

TwiBot-20: A Comprehensive Twitter Bot Detection Benchmark

Ningnan Wang

LUD Lab, Xi'an Jiaotong University

mrwangyou@stu.xjtu.edu.cn

September 26, 2021

Table of Contents

1 Introduction

2 TwiBot-20 Construction

3 Experiments

4 Conclusion & Data Access

What is a Twitter bot?

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**.



Why should we detect Twitter bots?

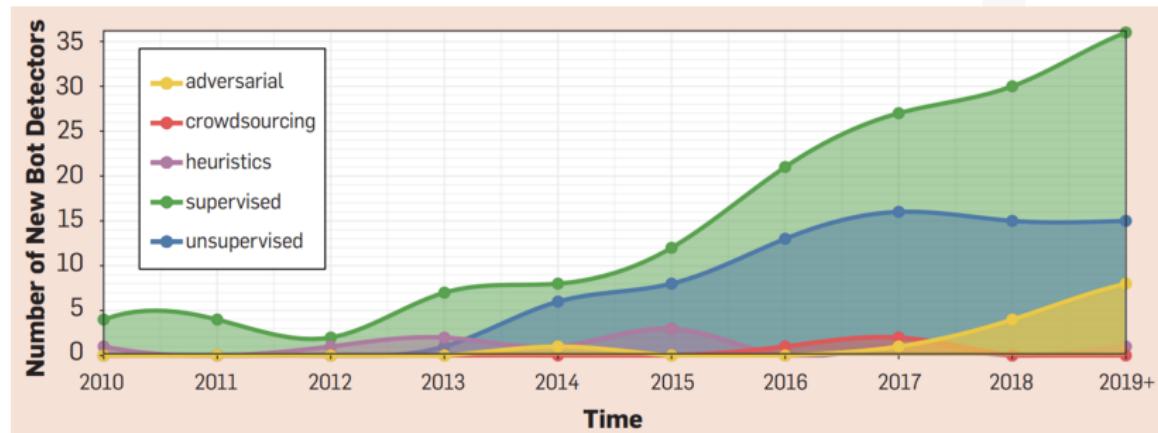
Twitter bots are responsible for:

- 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 Detection Methods

The majority of existing bot detection measures are **supervised**, which rely on annotated data sets.



Cresci et al., Communications of the ACM 2020

Existing Bot Detection Data Sets

The bot repository¹ keeps track of 18 publicly available bot detection data sets.

However, existing data sets suffer from:

- **user diversity.** different types of users and bots
- **limited user information.** multi-modal user information
- **data scarcity.** models are increasingly complex

We propose Twibot-20 to address such challenges.

1 <https://botometer.osome.iu.edu/bot-repository/>

Overview

Twibot-20 was constructed from July to September 2020, following the process:

- select diversified users from Twitter
- retrieve multi-modal user information
- derive trustworthy annotations

User Selection

We select **seed users** from politics, entertainment, business and sports to serve as starting points of controlled BFS.

Algorithm 1: Twibot-20 User Selection Strategy

Input: initial seed user u_0 in a user cluster

Output: user information set F

```

 $u_0.layer \leftarrow 0;$  // designate seed user as layer 0
 $S \leftarrow \{u_0\};$  // set of users to expand
 $u_0.expanded \leftarrow False;$ 
 $F \leftarrow \emptyset;$ 
while  $S \neq \emptyset$  do
     $u \leftarrow S.pop();$  // expand with user  $u$ 
     $T(u) \leftarrow get\_tweet(u);$ 
     $P(u) \leftarrow get\_property(u);$ 
    if  $u.layer \geq 3$  or  $u.expanded == True$  then
         $F \leftarrow F \cup u(T, P, N = \emptyset);$ 
        continue; // three layers max
     $u.expanded \leftarrow True;$ 
     $N^f(u) \leftarrow get\_friend(u);$ 
     $N^t(u) \leftarrow get\_follower(u);$ 
     $N(u) \leftarrow \{N^f(u), N^t(u)\};$ 
     $F \leftarrow F \cup u(T, P, N);$ 
     $S \leftarrow S \cup N^f(u) \cup N^t(u);$ 
    for  $y \in N^f(u) \cup N^t(u)$  do
         $y.expanded \leftarrow False;$ 
         $y.layer \leftarrow u.layer + 1;$ 

```

Return F ; // obtained one cluster of user information

User Information

We retrieve three aspects of user information on Twitter:

- semantic. the most recent 200 tweets for each user
- property. all property items accessible by Twitter API
- neighborhood. follow relationship between users

Data Annotation

We launch a crowdsourcing campaign and take the following steps to reach the final annotation of users in TwiBot-20:

- Firstly, if a user is **verified** by Twitter, we consider it to be a genuine user.
- For remaining users, if **four out of five annotators** believe that it is bot or human, we annotate the user accordingly.
- For other users with less mutual agreement from crowdsourcing, we use Twitter's **direct message** feature to send out simple questions in natural language, collect answers from users that respond and manually determine their annotations.
- Finally, remaining undecided users are **manually examined** within our research team. To ensure the trustworthiness of these ambiguous users, we **discard disputed cases** and only annotate when we reach a consensus on a Twitter user.

How does TwiBot-20 address these challenges?

TwiBot-20 is designed to tackle the challenge of _____ by:

- **user diversity.** seed users from multiple interest domains
- **limited user information.** retrieve semantic, property, and neighborhood user information
- **data scarcity.** controlled BFS for user collection

Next, we examine whether TwiBot-20 successfully addresses these challenges.

Challenge: User Diversity

Users in TwiBot-20 have diversified locations and interests.

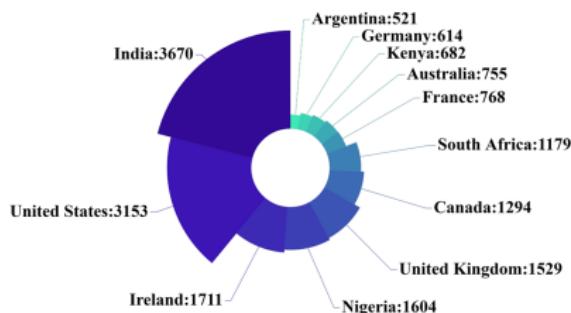


Figure 2: Most frequent countries of user location in TwiBot-20 and their number of appearances. There are 179 countries that appear less than 500 times and they collectively appear 11835 times in TwiBot-20, which are omitted in the figure to preserve clarity.

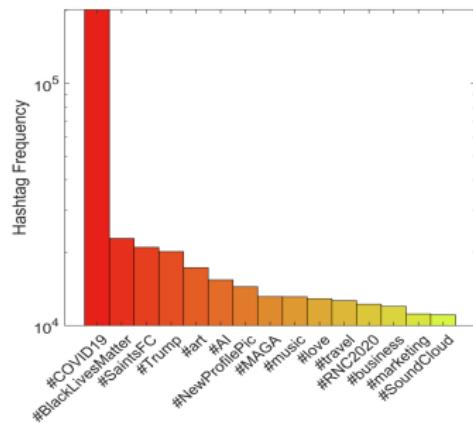
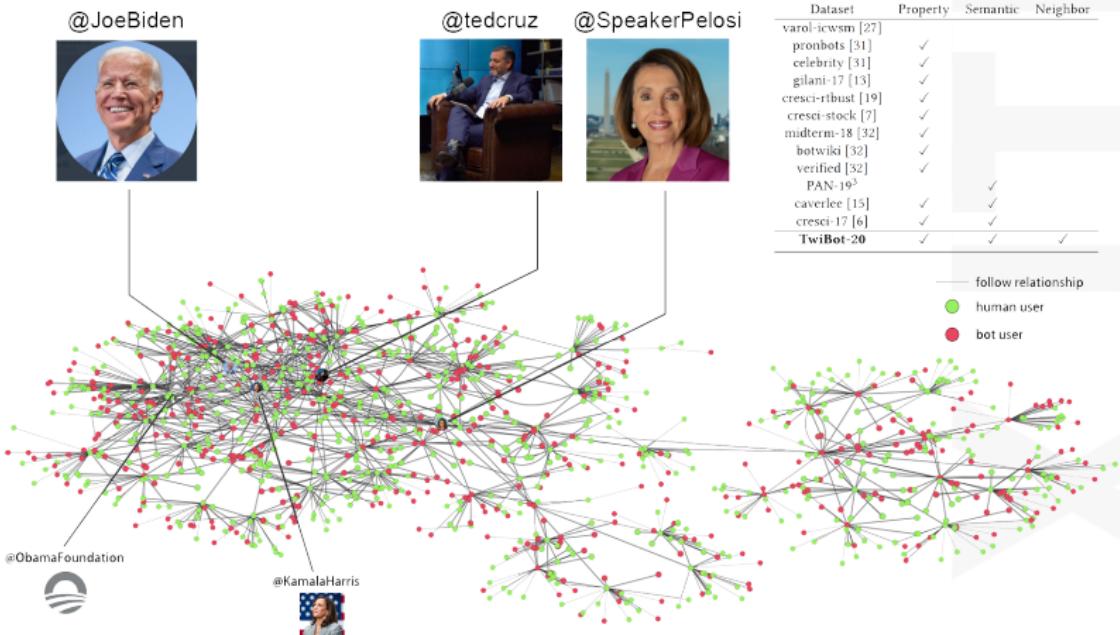


Figure 3: Most frequent hashtags in user tweets in our proposed TwiBot-20. We merge similar hashtags such as #COVID19 and #coronavirus and select the 15 hashtags of highest frequency.

Geographic Diversity

Challenge: Limited User Information

Twibot-20 is the first to provide neighborhood user information.



Challenge: Data Scarcity

Twibot-20 establishes the largest benchmark to date.

Table 1: Statistics of different bot detection benchmarks, from left to right, user count, user property item count, total tweet count and follow relationship count in each dataset.

Dataset	#User	#Property	#Tweet	#Follow
varol-icwsm [27]	2,573	0	0	0
pronbots [31]	21,965	750,991	0	0
celebrity [31]	5,971	879,954	0	0
gilani-17 [13]	2,653	104,515	0	0
cresci-rtburst [19]	693	28,968	0	0
cresci-stock [7]	13,276	551,603	0	0
midterm-18 [32]	50,538	909,684	0	0
botwiki [32]	698	29,082	0	0
verified [32]	1,987	83,383	0	0
PAN-19 ³	11,568	0	369,246	0
caverlee [15]	22,224	155,568	5,613,166	0
cresci-17 [6]	14,398	547,124	18,179,186	0
Twibot-20	229,573	8,723,736	33,488,192	455,958

Benchmarking

Twibot-20 sufficiently benchmarks existing methods and presents new challenges.

Table 4: The overall Twitter bot detection performance of various methods on our proposed Twibot-20 and two public datasets, Cresci-17 [6] and PAN-19³. “/” denotes that the dataset doesn’t have sufficient user information to support the method.

		Lee <i>et al.</i> [16]	Yang <i>et al.</i> [32]	Kudugunta <i>et al.</i> [14]	Wei <i>et al.</i> [29]	Miller <i>et al.</i> [21]	Cresci <i>et al.</i> [5]	Botometer [9]	Alhosseini <i>et al.</i> [1]
Twibot-20	Acc	0.7456	0.8191	0.8174	0.7126	0.4801	0.4793	0.5584	0.6813
	F1	0.7823	0.8546	0.7517	0.7533	0.6266	0.1072	0.4892	0.7318
	MCC	0.4879	0.6643	0.6710	0.4193	-0.1372	0.0839	0.1558	0.3543
Cresci-17	Acc	0.9750	0.9847	0.9799	0.9670	0.5204	0.4029	0.9597	/
	F1	0.9826	0.9893	0.9641	0.9768	0.4737	0.2923	0.9731	/
	MCC	0.9387	0.9625	0.9501	0.9200	0.1573	0.2255	0.8926	/
PAN-19 ³	Acc	/	/	/	0.9464	/	0.8797	/	/
	F1	/	/	/	0.9448	/	0.8701	/	/
	MCC	/	/	/	0.8948	/	0.7685	/	/

Conclusion

We present TwiBot-20, a comprehensive Twitter bot detection benchmark that:

- addresses the issues of user diversity, limited user information and data scarcity in previous datasets
- establishes the largest bot detection benchmark to date
- is the first to provide follow relationship to enable graph-based approaches
- sufficiently evaluates existing measures and presents new challenges to the task of bot detection

Data Access

We make TwiBot-20 publicly available to facilitate research efforts.
For a TwiBot-20 sample,

- TwiBot-20 github repo:

<https://github.com/BunsenFeng/TwiBot-20>

- bot repository:

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

Please contact wind_binteng@stu.xjtu.edu.cn for the complete TwiBot-20 dataset.

Thank You !

Ningnan Wang

LUD Lab, Xi'an Jiaotong University

mrwangyou@stu.xjtu.edu.cn

September 26, 2021