Don’t @ Me

Hunting Twitter Bots at Scale
Jordan Wright
Twitter: @jw_sec

Olabode Anise
Twitter: @JustSayO
The Problem
Justin livingston
@Justinl2233
Joined August 2016

Tweet to Justin livingston

Tweets

<table>
<thead>
<tr>
<th>Time</th>
<th>Text</th>
<th>Likes</th>
<th>Retweets</th>
</tr>
</thead>
<tbody>
<tr>
<td>21h</td>
<td>Get more twitter followers instantly. getmorefollowersnow.com</td>
<td></td>
<td>1.9K</td>
</tr>
<tr>
<td>22h</td>
<td>Get more twitter followers sent to your account. getmorefollowersnow.com</td>
<td></td>
<td>2.1K</td>
</tr>
</tbody>
</table>

Tweets & replies
Bots aren’t always malicious
TechCrunch
@TameWaylan

📍 Joined May 2018

Tweet to TechCrunch

<table>
<thead>
<tr>
<th>Tweets</th>
<th>Tweets &amp; replies</th>
</tr>
</thead>
</table>
| TechCrunch @TameWaylan · Jul 15 | Replying to @TameWaylan
Enter NOW! Don't miss it! |
| TechCrunch @TameWaylan · Jul 15 | Replying to @TameWaylan
You can use any wallet or exchange |
| TechCrunch @TameWaylan · Jul 15 | Replying to @TameWaylan
👇 More info in the link above 👇 |
Other Bot Hunters

@r0zetta  @DanielGallagher
@3r1nG  @geoffgolberg
We can help solve this problem as a community.
Gathering the Dataset

- Finding accounts
- Enriching the data
- All as efficiently as possible

Analyzing the Data at Scale

- Identify features
- Build a classifier
- Run the classifier over the dataset
Gathering the Dataset
Finding Accounts

- Finding user accounts
  - Screen name
  - User ID
  - Tweet Count
  - Followers/Friends Count

Enriching the Dataset

- Tweets
- Followers/Friends
Finding Accounts
Fetching Accounts

- users/lookup
- 100 account IDs per request
- 900 requests / 15 minute window
- Returns full user objects
Jordan Wright
@jw_sec
R&D @duo_labs
🔗 jordan-wright.com
Joined December 2014
8.6 million lookups per day
32-Bit Account ID’s

- Sequential 32-bit unsigned integers
- Enumerated a random 5% sampling
33-Bit Account IDs

- Observed accounts up to $5 \times 10^6$ ID range
- Creation dates up to 2016
Introduction of Snowflake ID’s

- Generated by workers
- Same format for all Twitter objects
- **Hard** to enumerate
Snowflake ID Format

- 63-bit

1019674853719228416

0001110001010011001110011100000100000100101010100000000000000
Snowflake ID Format

- 63-bit

- Timestamp (41 bits)
  - Custom Epoch (November 4, 2010)

1019674853719228416

0001110001010100110011100011000011000000100010101010101000000000000000
Snowflake ID Format

- 63-bit
- Timestamp (41 bits)
- Worker Number (10 bits)
Snowflake ID Format

- 63-bit
- Timestamp (41 bits)
- Worker Number (10 bits)
- Sequence Number (12 bits)
Snowflake ID Format

- 63-bit

1019674853719228416

0001110001010011100000100101000000000000000000000

It’s **hard** to guess these values.
Using the Streaming API

- Random sample of public statuses
- Contains full user object
- Filter on keyword, location, and more

- Introduces bias into the dataset
Enriching the Dataset
Fetching Tweets

- Latest 200 tweets
- Accounts with more than 10 tweets
- Declared English accounts
- Fetched original tweets
Mapping the Social Network
System Architecture
## System Speed

<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounts</td>
<td>8.6 million per day</td>
</tr>
<tr>
<td>Tweets</td>
<td>144,000 per day</td>
</tr>
<tr>
<td>Network</td>
<td>1,440 per day</td>
</tr>
</tbody>
</table>
Results

88m Accounts

3.1m Accounts with Tweets

576m Total Number of Tweets
Anatomy of a Twitter Bot
Types of Characteristics

Account Attributes  Tweet Content  Tweet Behavior
Heuristics - Account Attributes

- Entropy of the username
- Numbers at the beginning and end of the screen name
- Ratio of followers to following
- Number of tweets and likes relative to account age
Heuristics - Tweet Content

- Number of hashtags in tweets
- Number of urls in tweets
- Number of users mentioned in tweets
- Percentage of tweets with media
Heuristics - Tweet Behavior

- Number of unique sources
- Time to reply/retweet
- Unique users retweeted
- Distance between geo-location enabled tweets
- Distribution of tweet times
Many bots try to look human.
Types of Twitter Bots

- Spam Bots
- Fake Followers
- Amplification Bots
Classifying Bots at Scale
Data Science 101 - Process

- Gather or acquire data
- Build features
- Split the data
- Evaluate different algorithms
- Test the best model on unseen data
Data Science 101 - Terminology

- Models and Algorithms
  - Decision Trees, Random Forests, and AdaBoost
- Precision & Recall
- Cross-Validation
## Dataset

<table>
<thead>
<tr>
<th>Sets of Accounts</th>
<th>Number of Accounts</th>
<th>Number of Tweets</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genuine accounts</td>
<td>3,474</td>
<td>2,839,361</td>
<td>Labeled genuine user accounts</td>
</tr>
<tr>
<td>Social spam bots</td>
<td>4,912</td>
<td>3,457,133</td>
<td>Labeled bot accounts</td>
</tr>
<tr>
<td>Crypto-giveaway bots</td>
<td>7,250</td>
<td>92,515</td>
<td>Labeled bot accounts</td>
</tr>
</tbody>
</table>
Account Features

- Finding the best model
- Classifying accounts

Account + Tweet Feature

- Identifying important features
- Classifying accounts
Finding the Best Model

- Started with five models
- Split the data
- Used a variety of metrics
Finding the Best Model

- Started with five models
- Split the data
- Used a variety of metrics
Finding the Best Model

- Started with five models
- Split the data
- Used a variety of metrics
Finding the Best Model

- Started with five models
- Split the data
- Used a variety of metrics

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>AdaBoost Classifier</td>
<td>98.33%</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>97.75%</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>92.57%</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>83.57%</td>
</tr>
<tr>
<td>Random Forest</td>
<td>98.43%</td>
</tr>
</tbody>
</table>
Other Cross-Validation Metrics for Random Forest

99% Precision
97% Recall
Classification is hard.
There are a lot of Unknowns

- Calibration relies on estimates of number of bots
- Generalizable results require representative samples
- Optimal thresholds vary when the proportions of classes are different
Metrics for Test Set with Varying Amounts of Bots

92% Precision when estimating 10% of accounts are bots

94% Precision when estimating 15% of accounts are bots
Results using Account Attributes on Unknown Data

32% 1%

Accounts were more likely to be bots

Accounts with at least 90% chance to be bots
## Results using Account Attributes on Unknown Data

<table>
<thead>
<tr>
<th>Feature</th>
<th>Top 20%</th>
<th>Bottom 20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ratio_followers_friends</td>
<td>0.022</td>
<td>27.48</td>
</tr>
<tr>
<td>screen_name_entropy</td>
<td>2.94</td>
<td>2.93</td>
</tr>
<tr>
<td>numbers_at_beginning_of_screen_name</td>
<td>0.17</td>
<td>0.037</td>
</tr>
<tr>
<td>tweet_rate</td>
<td>0.22</td>
<td>10.91</td>
</tr>
<tr>
<td>favorite_rate</td>
<td>0.19</td>
<td>6.60</td>
</tr>
<tr>
<td>numbers_at_end_of_screen_name</td>
<td>2.09</td>
<td>0.48</td>
</tr>
</tbody>
</table>
Your model is only as good as your data

80% vs 53%

Accuracy predicting crypto-giveaway bots with model trained with social spam bots

Accuracy predicting social spam bots with model trained with crypto-giveaway bots
Adding content and behavior data
Identifying Important features

- 22 features across the three types of data
- Fewer features = less data, less complexity, and faster training
- Recursive feature elimination
Five Most Important Features

- is_default_profile
- avg_num_users_mentioned_in_tweets
- avg_distinct_hours_tweeted
- favorite_rate
- avg_reply_distance_id
Results on Unknown Dataset

21% Accounts were more likely to be bots

0.8% Accounts with at least 90% likelihood to be bots
Results on Unknown Dataset

<table>
<thead>
<tr>
<th>Feature</th>
<th>Top 20%</th>
<th>Bottom 20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>is_default_profile</td>
<td>67%</td>
<td>60%</td>
</tr>
<tr>
<td>avg_num_users_mentioned_in_tweets</td>
<td>0.67</td>
<td>0.88</td>
</tr>
<tr>
<td>avg_distinct_hours_tweeted</td>
<td>5.4</td>
<td>4.9</td>
</tr>
<tr>
<td>favorite_rate</td>
<td>4.84</td>
<td>22.85</td>
</tr>
<tr>
<td>avg_reply_id_distance</td>
<td>1.87e17</td>
<td>9.6e16</td>
</tr>
</tbody>
</table>
Classification is still hard with context.
More Data Helps

83% vs 62%

Accuracy predicting crypto bots with model trained with social spam bots vs Accuracy predicting social spam bots with model trained with crypto-giveaway bots
Limitations and Challenges
Context and Perspective

- Accounts change over time
- Network data tells the full story
- Insight into the costs of a false positive changes thresholds
Straightforward techniques do find bots.
Crawling Botnets

Gathering the Dataset

Identifying Twitter Bots

Crawling Botnets
SONM
@04ph1xsb0nNgXCIE
Joined June 2018

Tweets
5
Following
15
Followers
3

New to Twitter?
Sign up now to get your own personalized timeline!

Sign up

Worldwide trends

#FIFAxEXOPower
3,259 Tweets

#BTSxFifaWorldCup
2,118 Tweets

#NB17
178K Tweets

#نسر_النجم_الشام_البرازيل
17K Tweets

#FifaJuve
37.4K Tweets

Claude Lanzmann
19.4K Tweets

C.tolist
8,837 Tweets

Gizem Özdemir
7,018 Tweets

Polito
8,239 Tweets

Kukushkin
5,672 Tweets

© 2018 Twitter
About | Help Center | Terms
Privacy policy | Cookies | Ads info

To celebrate 10,000,000 worth of ETH transactions
We are giving back to the community with

10 000 ETH giveaway

Send 0.5 10 ETH to verify your address, and get 5 100 ETH back (limit is 100 ETH per address), far this tweet when you get yours!

10,000 ETH to Give Away! We will immediately send you 5-100 ETH
plus.google.com
To celebrate 10,000,000$ worth of ETH transactions
We are giving back to the community with

10 000 ETH giveaway

Click Here

GET IT HERE

In order to get 5-100 ETH
back on your wallet,

send us your shear link

plus.google.com

6:13 PM - 24 Sep 2017

Aгриппина Соболе...
@19dhvTPzspYd0R

Caution: This account is temporarily restricted
<table>
<thead>
<tr>
<th>Likes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stakenet @9njR8PefXw95mt · Jun 30</strong></td>
</tr>
<tr>
<td>Replying to @9njR8PefXw95mt</td>
</tr>
<tr>
<td>Enter NOW! Don’t miss it!</td>
</tr>
<tr>
<td>🗣️ 23</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Stakenet @9njR8PefXw95mt · Jun 30</strong></td>
</tr>
<tr>
<td>Replying to @9njR8PefXw95mt</td>
</tr>
<tr>
<td>You can use any wallet or exchange!</td>
</tr>
<tr>
<td>🗣️ 1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Stakenet @9njR8PefXw95mt · Jun 30</strong></td>
</tr>
<tr>
<td>Replying to @9njR8PefXw95mt</td>
</tr>
<tr>
<td>According to the rules you may take part only once!</td>
</tr>
<tr>
<td>🗣️ 1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Stakenet @9njR8PefXw95mt · Jun 30</strong></td>
</tr>
<tr>
<td>Replying to @9njR8PefXw95mt</td>
</tr>
</tbody>
</table>
The Bitter Battle to Turn an Old Factory Into a 21st Century ‘Eco-Vill...'
A bold plan to redevelop Ford’s Twin Cities Assembly Plant has become the focus of a fight over density, traffic, and the future of cities.

WIRED  @WIRED  -  Jul 16

Belgium beats England 2-0, finishing third place in the World Cup cnn.it/2zFhgrR

CNN  @CNN  -  Jul 14

We got good news!

CCN (Breaking NEWS)  @AlvinOsmond  -  Jul 14

iFixit finds dust covers in latest MacBook Pro keyboard

TechCrunch  @TechCrunch  -  Jul 15

To celebrate 10,000,000 ETH transactions.
We are giving back to the community with 10 000 ETH giveaway.

Twitter Link: twitter.com
Taking over verified accounts

Coinbase 🔄 @Se_1h
Replying to @Se_ @coinbase

You can use any exchange or wallet to make a transaction.

Left Ethereum - 5 000.
Left Bitcoin - 500.

All users can take part in this giveaway!
Yup.

There’s still some time to squeeze in more easter eggs.

I have one easter egg right here, in fact:
They even found Twitter recommending users follow other spam bots in the botnet under the “Who to follow” section in the sidebar. Ouch.

Duo Security researchers' Twitter 'bot or not' study unearthed...
A team of researchers at Duo Security has unearthed a sophisticated botnet operating on Twitter — and being used to spread a cryptocurrency scam. The botnet was discovered dur...
techcrunch.com

11:35 AM - 6 Aug 2018

8 Retweets 38 Likes

Jameson Lopp @GP3NJEbq4EioO1d · 2h
Replying to @lopp
As a thanks for your support, with "ETHEREUM", we have big share for all of you!
twitter.com

To celebrate $10,000,000 worth of ETH transactions
We are giving back to our community!

10 000 ETH giveaway on ethgift.net
Visit www.ethgift.net to participate on this giveaway!
Send 2-20 ETH to verify your address, and get 20-200 ETH back immediately! Favorite this tweet when you get yours!
A thread can be followed unraveling an entire botnet.
Sharing with Twitter
“Twitter is aware of this form of manipulation and is proactively implementing a number of detections to prevent these types of accounts from engaging with others in a deceptive manner. Spam and certain forms of automation are against Twitter’s rules…”
“...In many cases, spammy content is hidden on Twitter on the basis of automated detections. When spammy content is hidden on Twitter from areas like search and conversations, that may not affect its availability via the API. This means certain types of spam may be visible via Twitter's API even if it is not visible on Twitter itself...”
“…Less than 5% of Twitter accounts are spam-related.”

- Twitter
Sharing with the Community
Open Source

github.com/duo-labs/twitterbots
Get the Paper

[duo.sc/twitter-bots]
Summary

- Large Twitter datasets can be gathered quickly and efficiently.
- A practical data science approach is effective at finding bots across a large dataset
- Mapping social connections can unravel entire botnets
We can help solve this problem as a community.
Thank You!

GitHub: github.com/duo-labs/twitterbots  
Duo: duo.sc/twitter-bots

Twitter: @JustSayO  
@jw_sec