



Studying Twitter User Accounts: Spotting Suspicious Social Bot Behavior

Master candidate: Marcelo Rovai

Professors: Eduardo Graells-Garrido (Sup.)
Loreto Bravo
Leo Ferres

Santiago, December 2019

Motivation

Piñera encabeza la lista de candidatos con más seguidores "falsos" en Twitter

por Felipe Salas | 17 octubre, 2017

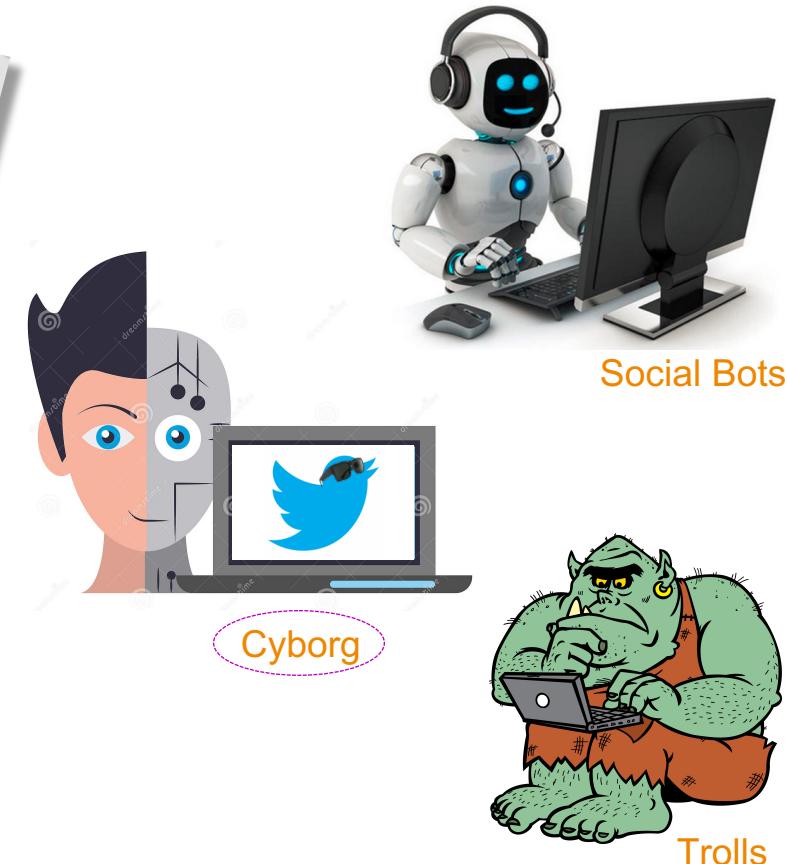
Após eleição, perfis falsos e robôs pró-estudo Bolsonaro continuam ativos, aponta estudo

México: los "bots", "trolls" y otros trucos de manipulación en internet que amenazan las próximas elecciones presidenciales

Marcos Martínez
BBC Monitoring

Michael Kozak:
Jefe de la diplomacia de EE.UU. para América Latina ve "actores extranjeros" en protestas en Chile

Bot-Like Accounts and Pro-Government Hashtags in Colombia
Suspicious accounts boosted pro-government hashtags in Colombia

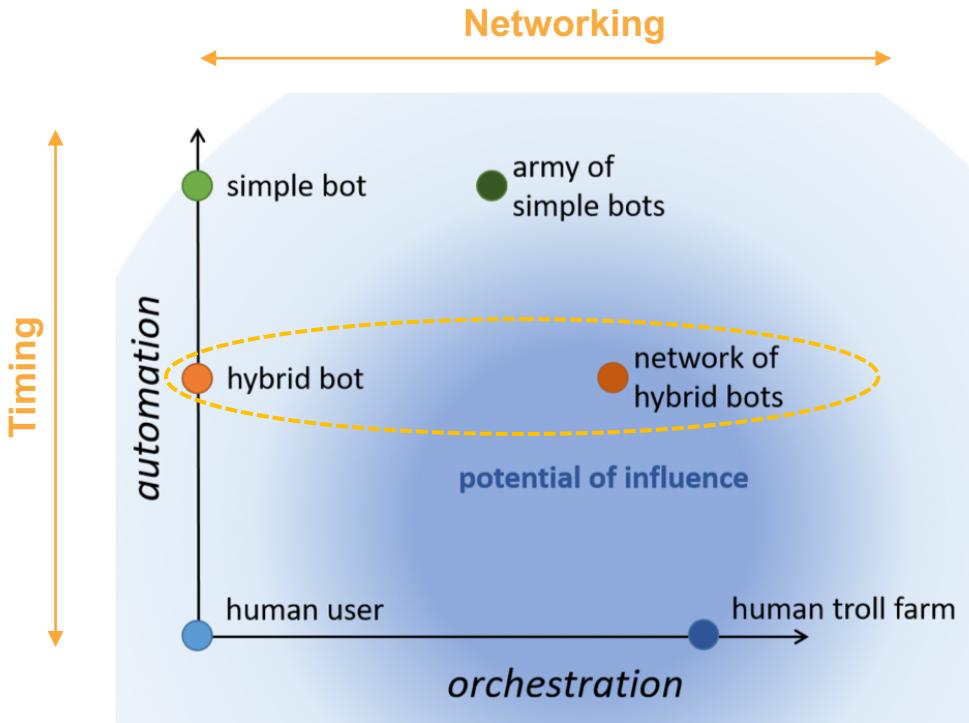


Social Bots

Cyborg

Trolls

Social Media – Potential Influence

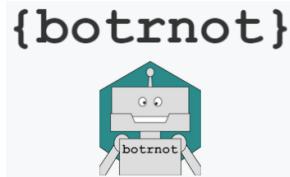


Qualitative classification of the potential influence of humans and bots in social media with respect to automation and orchestration (GRIMME et al. 2017)

Related Works



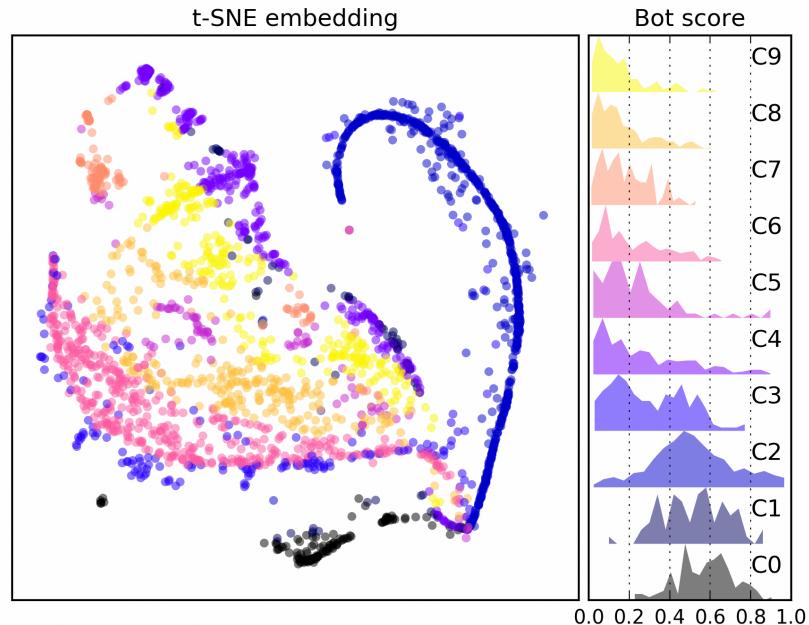
- APP and API (Python)
- Indiana University
- SML – Random Forest
- 1,000 features



- APP and API (R)
- University of Missouri
- SML – Gradient Boosted
- 100 features



- APP and API (JS)
- ITS Rio
- Score Calculator
- 40 features

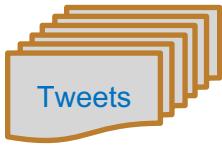


- Varol et al. 2017
- T-SNE Clusters
- 100 features

Objective

- This work aims to **identify user accounts** that could be **potential bots** based on their **behavior**
- **User behavior** is based on **tweet features**, retrieved from user meta-data, friends' information, content, network, and time-series data
- Each **feature** is studied individually and in **different subsets of data**, such as:
 - High-frequency vs low-frequency tweet users
 - Simple bots
 - Younger Mid-Frequency accounts
 - Clusters of users with similar behavior (UMAP)

Main Dataset



- The source of this study is an **unstructured tweet dataset** captured during the last Chilean presidential election (2017) by Eduardo Graells-Garrido (UDD/BSC)
- The raw dataset with a size of **7.5Gb** is a text file that was imported primarily in a JSON format
- The dataset covers **30 days (full November) of 1.6M tweets** generated "around" the first round of 2017's elections (November 19, 2017), posted by **92K distinct users**

Methodology

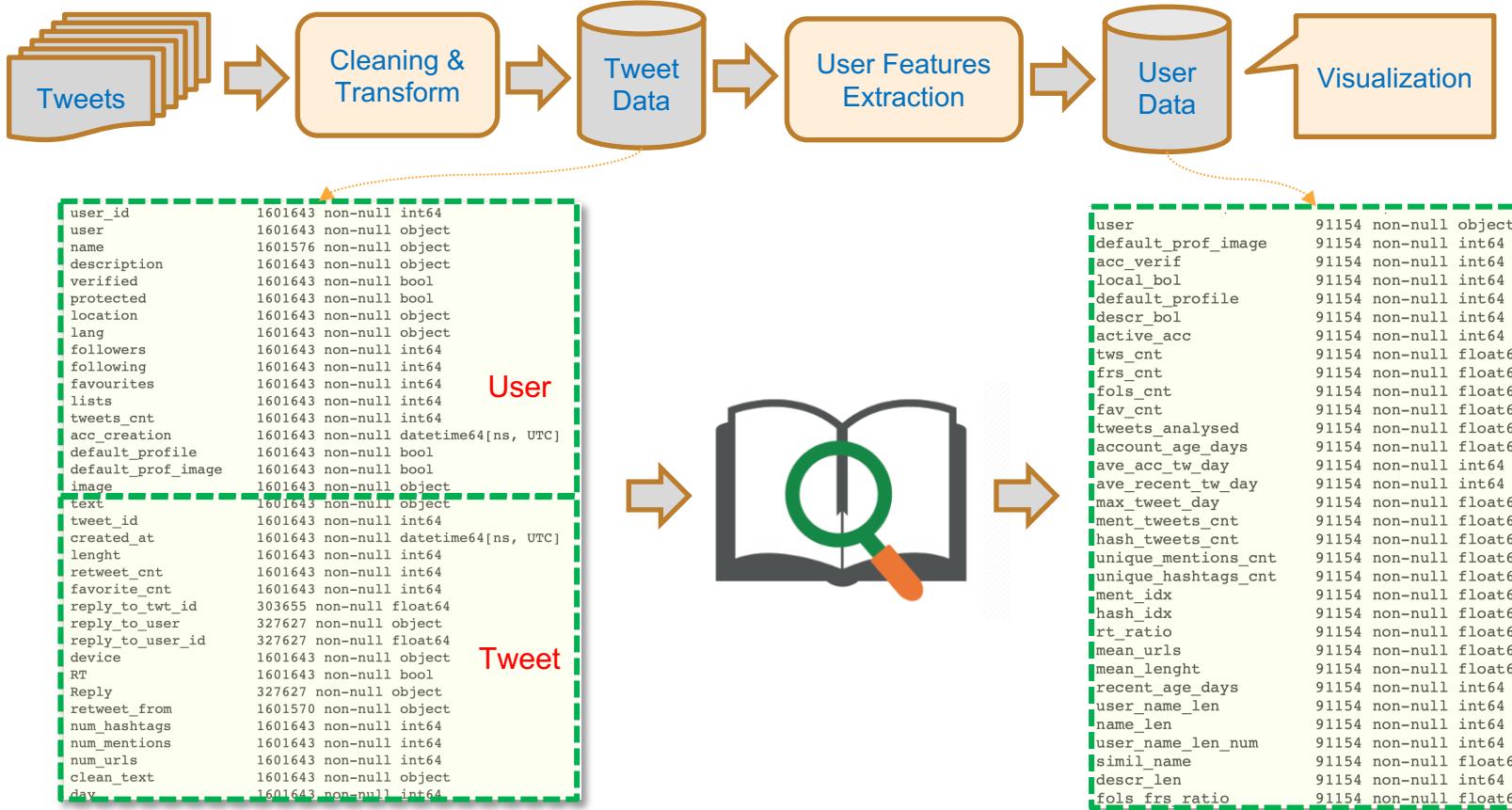


user_id	1601643	non-null	int64
user	1601643	non-null	object
name	1601576	non-null	object
description	1601643	non-null	object
verified	1601643	non-null	bool
protected	1601643	non-null	bool
location	1601643	non-null	object
lang	1601643	non-null	object
followers	1601643	non-null	int64
following	1601643	non-null	int64
favourites	1601643	non-null	int64
lists	1601643	non-null	int64
tweets_cnt	1601643	non-null	int64
acc_creation	1601643	non-null	datetime64[ns, UTC]
default_profile	1601643	non-null	bool
default_prof_image	1601643	non-null	bool
image	1601643	non-null	object
text	1601643	non-null	object
tweet_id	1601643	non-null	int64
created_at	1601643	non-null	datetime64[ns, UTC]
length	1601643	non-null	int64
retweet_cnt	1601643	non-null	int64
favorite_cnt	1601643	non-null	int64
reply_to_twt_id	303655	non-null	float64
reply_to_user	327627	non-null	object
reply_to_user_id	327627	non-null	float64
device	1601643	non-null	object
RT	1601643	non-null	bool
Reply	327627	non-null	object
retweet_from	1601570	non-null	object
num_hashtags	1601643	non-null	int64
num_mentions	1601643	non-null	int64
num_urls	1601643	non-null	int64
clean_text	1601643	non-null	object
day	1601643	non-null	int64

User

Tweet

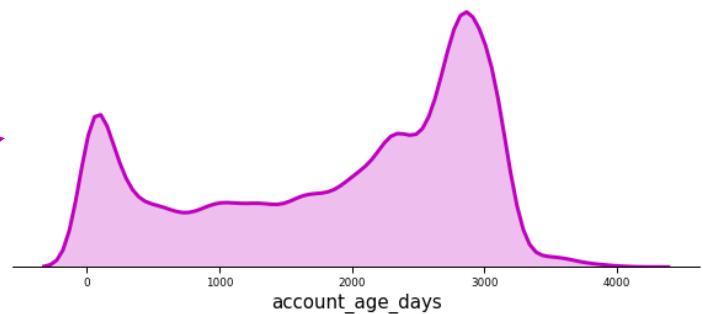
Methodology



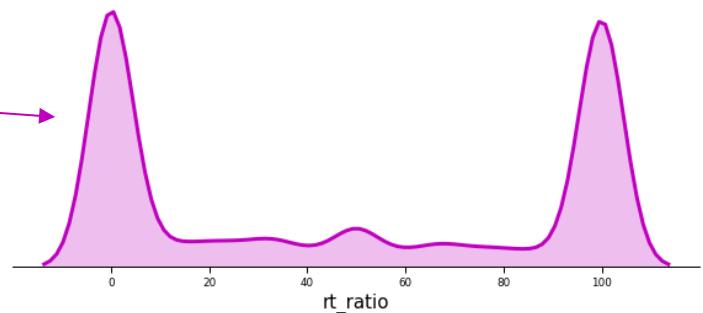
User Dataset Exploration

DF USER DESCRIPTION									
Feature	count	mean	std	min	25%	50%	75%	max	
default_prof_image	91,154	0.1	0.3	0	0.0	0.0	0.0	1.0	
acc_verif	91,154	0.0	0.0	0	0.0	0.0	0.0	1.0	
local_bol	91,154	0.6	0.5	0	0.0	1.0	1.0	1.0	
default_profile	91,154	0.5	0.5	0	0.0	1.0	1.0	1.0	
descr_bol	91,154	0.7	0.5	0	0.0	1.0	1.0	1.0	
active_acc	91,154	0.9	0.3	0	1.0	1.0	1.0	1.0	
tws_cnt	91,154	9,198.7	28,376.3	1	231.0	1,406.0	6,627.5	1,386,920.0	
frs_cnt	91,154	753.6	4,236.8	0	106.0	302.0	775.0	761,106.0	
fols_cnt	91,154	1,599.4	29,394.8	0	33.0	141.0	499.8	2,882,527.0	
fav_cnt	91,154	2,315.7	8,656.2	0	55.0	337.0	1,556.8	897,789.0	
tweets_analysed	91,154	164	165.0	1	1.0	2.0	7.0	25,900.0	
account_age_days	91,154	1,889.8	1,058.6	0	978.0	2,241.0	2,819.0	4,065.0	
ave_acc_tw_day	91,154	5	22.4	0	0.0	1.0	3.0	1,246.0	
ave_recent_tw_day	91,154	1.7	12.9	0	0.0	1.0	1.0	1,669.0	
max_tweet_day	91,154	5.8	53.8	0	1.0	1.0	3.0	8,044.0	
ment_tweets_cnt	91,154	12.4	96.0	0	1.0	2.0	6.0	15,603.0	
hash_tweets_cnt	91,154	4.4	53.6	0	0.0	0.0	1.0	7,140.0	
unique_mentions_cnt	91,154	10.2	28.6	0	1.0	3.0	8.0	1,342.0	
unique_hashtags_cnt	91,154	1.7	5.5	0	0.0	0.0	1.0	220.0	
ment_idx	91,154	1.6	1.5	0	1.0	1.2	2.0	49.0	
hash_idx	91,154	0.5	0.7	0	0.0	0.0	1.0	16.0	
rt_ratio	91,154	48.5	44.4	0	0.0	50.0	100.0	100.0	
mean_urls	91,154	0.2	0.4	0	0.0	0.0	0.3	3.0	
mean_lenght	91,154	134.9	45.4	14	116.0	132.0	140.0	987.0	
recent_age_days	91,154	8.2	9.6	1	1.0	1.0	16.0	30.0	
user_name_len	91,154	11.1	2.7	3	9.0	11.0	13.0	15.0	
name_len	91,154	13.2	4.8	0	10.0	14.0	17.0	50.0	
user_name_len_num	91,154	0.8	1.8	0	0.0	0.0	1.0	14.0	
simil_name	91,154	63.1	27.6	0	41.7	63.6	88.9	100.0	
descr_len	91,154	58.5	56.2	0	0.0	46.0	108.0	179.0	
fols_frs_ratio	91,154	3.1	120.4	0	0.2	0.5	0.9	22,797.0	

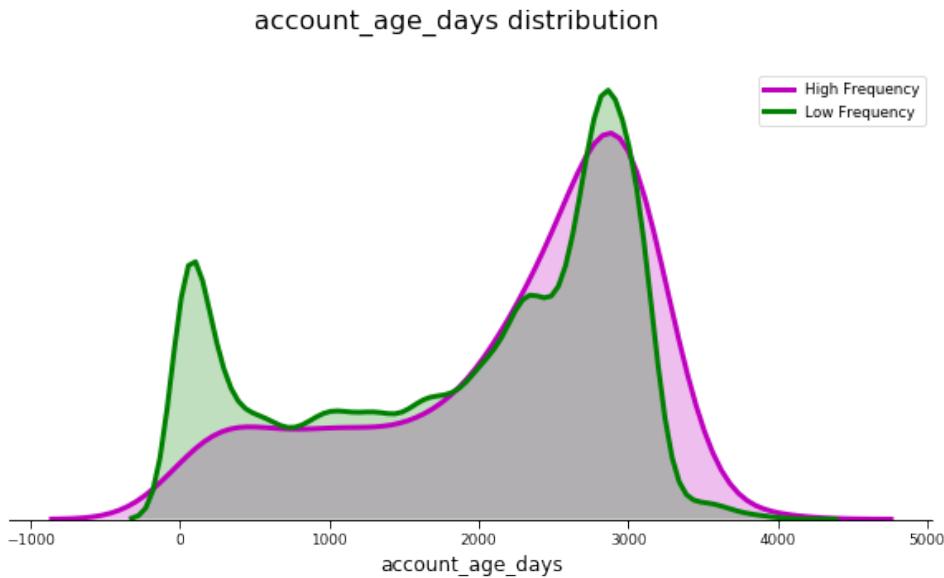
account_age_days distribution



rt_ratio distribution



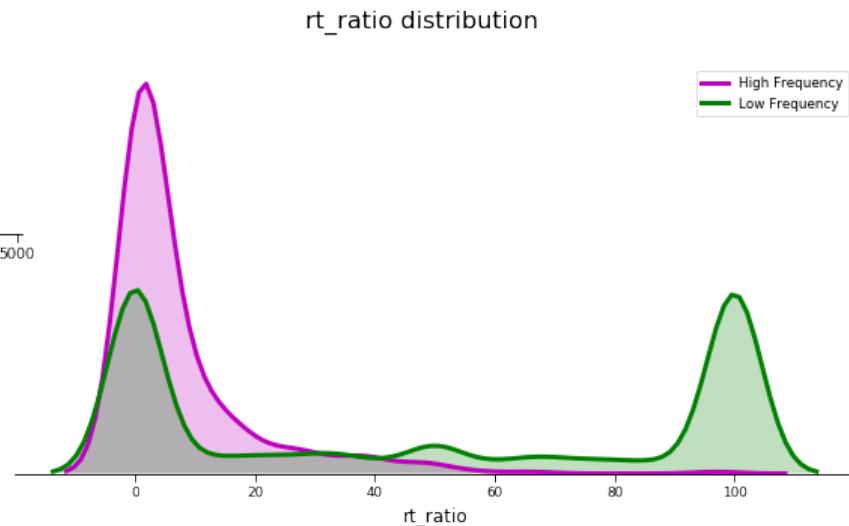
High/Low-Frequency User EDA



→ 356 users* → 25% tweets
→ 90,000 users → 52% tweets

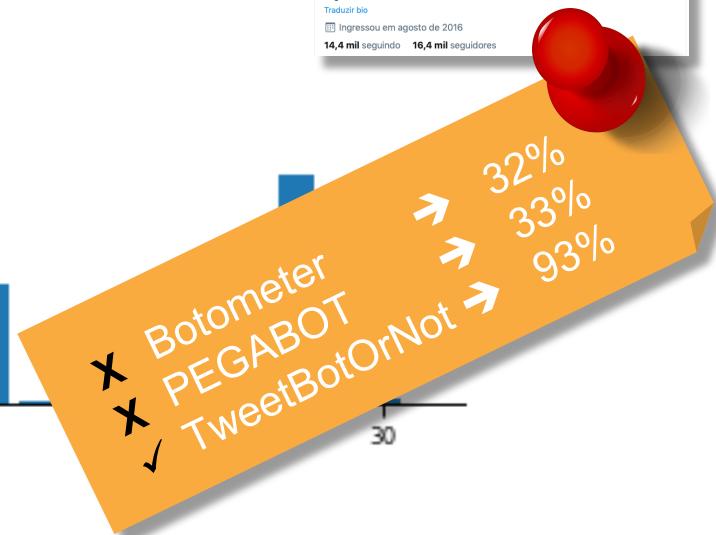
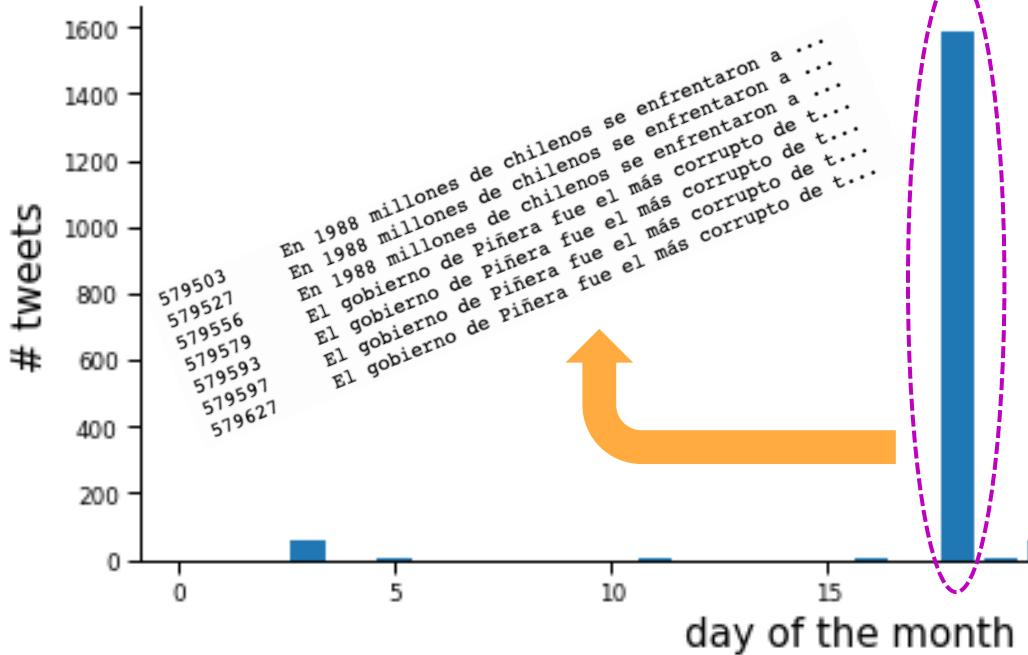
* Non-Verified

→ High-Frequency → > 144 tweets/day
→ Low-Frequency → < 50 tweets/day



Spotting High/Low-Frequency Bots

@andres20ad - Number of Tweets per day

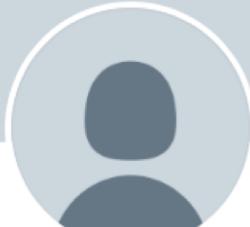


Spotting Simple Bots

fedoraletelier
410,1 mil Tweets

100K tweets/year

No Photo



fedoraletelier
@fedoraletelier

Ingressou em abril de 2014

3.085 seguindo 3.575 seguidores

Seguido por El Mercurio e Soledad Alvear

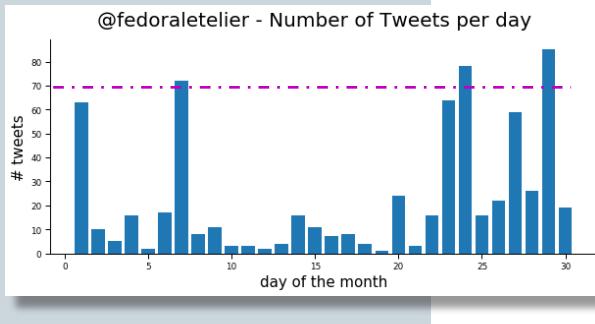
No Verified

No Description

High Network

... Seguir

@fedoraletelier - Number of Tweets per day

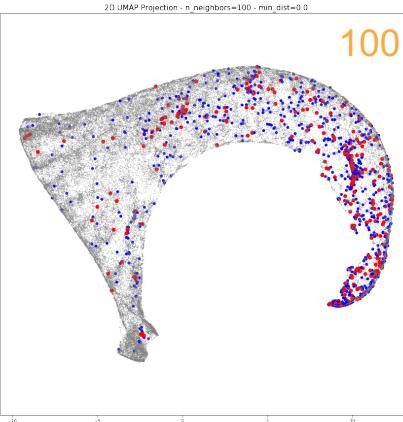
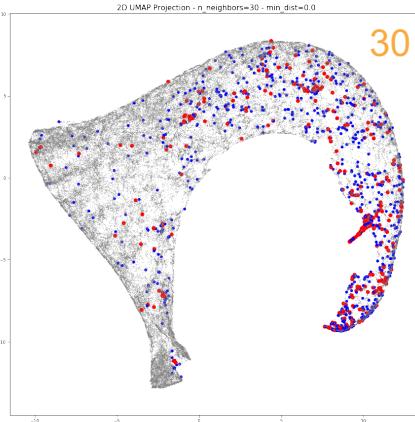
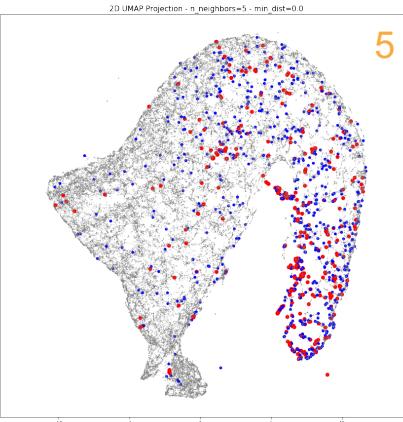
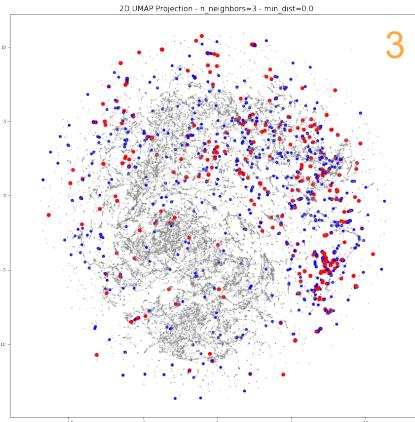


Day of the Month	# tweets
0	65
1	10
2	5
3	15
4	10
5	5
6	15
7	10
8	70
9	10
10	5
11	5
12	5
13	5
14	15
15	10
16	5
17	5
18	5
19	5
20	25
21	5
22	15
23	20
24	75
25	20
26	25
27	20
28	25
29	20
30	20

X Botometer
✓ PEGABOT
✓ TweetBotOrNot → 30%
→ 100%
→ 99%

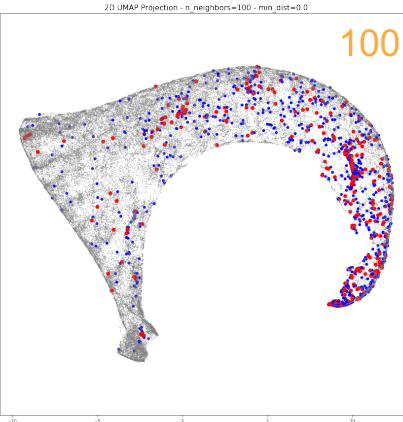
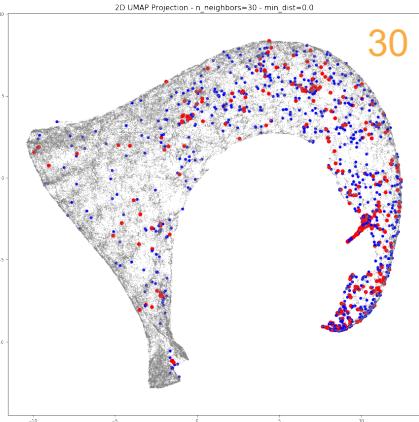
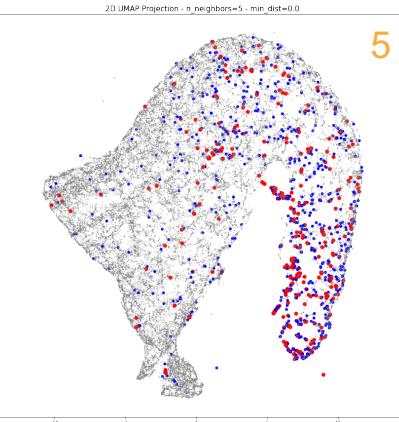
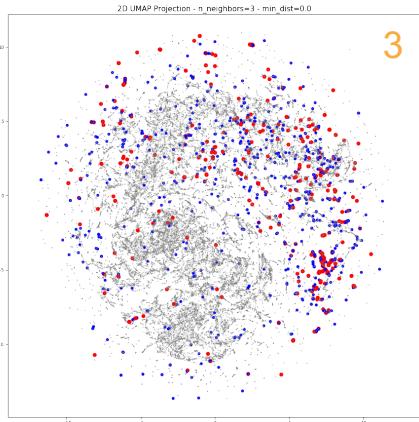
UMAP Hyperparameters Selection

component = 2, min_dist = 0 n_neighbors =

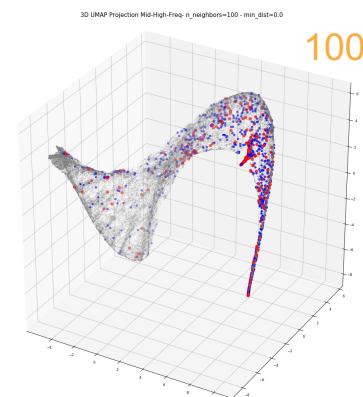
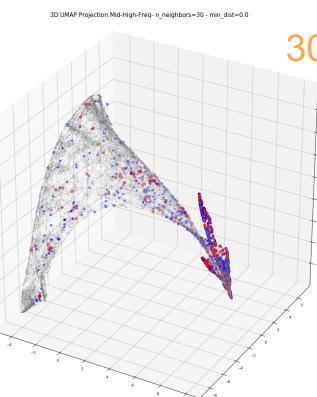
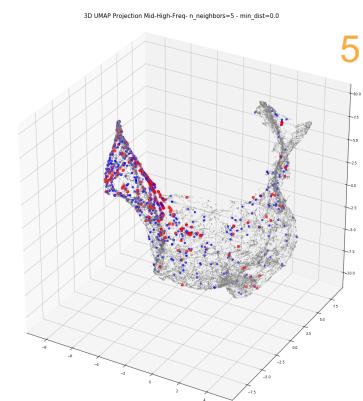
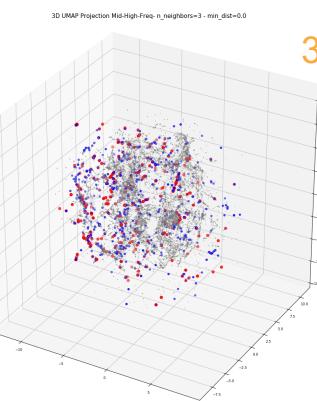


UMAP Hyperparameters Selection

component = 2, min_dist = 0 n_neighbors =

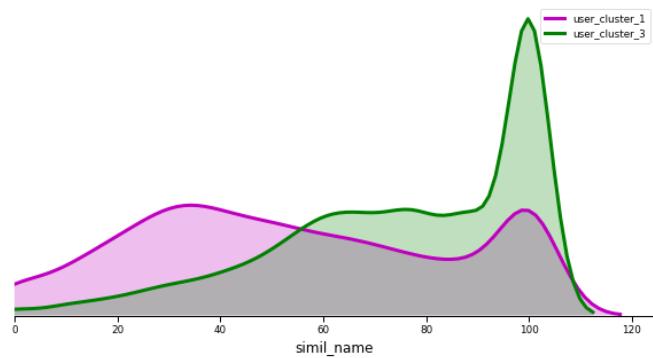


component = 3, min_dist = 0 n_neighbors =

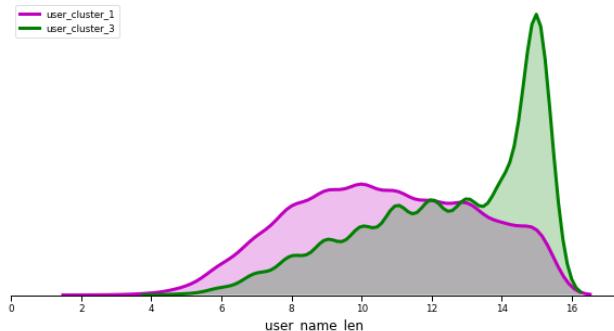


UMAP 2D Possible Clusters

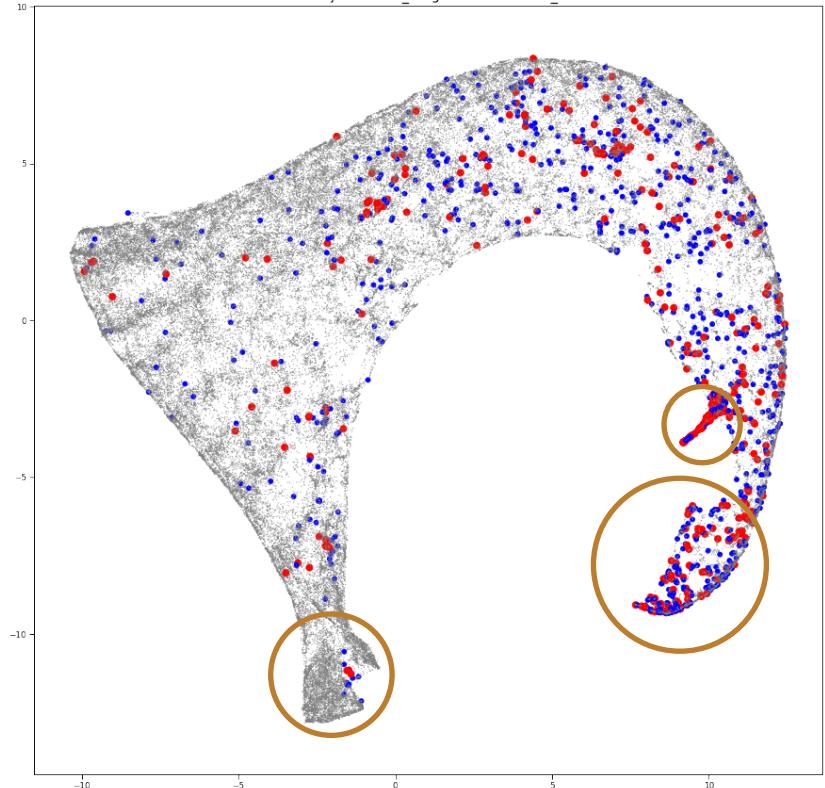
user_cluster_1-user_cluster_3 simil_name distribution



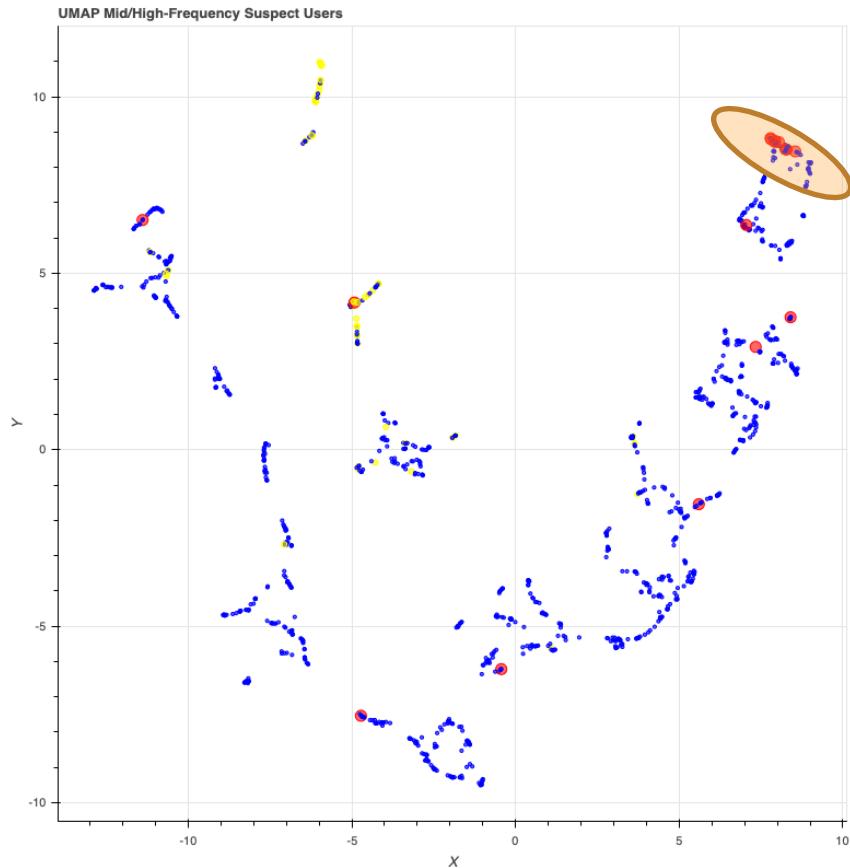
user_cluster_1-user_cluster_3 user_name_len distribution



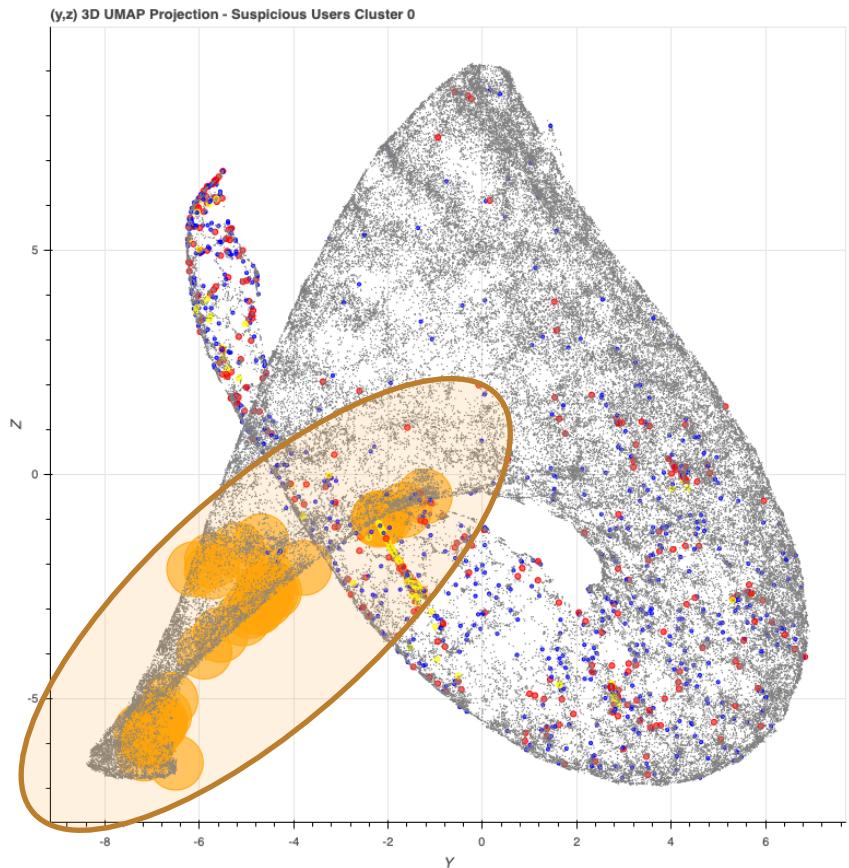
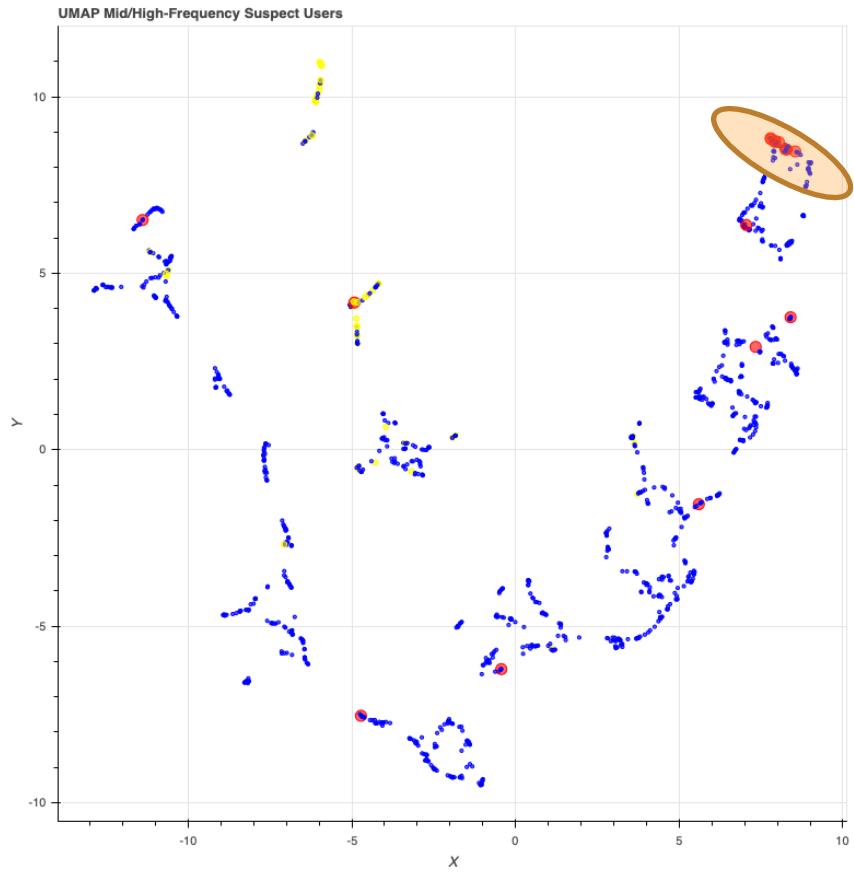
2D UMAP Projection - n_neighbors=30 - min_dist=0.0



UMAP Mid/High-Freq Users



UMAP – Cluster 0



Conclusion

- Studying the **frequency of tweets (timing)** together with **networking** seems the **right** path to follow
- Applying dimension reduction techniques as **UMAP** shows that **high-frequency accounts have several features in common**; and it is worth to be explored deeply to **speed-up bot detection**
- **Simple bots** are **scarcely found** on social media space
- Humans controlling suspicious accounts, alternating its activities between "human" and "bot" (**cyborgs**), **difficult their detection**
- **Classical methods** of bot's detection using Supervised ML models, **is not effective enough**

Future Works

- Capture **additional features**, focusing on **timing and networking**
- **Development** of a dashboard or an **App** similar to UDD-SpotBot to help **analyze Twitter user accounts**
- Test the **methodology used** in this study on **different datasets**
- Test **UMAP** with different **hyperparameters**

The screenshot shows a web application titled "UDD SpotBot User Metrics". At the top, there's a header with the URL "uddspotbot.herokuapp.com", a search bar, and some navigation links. Below the header, there's a user profile section for Alonso Astroza, featuring a small profile picture, the screen name "@aastroza", and other details like Name, Id, Location, and a "There is description" field set to True. To the right of the profile is a large table titled "Metrics" containing various statistics about the user's activity. The table includes rows for Number of Tweets analysed (199), Older Tweet Analysed (2018-08-31 21:54:00), Last Tweet Date (2019-07-31 16:27:12), Account Active Days since start (3710), Account Inactive Days (8), Average recent tweets per day (0.6), Average Tweets per day since start (3.08), Total Tweets since start (11438), Following (2221), Followers (997), Likes (5254), Retweet Ratio (54%), Mentions (298), Mentions Unique Index (0.59), Hashtags (40), and Hashtag Unique Index (0.8). At the bottom of the page, there are links for "Manage User on Twitter" and "Return to Main Page". The footer of the page also contains copyright information.

User	@aastroza
Screen Name	@aastroza
Name	Alonso Astroza
Id	44369850
Location	Santiago, Chile
There is description	True

Metrics	
Number of Tweets analysed	199
Older Tweet Analysed	2018-08-31 21:54:00
Last Tweet Date	2019-07-31 16:27:12
Account Active Days since start	3710
Account Inactive Days	8
Average recent tweets per day	0.6
Average Tweets per day since start	3.08
Total Tweets since start	11438
Following	2221
Followers	997
Likes	5254
Retweet Ratio	54%
Mentions	298
Mentions Unique Index	0.59
Hashtags	40
Hashtag Unique Index	0.8

Info retrieved at: Thu Aug 8 18:51:07 2019

Manage User on Twitter:

Return to Main Page

©2019 UDD - MDS18 - P0

Acknowledgments & Source Code

The analysis was performed in Python using Jupyter Notebooks, jointly with the Scikit-Learn, Pandas, Seaborn, Bokeh, Plot.ly, UMAP, Tweepy, and Botometer

It is essential to recognize the valuable advice, considerations, and share of experiences provided by Eduardo Graells-Garrido (UDD/BSC), Loreto Bravo (UDD), Leo Ferres (UDD), Maristela Rovai (SDSU), Ilza Rovai and Mauricio Pinto

Source code can be found at <https://zenodo.org/record/3576148#.XfUbsi2ZNTY>



Studying Twitter User Accounts: Spotting Suspicious Social Bot Behavior

Master candidate: Marcelo Rovai

Professors: Eduardo Graells-Garrido (Sup.)
Loreto Bravo
Leo Ferres

Thank you

Santiago, Chile - December 21, 2019