Brazilian Capitol: A Network of Criminals

Emerson Rodrigues Vero filho   Ali Ahmaadi

February 21, 2023

Abstract

A social network of far-right criminals directly involved with the episode of Brazilian congress building breakthrough was studied based on their twitter connections, the nodes being one user and the edges being the connections between them. Different centrality measures were applied to discover the most influential users of this network.

1 Introduction

Thousands of far right organized rioters broke into Brazilian Government building on January 8, in a similar fashion to the Capitol episode in the United States two years ago. These people claim that the elections won by Lula (current president) against Jair Bolsonaro (former president) were somehow a fraud. They explicitly ask the army to implement a military dictatorship, invoking the darkest period of Brazilian recent story, that begun in 1964's coup d'etat and lasted 24 years, marked by the imprisonment, torture and death of thousands of activists, repression to any kind of worker's or student's political organization or even to any kind of art manifestation that could slightly sound as a challenge to "good manners" and the "christian Brazilian traditional family." [1]

Since 1989, Brazilian government regime is a regular presidential parliamentarism that theoretically guarantees the right to protest but also formally criminalizes open attacks to democracy and to the regime itself [2]. It is also a crime to attack State public property.[2] This is enough to state that the people who participated in the attack in Brasilia are indeed criminals that deserve some kind of punishment, even though the State is quite lenient towards far-right criminal attacks in general.

So little was the fear of punishment, that during the episode, the invaders made tweets providing the proofs of their own participation. Not only that, over the whole country in the biggest cities, during two months, people set camp in front of Brazilian Army facilities in preparation for the tentative of coup d'etat about to come. This people ultimately form together a social network that is the study object of this work.

One is able to query tweets based on the geographic coordinates of the building and the exact time that the invasion happened, and also query based on the coordinates of the camps, and with a span of time corresponding to their duration.

The idea is to analyse the network of people who actively participated on the event, the nodes being a twitter user, and the links being their followers. This way, centrality measures can be applied to reach a conclusion on who are the most influential users, who are the ones responsible for spreading information the most, etc.

2 Dataset Preparation

Queries were made directly from twitter website engine, specifying the geographic coordinates of Brazilian government building and the data of the episode. The following line contains the text of the first query:

geocode:15.7990, 47.8608, 0.02km, since:2023-01-08 until:2023-01-09

It was also queried a period of two months (the queried period comprises the day after the election of Lula, 31/10/2022, until the date of the attack ,08/01/2023) for the tweets coming from the geographic location of the far-right organized camps in the following cities:

Figure 1: Crowd of Far-Right Rioters outside the building, the poster says "Military Intervention, now!"

Figure 2: Rioters inside the government building, clashing against police forces.

- Brasilia
- Rio de Janeiro
- Recife
- Curitiba
- Sao Paulo
- Porto Alegre

Only the users that were clearly identified as far right were chosen to be part of the Dataset. This was an easy task because their aesthetics is extremely flashy.

Figure 3: Flashy aesthetics of a regular Brazilian right wing twitter account.

Gathered a list of twitter users, vicinitas website [4] was used to get csv files containing the list of followers of each one of the users. This is a free online tool used to scrape twitter's users data. It can only download a maximum of 6500 followers for users, so it limits the research to people who have less than 6500 followers, which is also interesting, because doing so the focus will be on "ordinary" people (who are not web influencers somehow) who had a significant role in the event. Of course greater influencers were very relevant, but because they are already widely known, concentrate on "ordinary" people may be more informative. A total of 247 users were identified, and had their data scraped by vicinitas. The dataset created was made available on Kaggle website.[5]

From a folder full of xlsx files, it was created a Python dictionary with its keys being the users and the value being the list of their followers. 134 of these users were found to have no connection with the others, so they were not relevant to further kinds of analysis of the network, other than a simple degree analysis.

3 Data Analysis

Python package networkx (3.0) [6] was used to deploy and study the network. The full network has 247 nodes, but only 113 of them were displayed, given that the others have degree equal to zero. This is a directed network, unweighted, with nodes labeled as the account of each user. Kamada Kawai style of visualization, associated with Kamada Kawai method for determining the nodes position [6] was chosen empirically as the available one that provides the clearest and most useful visualization.

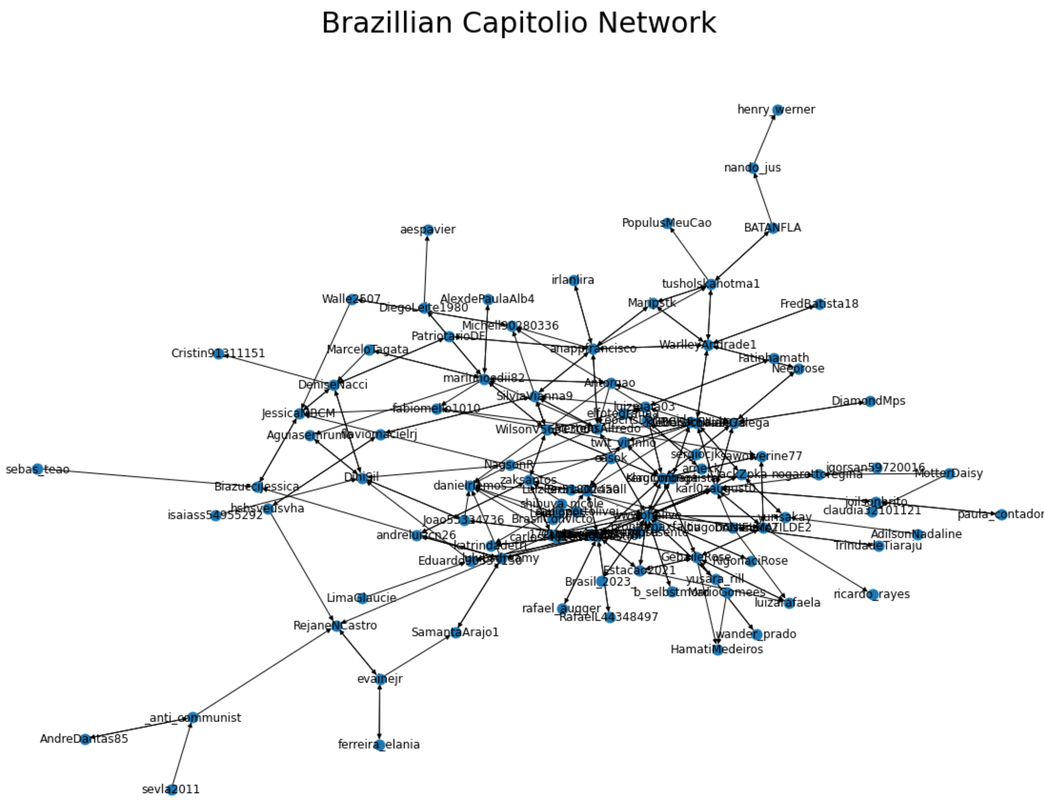

Figure 4: Original Network.

Another way to represent the network is by visualizing its adjacency matrix. The adjacency matrix is a matrix that describes which nodes are connected to each other by the edges. . The rows and columns of the adjacency matrix correspond to the nodes in the graph. If node i is connected to node j, then the element (i, j) in the adjacency matrix is 1. If there is no edge connecting nodes i and j, then the element (i, j) in the adjacency matrix is 0. Due to the significant amount of dots, a visualization of the adjacency matrix is performed as a square table with blue dots where a connection is present and a grey space, otherwise.

It is also possible to model the weights for the edges as being an inverse tangent of the log of the ratio between the amount of followers of the two nodes:

$$weigths(i,j) = \arctan(k \log \frac{followers_i}{followers_j}) \tag{1}$$

This formula makes sense if one thinks that a strong twitter profile will have higher influence on a small one (up to a certain saturation) and a small twitter profile will have low influence on strong one. Also, accounts with similar amount of followers will influence each other similarly. The following image shows the computation of the weight function for all the edges of the dataset.

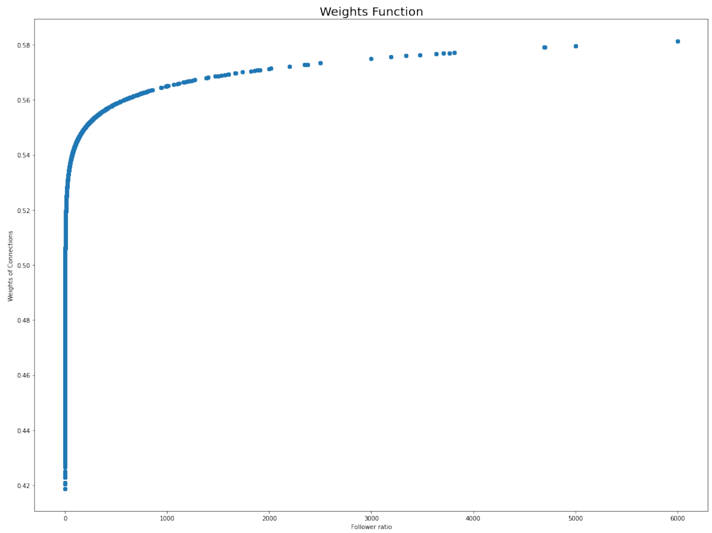

Figure 5: scattered plot of Weights function computed for followers ratio existing values. k=0.03

With the weights of the networks calculated, a new adjacency matrix was created using the standard damping factor . A color bar between blue and red denoting respectively the weakest and the strongest connection was used to better visualize it.

Figure 6: Adjacency Matrix Visualization.

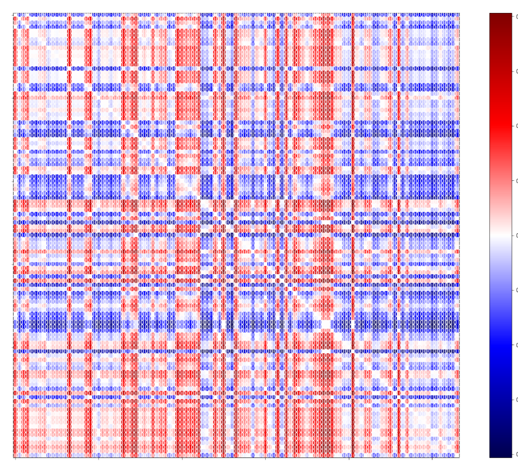

Figure 7: Weighted Adjacency Matrix

The degree of a node in a network is the amount of direct connections to other nodes. A degree distribution can be built showing the amount of nodes for the possible degrees. Histogram like degree distribution were deployed both considering and neglecting the zero degree nodes. The results both for the average computed considering zero nodes and neglecting them are deployed in the Table below.

3.1 Centrality measures

Centrality measures are used to identify the most important or influential nodes in a social network. The following centrality measures were analyzed:

	Average degree
Including zero-degree nodes	2.05
Excluding zero-degree nodes	4.62

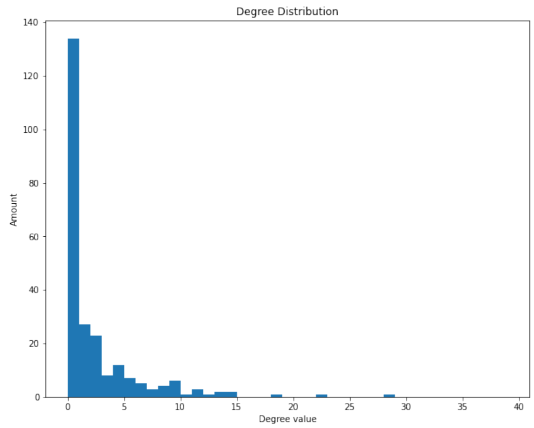

Figure 8: Degree distribution containing zero degree nodes.

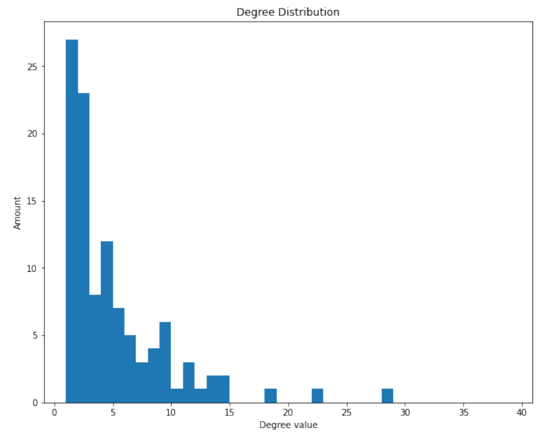

Figure 9: Degree distribution excluding zero degree nodes.

- Degree centrality
- Betweenness centrality
- Closeness centrality
- PageRank centrality:

3.1.1 Degree Centrality

Degree centrality measures the number of connections that a node has in the network. Nodes with high degree centrality are well-connected and often have a lot of influence in the network.

To highlight the relative centrality between nodes, the results of the degree centrality networkx algorithm for each node were given as a building parameter of a graph, as the color intensity and size of the nodes. So, the higher the centrality, the bigger and darker the nodes deployed are. Also, the five most central ones are labeled, and the values associated to them are shown in the table bellow.

Id	**Degree**	**Score**
57	wwwbhalive	0.264151
56	predsoncastro	0.207547
53	anappfrancisco	0.169811
24	KlauCoronga	0.132075
4	karl0zaugust0	0.132075

3.1.2 Betweenness Centrality

Betweenness centrality measures the extent to which a node lies on the shortest paths between other nodes in the network. Nodes with high betweenness centrality act as bridges between different parts of the network and can have significant influence on the flow of information or resources in the network.

The same procedure as before was repeated to indicate betweenness centrality graphically, so the darker and biggest the nodes, the more central they are.

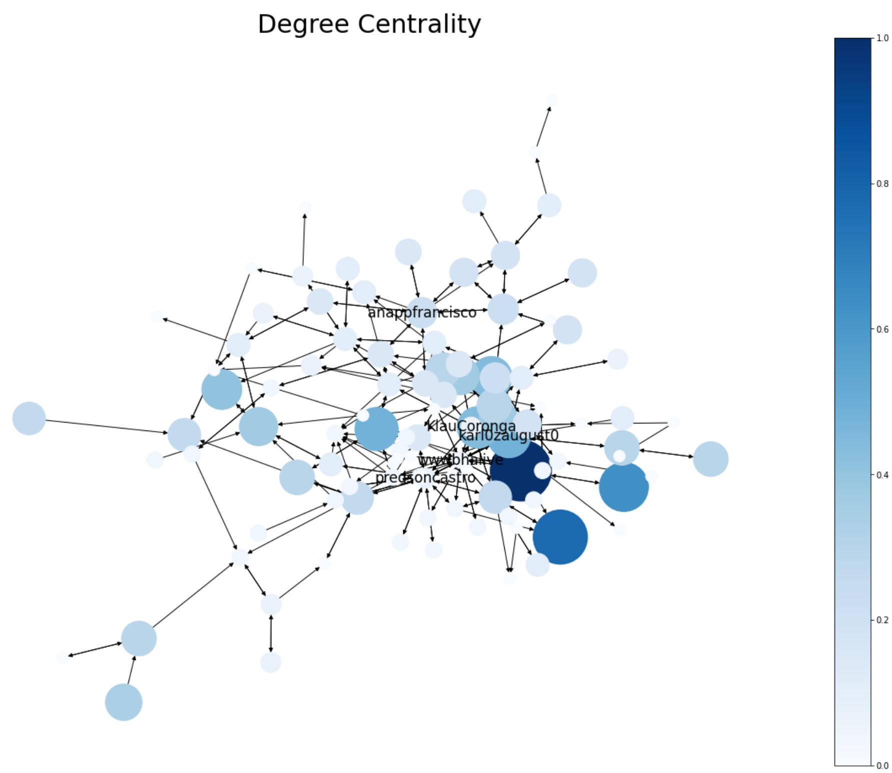

Figure 10: Degree Centrality visualization.

Id	Names	Betweenness
57	wwwbhalive	0.151239
56	predsoncastro	0.124730
24	anappfrancisco	0.089177
53	KlauCoronga	0.087842
35	MertensAlfredo	0.085793

3.1.3 Closeness centrality

Closeness centrality: Closeness centrality measures how close a node is to all other nodes in the network. Nodes with high closeness centrality can reach other nodes quickly and efficiently, making them important for the flow of information or resources in the network.

Again, the darker and bigger the nodes, the more central they are.

Id	Names	Closeness
56	predsoncastro	0.256578
57	wwwbhalive	0.251955
53	KlauCoronga	0.242138
4	karl0zaugust0	0.238017
35	MertensAlfredo	0.227374

3.1.4 PageRank centrality

PageRank is a measure of centrality originally used to rank web pages in search engines. In a social network, PageRank centrality measures a node's importance based on the importance of the nodes that are connected to it. So, the idea was to feed PageRank algorithm with the weighted adjacency

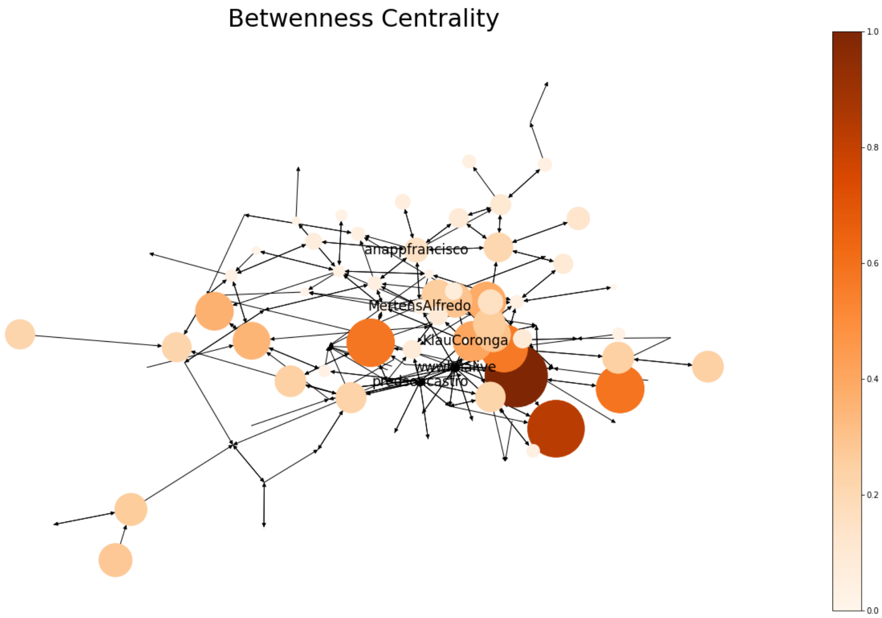

Figure 11: betwenness Centrality visualization.

matrix, that carries the information of the relative importance of connections. The top PageRanks scores are indicated in the table below.

Name	**Pagerank**
flaviomacielrj	0.009442
wwwbhalive	0.009427
AndreDantas85	0.009422
LimaGlaucie	0.009419
marinhoedii82	0.009411

4 Conclusion

In total, 10 different twitter profiles were accused by some of the top centrality measures tables.

These are the profiles retained by this analysis as the most influential ones to the episode of the breakthrough of Brazilian government building on the 8th of january of 2023.

- wwwbhalive
- predsoncastro
- KlauCoronga
- marinhoedii82
- karlz0august0
- LimaGlaucie
- flaviomacielrj
- MertensAlfredo
- anappfrancisco

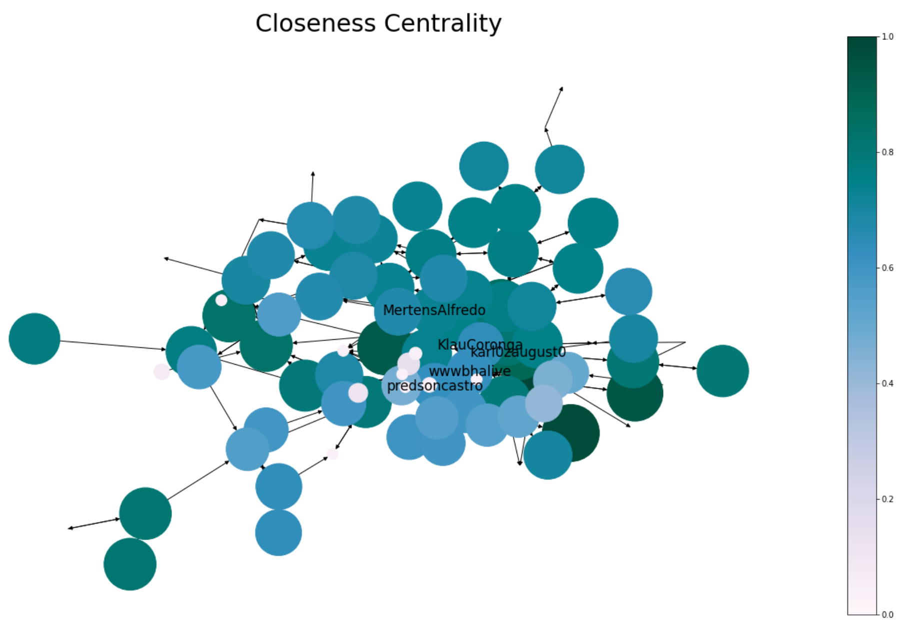

Figure 12: Closeness centrality visualization.

- AndreDantas85

  Particularly, the user ’wwwbhalive’ appeared in every single centrality measure realized, so this user is the great ”champion” of the analysis, the best guess for the most influential node of the network of far-rightcriminals

Figure 13: The greatest influencer of the network of far-right criminals

References

[1] CNN. (2023, January 9). Brazil’s Congress attacked by far-right rioters. CNN. https://edition.cnn.com/2023/01/09/americas/brazil-congress-attack-explained-intl/index.html

[2] Brazil. (1988). *Constituição da República Federativa do Brasil*. [Translation: Constitution of the Federative Republic of Brazil].

[3] Vicinitas. (n.d.). Vicinitas: Data-Driven Intelligence for Legal Professionals. https://www.vicinitas.io/.

[4] Camarada Charlie Brown. (2021). Bolsominions. Kaggle. https://www.kaggle.com/datasets/camaradacharliebrown/bolsominions.

[5] Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. Information Processing Letters, 31(1), 7-15.