Social Network Analysis of Twitter Bots

Davide Cremonini, Artificial Intelligence, 0001137778
Alessia Crimaldi, Artificial Intelligence, 0001145505
Fabio Giordana, Artificial Intelligence, 0001145924
Gabriele Nanni, Artificial Intelligence, 0001146107

1 Introduction

In the current historical moment, the widespread diffusion of AI systems designed to simulate human behaviour is increasingly evident. Our project aims to evaluate the effectiveness of these systems in the context of bot detection on Twitter, seeking to identify behavioural patterns that differentiate automated users (aka *bots*) from human ones. To this end, we make use of Social Network Analysis techniques.

2 Problem and Motivation

Nowadays it is clear the impact of social medias in the everyday life of people. These applications are not only a way to interact with friends or observing celebrities but they are becoming one of the most impacting form of information for all generations, from middle-aged demographics to young adults, interacting also with children most of the population extracts their knowledge through social media. The creation of automated accounts on these platforms can in fact impact the perception of people spreading false information or polarizing content, causing the clash between two sides not willing to discuss and find a middle ground. Bots can also, posting and interacting with other accounts, show a different distribution of opinions with respect to the real one, causing public opinion to be shifted and inducing entities like companies or public organizations to move in certain directions because of the feedback given by them.
For this reasons we wanted to observe the behaviour of bots and human account to find a way to detect patterns that characterize one's relations. The focus of this project are the relations between the accounts, in the form of the "follow" action. This action is performed from one account to another signaling that the user is interested in receiving updates on the interactions of that account and it can be reciprocated.

3 Datasets

We conducted our analysis using the TwiBot-22 dataset [2], which is publicly available on GitHub[1].
TwiBot-22 is a comprehensive, graph-based benchmark for Twitter bot detection, featuring the largest dataset available up to date. It offers a diverse range of entities and relationships within

[1] https://github.com/LuoUndergradXJTU/TwiBot-22

the Twitter network, and boasts significantly improved annotation quality compared to previous datasets.
Given the extremely large size of the dataset it is impossible to work with the full dataset with our resources. For this reason preprocessing has been used to elaborate the dataset in a more efficient way.

3.1 Preprocessing

The two files we focus on are tweets and edges. The preprocessing is needed to extract the relation *follower-following* for every user to build the network and to obtain the number of tweets of each user.
We create chunks for edges and for tweets that is possible to store in the avaliable RAMs.
After this technical solution to be able to work with all data we proceed in creating the communities. The communities are sub-networks of people that used the same hashtag in their tweets. They will be used to check if the patterns found in the network are robust to network chances or are network specific. To test the robustness of the patterns we divided the subnetworks in different categories (large, medium and small communities) to test the results on different scales.

3.2 Adopted tools

The python libraries used in this project are:

- the *Polars* library and the *ijson* parser for the manipulation of the dataset.
- the *NetworkX* and *NetworKit* libraries to create the network and to calculate the measures on it.
- the *seaborn*, *SciPy* and *scikit-learn* libraries to search possible patterns in the bots/humans behaviour.

3.3 Analyzed networks

As discussed above, we divided subnetworks into three classes, based on the number of nodes (i.e. users) they contain: small networks have less than 1000 nodes, medium-size have between 1000 and 10000 and large have more than 10000. Among each class we chose to focus on communities which may present more polarising opinions as they discussed hot topics at the time of the dataset creation.

- Large $\rightarrow$ Ukraine, Ai, Covid
- Medium $\rightarrow$ Nato, Deeplearning, Nftcommunity
- Small $\rightarrow$ Ruleoflaw, Feminist, Agenda2030

4 Validity and Reliability

As discussed in its official paper [2], the TwiBot-22 dataset was created trying to address and mitigate known problems of previous datasets, such as poor annotation quality and low dataset

scale. This led to the construction of a large social graph with real world tweets, relationships between entities and metadata. This design allows the results to be statistically relevant and it accurately reflects Twitter's social dynamics. Moreover, the dataset benefits from a strong annotation pipeline, which guarantees reliability and consistency, and reproducibility for a wide range of bot detection and behaviour analysis studies. It is also easily accessible and freely available.
In our work, we exploited the dataset focusing only on *follower-following* relationships between users. Additionally, we experimented with applying the same analytical measures to the full graph and to subgraphs based on shared hashtags, in order to introduce a topical dimension to our analysis. This approach maintains the validity of the dataset, as both the follower links and the shared hashtag activity represent authentic user behaviours and capture meaningful patterns within Twitter's social structure. Our results are also fully reproducible and reliable, as we provide a detailed description of the preprocessing pipeline used to tailor the dataset to our specific research goals.

5 Measures and Results

The experiments were conducted on an NVIDIA GeForce RTX 3090 GPU (24GB VRAM).

5.1 Measuring the Network

5.2 Measuring the Nodes

Being tasked with finding notable features which may highlight the differences between human and bot users, we decided to compute well-established network measurements on the graphs. As the objective of this paper is to identify anomalies in the behaviour of specific users, we decided to focus our attention on node measures. Alongside these we added a few broader scope measures to show the presence of local groups around certain users.

Centrality.

- **Degree centrality.** Considering the entire population of users, the degree centrality is the total sum of the number of followers a user has and the number of accounts they follow. We might expect moderate values for humans, with followers growing organically, while it is probably more common for bots to have extreme values, either high, if artificially inflated, or low, for simple spam bots. To better investigate this hypothesis, we considered three different measures: *in_degree* (the number of followers), *out_degree* (number of followed), *degree_centrality* (sum of the two).

- **Reputation.** On top of the simple degree centrality, we decided to introduce the reputation measure as described in [1]. This acts as a ratio between the *in_degree* and the *degree_centrality* and highlights how unbalanced the distribution between followers and followed is for each user. While most humans should have a ratio close to 0.5, some notable ones may be closer to 1, if they have proportionally more followers. On the other side we expect bots programmed to boost following to have a reputation close to 0.

- **Reciprocity.** Leveraging the directed nature of the network, it is possible to compute how many of the follow relationships are reciprocated by each user. This may give us an

insight into the nature of users, as we can expect most humans to follow each other back, while bots prefer one-sided relations.

- **Betweenness centrality.** This measure allows us to understand how much a node acts as a crossroad between paths from other users. This may be an index of how much a user acts as a "common friend" between others. We expect higher values for some humans acting as bridges between communities, while bots are more likely to be peripheral.
- **Eigenvector centrality.** This centrality highlights the importance of a node depending on its neighbours. In our case, having a directed network, we followed NetworkX approach and computed the left-eigenvector, which adds the centrality of the predecessors. This means that a node will be given more relevance if it is followed by important nodes. We expect most bots to have a low eigenvector centrality, as it is unlikely for them to gather too much legitimacy from human users, while we expect a more balanced distribution for the latter, with some notables gaining more relevance.
- **PageRank.** Given that eigenvector centrality suffers because of zero-trailing, we noticed that these types of networks contained a high number of elements with 0 *in_degree*, so we decided to introduce the PageRank measure to better investigate the phenomena discussed before, accounting for the specifics of our problem.
- **Hubs & Authorities.** Another approach we attempted to exploit the directed nature of the graph is the deployment of the HITS algorithm. Supposing bots are less followed than humans, we expect them not to act as authorities and to be moderate hubs at best, while humans should have a more evenly distributed behaviour.

Clustering Coefficient. For each user u and the set of its neighbours N_u (users that follow it or are followed by it) this measure is the ratio between the number of couples of N_u that have a relationship between each other and their total number. This gives us insight on how a user acts as a centre of its local community. We expect human to be part of more meaningful communities, thus resulting in higher values.

Average Neighbourhood Degree.

Core Number.

Number of Triangles.

Number of Tweets.

6 Conclusion

7 Critique

The *follower* and *following* connections may not represent meaningful relations for bot detection, as the TwiBot-22 dataset lacks a clear separation between human and bot accounts in this aspect.

References

[1] David M. Beskow and Kathleen M. Carley. You are known by your friends: Leveraging network metrics for bot detection in twitter. In Babak Akhgar, Hamid R. Arabnia, and Petra Saskia Bayerl, editors, *Open Source Intelligence and Cyber Crime*, pages 53–88. Springer, 2020.

[2] Shangbin Feng, Zhaoxuan Tan, Herun Wan, Ningnan Wang, Zilong Chen, Binchi Zhang, Qinghua Zheng, Wenqian Zhang, Zhenyu Lei, Shujie Yang, Xinshun Feng, Qingyue Zhang, Hongrui Wang, Yuhan Liu, Yuyang Bai, Heng Wang, Zijian Cai, Yanbo Wang, Lijing Zheng, Zihan Ma, Jundong Li, and Minnan Luo. Twibot-22: Towards graph-based twitter bot detection, 2022.