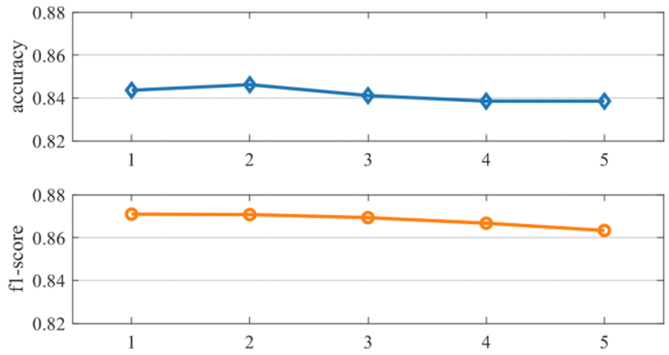


Fig. 4. BotRGCN performance with different number of R-GCN layers.

which proves the effectiveness of our graph-based approach.

Conclusion

We proposed BotRGCN, a novel bot detection framework that is

- **graph-based**
- **community-aware**
- **disguise-revealing**
- **semi-supervised**

BotRGCN achieves best performance on a comprehensive benchmark TwiBot-20.

Resources

BotRGCN code

https://github.com/BunsenFeng/BotRGCN

BotRGCN data: TwiBot-20 benchmark

"TwiBot-20: A Comprehensive Twitter Bot Detection Benchmark." In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (CIKM).

The Bot Repository

https://botometer.osome.iu.edu/bot-repository/

Thank You !

Herun Wan

LUD Lab, Xi'an Jiaotong University

wanherun@stu.xjtu.edu.cn

November 11, 2021