International Journal of Communication 16(2022), 2431–2455 1932–8036/20220005

WhatsApp and Digital Astroturfing: A Social Network Analysis of Brazilian Political Discussion Groups of Bolsonaro's Supporters

VIKTOR CHAGAS[1]
Fluminense Federal University, Brazil

Digital astroturfing and computational propaganda have drawn a lot of attention in recent years because of the malicious effects on the political environment, especially in the face of the emerging far right. But most studies on astroturfing are limited to seeking theoretical concepts. The present article suggests that the concept of astroturfing can be conceptually defined and empirically investigated through a social network analysis (SNA) approach. The article is specially focused on understanding the use of mobile instant messaging services (MIMS) like WhatsApp as a stage for astroturfing practices in 2018 Brazilian elections. Its main hypothesis is that SNA methods can help in understanding how a Bolsonarist influence operation and misinformation network was structured. Results show different thematic groups and several functional clusters and lead to the identification of practices that match with those of agents from the professional field of politics.

Keywords: digital astroturfing, social network analysis, misinformation networks, Jair Bolsonaro, WhatsApp

While attention to the effects of astroturfing and computational propaganda on the political environment has grown substantially in recent years (Benkler, Faris, & Roberts, 2018; Chakrabortty, 2020; Henrie & Gilde, 2019; Woolley & Howard, 2018), few empirical works offer more than merely a conceptual definition for this phenomenon. Most of the literature on astroturfing still shows much interest in offering theoretical outlines and little in allowing researchers to identify this kind of action in progress (Kovic, Rauchfleisch, Sele, & Caspar, 2018; Ratkiewicz et al., 2011). This study aims to shed some light on how

Viktor Chagas: viktor@midia.uff.br
Date submitted: 2021-02-01

[1] The author would like to acknowledge the anonymous referees for the insightful comments during the review process. He also gratefully thank Wilson Gomes, Samuel Barros, Júnia Ortiz, Crysttian Paixão, João Guilherme Santos, and Miguel Freitas. This article benefits from the support of the National Council for Scientific and Technological Development–CNPq (Fellowships No. 306791/2021-8 and 243218/2018-2) and the Carlos Chagas Filho Foundation for Research Support of the State of Rio de Janeiro (Fellowships No. 259788 and 249104). This study was approved by the Fluminense Federal University ethics committee, permission No. 29720620.8.0000.5243. Requirement for consent from all subjects was waived by the ethics committee.

social network analysis (SNA), when combined with content analysis and other descriptive statistics, offers interesting tools for identifying astroturfing in the political environment.

This article draws from a case study on the 2018 elections in Brazil, particularly the widespread use of mobile instant messaging services (MIMS) in the digital campaign of then candidate Jair Bolsonaro. Among these services, WhatsApp is one of the most widespread in the world and has a large user base in Brazil.

Although the mobile app is generally highlighted by its end-to-end encryption system, this same system can contribute to the viral spread of messages in discussion groups, including the wide dissemination of misinformation, as occurred in Bolsonarist political groups in 2018 (Chagas, Mitozo, Barros, Santos, & Azevedo, 2021; Chagas, Modesto, & Magalhães, 2019; Kischinhevsky et al., 2020; Piaia & Alves, 2020). In addition, because WhatsApp does not have a public API and does not offer public metadata about the content circulated on the platform (such as the number of views or shares of a given message, groups with which a given user is associated, etc.), its opacity presents a challenge for researchers interested in understanding the degree of coordination performed by some groups of users on the platform.

Therefore, this study draws attention to how SNA, and, thereby, a perspective that allows for assessment of how actors relate to each other within a network, can provide researchers with useful information about how some digital astroturfing actions are developed. Even amid the scarcity of metadata from a difficult-to-monitor platform like WhatsApp, the analysis of relational data allows for understanding (1) how networks evolve over time, and therefore whether or not they are growing organically, (2) how actors and clusters perform or distribute among themselves the tasks that they are responsible for, (3) the sociodemographic profiles of users, and (4) the engagement and centrality of some of these actors within the network, which is essential to understanding whether or not they are part of an alternative network of influence (Lewis, 2018).

To a large extent, the Cambridge Analytica scandal helped to draw attention to the use of computational propaganda and micro-segmented data in political campaigns (Benkler et al., 2018). However, most studies that address this topic have invested in methods for identifying bots in social media (Arnaudo, 2017), rather than looking at the phenomenon more broadly and focusing on the participation of actors from the professional field of politics who emulate spontaneous and disinterested behaviors.

In comparison with those on other social media platforms, campaigns based on instant messaging services like WhatsApp are fundamentally different on at least three distinct levels. These services not only provide little metadata about the content that circulates through it, but they also do not rely mainly on algorithms, as other social media platforms do (Santos, 2019). The result is that bots do not have the same prominence as on platforms such as Twitter or Facebook, and most of the automatized behavior is performed by click farms and mass message triggers, which are mainly human operated (Friedberg & Donovan, 2019). Add to this the fact that the MIMS are generally developed in an end-to-end encryption model, which allows for the free flow of information with greater security, even within authoritarian contexts and strong media regulation, but ends up becoming an inhospitable environment in which opacity makes it difficult to perform

even simple actions, such as tracking misinformation (Resende et al., 2019). Therefore, MIMS are fertile ground for the development of disingenuous activities.

A second relevant aspect is that these services are based on a model of private communication, meaning that voters feel that they are in a close communication environment, surrounded by people they trust; this facilitates the formation of ideological bubbles (Rossini, Stromer-Galley, Baptista, & Veiga de Oliveira, 2021). WhatsApp routinely presents itself as a peer-to-peer communication app. According to data from the WhatsApp office in Brazil, more than 90% of messages exchanged through the application circulate in a flow of communication between two people, and the groups have an average of up to seven people. However, political use of the tool is based mainly on so-called discussion groups. These groups have a maximum capacity of 256 people and are often composed of family members, friends, coworkers, neighbors, or people who have close interests of various kinds.

Finally, these services are mainly developed for mobile devices. This implies not only a certain pervasiveness, especially in cultures with strong online sociability, such as that in Brazil, but also an integrated environment of communication, which could easily be recognized as a hybrid media system (Chadwick, 2013). Through WhatsApp groups, it is common to receive links to other platforms, such as YouTube, screenshots of tweets, and memes. The tool is a kind of hub for other services and content, including calls to action and antiscience posts (Chagas et al., 2021; Massuchin, Tavares, Mitozo, & Chagas, 2021).

Thus, the present study aims to advance the use of digital methods, especially from an SNA background, that allow one to identify digital astroturfing dynamics within WhatsApp's environment. This study is based on the following hypotheses:

H1: The social network analysis combined with the enunciative content analysis of Bolsonarist WhatsApp groups can elucidate the nuances that make up the conservative electorate in Brazil.

H2: Regular monitoring of a large number of public discussion groups on WhatsApp allows the identification of routines and behaviors typical of professional actors from the political field, which can ultimately be characterized as digital astroturfing.

To test these hypotheses, 124 WhatsApp public groups were monitored between May and October 2018, covering the pre-campaign, campaign, and after-campaign periods in Brazil. Around 12,000 users were members of the observed groups, and approximately 760,000 messages were exchanged during these five months. Results allow a deeper understanding of the topological structure of the Bolsonarist network. Although it is possible to understand this network as a distributed and organic one, this study also shows patterns of activity among members and group administrators that resemble hierarchically organized networks performing coordinated actions with a high degree of centrality. Therefore, this article presents at least two major contributions to the debate on astroturfing, political mobilization, and digital campaigns. The first is a theoretical-methodological gain; the article argues that it is possible to recognize coordinated actions based on an SNA approach. Thus, unlike most studies that are limited to conceptualizing what astroturfing is, here, it is possible to empirically operationalize the concept. The second contribution is specifically related to the Brazilian case;

this study shows, as other authors have suggested (Kischinhevsky et al., 2020), that WhatsApp played a central role in the 2018 elections.

To account for these observations, the article is divided into four sections. The first discusses the definitions of digital astroturfing and its political effects. Then, briefly relying on contextualization of the Brazilian case in the 2018 elections, the second section frames how WhatsApp's opacity can foster a worrying context for astroturfing to boost far-right extremist messages and misinformation. Finally, the last two sections present and discuss the empirical use of SNA methods to identify astroturfing practices.

Astroturfing and Digital Campaigns

Over the past 10 years, astroturfing has gained the spotlight of academic research in digital democracy. Some scholars (Henrie & Gilde, 2019; Lee, 2010) suggest that the intensive use of social media by Obama in 2008 and, of course, the use of micro-segmented data by Trump in 2016 may have contributed to launching the topic under scrutiny. Most studies on astroturfing are dedicated to defining it theoretically, but they lack ways to empirically recognize it. How does one know when actions based on astroturfing are taking place?

Astroturfing is defined by Silva (2015) as a "staged demonstration." It is a practice employed by political organizations in which professional lobbyists emulate an organic public (Howard, 2006). This emulation is performed by what Leiser (2016) classifies as a "repetition heuristic"—that is, the reiteration of false claims over a certain issue—to impact public opinion and attract other sympathizers.

The first conceptual definitions of astroturfing do not differentiate regular astroturfing from digital astroturfing. An important contribution was developed by Zhang, Carpenter, and Ko (2013), for whom digital astroturfing, or online astroturfing as they call it, can be defined as "the dissemination of deceptive opinions by imposters posing as autonomous individuals on the Internet with the intent of promoting a specific agenda" (p. 3). The researchers infer that the basic characteristics of online astroturfing include the low cost of sending mass messages online and the great availability of computational resources and strategic protocols developed by campaign managers.

Leiser (2016) is another author to attempt a definition. He describes cyberturfing—that is, cyberspace astroturfing—as "the manipulation of a digitally mediated platform in order to advance a specific agenda, product, or political ideology" (p. 5). According to Henrie and Gilde (2019), opinions expressed in comment sections, electronic forums, online public consultations, and even rating-and-ranking reviews of products and services are favorable environments for astroturfing. They argue that an association exists between astroturfing and trust. This view is consistent with that of Kovic et al. (2018), for whom digital astroturfing is primarily political and polarized.

But, unlike what happens with regular astroturfing, digital astroturfing is always performed as a clandestine activity (Kovic et al., 2018), is difficult to detect and constrain (Leiser, 2016), and has much greater potential for inflicting more severe damage and consequences on the information

environment. Thus, authors such as Ratkiewicz et al. (2011) identify it as a form of abuse or online misinformation.

The most likely origin for its clandestinity lies in several unique features of computer-mediated interaction. As Stromer-Galley (2000) argues, politicians do not have many incentives to interact with their constituencies online. Political actors tend to avoid interaction and usually prefer pseudo-interactive environments where they can control third-party responses, whether due to financial costs, lack of human resources, or simply not wanting to lose their strategic ambiguity. However, as online campaign models evolve, there is a greater sense of direct interaction between politicians and their constituents. Gomes, Fernandes, Reis, and Silva (2009) call attention to how the Obama campaign used an expedient in which the candidate was present on different social network sites, but in reality, there was little direct contact with the voter.

Astroturfing is often used as a way of building a climate of opinion or setting the agenda for media coverage. Two of the main characteristics of digital astroturfing are the viral spread of their messages and the opacity of the campaign sponsors. Given these aspects, one can understand astroturfing campaigns as coordinated actions, not collective or connective ones—and this leads to the use of SNA. These actions are often based on a small group of digital influencers or performed on a small number of super posters (Graham & Wright, 2013), with a strong investment in a broadcast model emulating direct interaction.

Digital astroturfing takes advantage of these characteristics of computer-mediated interaction to generate a bandwagon effect. Without properly interacting with audiences, online campaigns typically emulate public adhesion, using combined networks of social media bots, professional spammers, and other relatively low-cost techniques. Studies on computational propaganda, including those presented by Woolley and Howard (2018), draw attention to the use of these techniques in countries such as Brazil, Ukraine, Taiwan, and many others. In some cases, politicians hire bot networks to fight for Twitter trending topics with favorable hashtags.

But the concept of computational propaganda does not emphasize the clandestinity as much as astroturfing does. In an empirical sense, this results in the fact that several computational propaganda studies have invested in target advertising or botnet recognition techniques (Woolley & Howard, 2018), although astroturfing actions are not limited to this. Often, astroturfing uses real accounts and relies on influence operations networks to leverage their agendas (Friedberg & Donovan, 2019; Lewis, 2018).

For this reason, the way in which these campaigns build their coordinated actions must be studied in detail. A relational approach can highlight the nonorganic growth of the network, the functional behavior of certain user clusters, and the levels of engagement and individual participation of specific users in relation to others.

Astroturfing and Brazilian WhatsApp Far-Right Groups

Although critics of digital activism usually describe it as ineffective or uncertain because of the common understanding that protests in virtual space generally lead to no real effects, as in the so-called slacktivism literature (Klang & Madison, 2016; Vie, 2014), digital astroturfing cannot be neglected. As Christensen (2011) argues, when it comes to their efficaciousness, online political actions are generally underestimated. And as Chagas (2019) demonstrated with the Vomitaços protests, a cycle of humorous online protests in which users posted a "puke" sticker on former president Michel Temer's official Facebook fan page, sometimes it seems much more difficult to censor and disperse online political actions than to break up street demonstrations.

As Benkler et al. (2018) note, the crisis with far-right and negationist groups emerging recently "is more institutional than technological" (p. 20). This is why the same platforms that enable online progressive activities can be easily converted into arenas of extremism. Bolsonarist networks on WhatsApp are an example of how a private communication service with end-to-end encryption can ultimately favor the creation of largely ideological bubbles, in which it is possible to identify the circulation of misinformation, campaigns, and the most diverse attacks on democratic institutions.

WhatsApp was created in 2009 by Brian Acton and Jan Koum. The application was originally an instant messenger and later incorporated voice over IP and end-to-end encryption functions, making it a popular communication platform on mobile devices. In 2014, WhatsApp was acquired by Facebook. Since then, the application has become one of the most popular in Latin America.

According to data from Latinobarómetro, use of WhatsApp averages 64% on the continent; Brazil is among the countries with the largest population of users, with 66% of the population connected. According to data from the company itself, more than 120 million WhatsApp accounts were active in Brazil in 2017, representing around 8% of users worldwide.

Recent changes in Internet users' preferences can be seen by comparing the data from the Brazilian Internet Steering Committee (CGI.br) regarding online users' activities. The survey compares the use of social network sites, instant messaging, e-mail, voice conversation apps, online blogs, and online forums or mailing lists. A significant change in users' preferences since 2014 is apparent, with the number of users sending instant messages via private chats notably increasing (see Table 1).

Table 1. Growth in the Use of Private Communication Services in Brazil (Percentages).

Year	Use social network sites	Send instant messages	Send or receive e-mails	Chat via voice apps	Use blogs or news sites	Use forums or mailing lists
2018	75	92	57	70	9	7
2017	77	90	58	67	9	10
2016	78	89	60	60	9	12
2015	77	85	60	54	9	11
2014	76	83	64	26	13	11
2013	77	74	72	32	18	17
2012	73	59	70	23	15	12
2011	70	70	80	25	23	15
2010	69	74	79	17	14	10
2009	67	70	79	17	*	12
2008	70	61	77	17	*	15
2007	64	55	78	17	*	11

* Not measured this year.
Source: The author, based on data from Brazilian Internet Steering Committee (CGI.br).

News consumption on WhatsApp is also an increasingly relevant activity. According to the Digital News Report 2019, 53% of survey respondents in Brazil say they use WhatsApp as a news source, the highest index among the 38 Latin American countries evaluated (Reuters Institute, 2019). In October 2018, a Datafolha survey noted out that 62% of respondents believe the news they receive through WhatsApp, compared with only 8% who do not (Passos, 2018). Another survey from the same organization found that six of 10 Bolsonaro voters were informed mainly by WhatsApp and use the platform to share political news ("Datafolha," 2018).

It is noteworthy that then deputy Bolsonaro presented a bill in July 2017 in which he intended to inhibit suspensions or interruptions of services such as WhatsApp. In March 2017, Bolsonaro would meet Steve Bannon for dinner in Washington. In August 2018, weeks before the elections, in a photo widely circulated on social media, Eduardo Bolsonaro, son of Jair Bolsonaro, appears beside Bannon in a private meeting. The evidence points to long-term campaign planning that started at least two years before the elections.

On October 18, 2018, the eve of the second round of elections, journalist Patrícia Campos Mello published an investigative story attesting that businessmen had purchased mass mailing services via WhatsApp to disseminate negative information about Bolsonaro's opponents (Mello, 2018). The practice, illegal under Brazilian legislation, led WhatsApp itself to take a series of measures to contain the spread of fake news. In Brazil, the message forwarding limit of 250 contacts was reduced to 20. The measure is similar to that adopted in India after a series of lynchings (Arun, 2019). WhatsApp also banned around 100,000 users (Hous, 2018). Bolsonaro expressed his disagreement with WhatsApp measures publicly. However, his online campaign does not appear to have suffered a major setback, and Bolsonarist WhatsApp groups

continue to engage in political proselytism and spread misinformation and negative campaign messages (Chagas et al., 2019; Massuchin et al., 2021).[2]

For many people, the Bolsonarist network on WhatsApp was an organic phenomenon in which supporters spontaneously discussed politics. However, a closer look at this network reveals a structure composed of a set of administrators of multiple groups and users who can send thousands of messages. These factors, along with the opaque environment of the platform itself, make the use of WhatsApp during the 2018 elections in Brazil an important case study for understanding the effects of digital astroturfing in contexts of political polarization.

Methodology

The body of studies on WhatsApp is growing (Caetano, Oliveira, Lima, Marques-Neto, & Magno, 2018; Moura & Michelson, 2017; Zúñiga, Arvèvol-Abreu, & Casero-Ripollés, 2019), and some seek to emphasize its impact on the electoral scenario in Brazil (Chagas et al., 2019; Kischinevsky et al., 2020; Piaia & Alves, 2020) and other parts of the world (Arun, 2019; Pang & Woo, 2020; Treré, 2020). Although it is debatable whether to WhatsApp can be described as social media, the platform is undoubtedly rich in elements that characterize the organization of online social networks. However, because of its architecture, conducting scientific research on WhatsApp has presented ethical and methodological challenges for researchers (Barbosa & Milan, 2019).

This study started from a systematic monitoring of public WhatsApp political discussion groups. WhatsApp has three different types of groups: broadcast lists, secret groups, and public groups. Public groups are those whose invitation links circulate publicly for users to access. These groups usually house users from different backgrounds but driven by the same interests. Unlike with broadcast lists, users in public groups can freely exchange messages, so the environment is characterized by a principle of horizontality.

To monitor these groups, this study started months before the elections, in the first semester of 2018. Two public groups were initially selected, and from those groups, it was possible to access other several groups supporting Jair Bolsonaro's candidacy, always through public invitation links. New groups were added using a snowball technique. The basic criterion for selecting new groups was that they had a title or description explicitly indicating support for Bolsonaro's candidacy or his party. Around the beginning of August, the sample reached about 120 groups, and public invitations to access new groups started to repeat at most, which suggested a saturation point.

The study took place using a covert research protocol, one in which the members of the observed groups were not previously informed of the ongoing investigation because it is an environment hostile to academic research. The covert research is defined by the Brazilian National Health Council, which regulates and supervises ethical research procedures, such as "research conducted without the participants being

[2] Although WhatsApp groups for other parties also exist, they are less articulated and less numerous than those from the far right, and they have not contributed to the massive circulation of misinformation (Santos, 2019).

informed about the objectives and procedures of the study, and without their consent being obtained previously or during the research" (National Health Council, 2016, p. 3). It is a common strategy, for example, giving a placebo to a control group without the participants being informed about what substance is being administered to them. According to a 2016 resolution of this committee, this model is justified "in circumstances in which information about the research goals and procedures would change the target behavior of the study or when the use of this method presents itself as the only way of conducting the study" (National Health Council, 2016, p. 3).

Although there are a number of critics of this method, it has proved to be the only one that allows for providing data for an investigation based on an environment formed mainly by groups that are averse to scientific research and mostly antiscience (Massuchin et al., 2021). Because users, and the candidate himself, have repeatedly criticized universities and scholarly research, once presented as a researcher, the author would have been banned immediately from the field of observation.

Thus, groups were continually monitored for months. During this time, ethnographic observations were also gathered. It should be noted, however, that all data used in this and other analyzes respect the private nature of the metadata available in this environment. SNA uses only data on an aggregate scale; no personal information not made public has been used as a result of this research.

The database for this study comprises 124 public groups. The relational data refer to the connection between users and monitored groups, between administrators and groups, and between messages sent and users. Most studies on WhatsApp have mostly focused on the messages sent, given that chats exported from the native WhatsApp tool return nonparsed text files with backups of messages to each group (Piaia & Alves, 2020). Because of the scarcity of metadata provided by the platform, few studies deal with users' belonging to groups. These issues make the research based on SNA a difficult enterprise.

Data such as which users and administrators belonged to which groups were collected semiautomatically based on a data-scraping procedure from the WhatsApp Web service. Additionally, these ties are very volatile and impermanent because users enter and leave groups quickly for various personal and other reasons. As a strategy to cope with this scenario, this study focused on three snapshots of the network over the campaign period in 2018. The data were collected at three different times: during the campaign, immediately after the first voting round, and immediately after the second round, which gave the final result of the electoral race. The data were collected on the following dates: September 17, 2018, when Bolsonaro promoted a series of demonstrations across the country, taking advantage of his candidate's party number being 17; October 7, 2018; and October 28, 2018.

After scraping the data, users' private data were anonymized, except metadata referring to the geographic origin of mobile phone numbers. In Brazil, in addition to international dialing codes, cell phones have local dialing codes, which allow a given phone number's region of origin to be identified. While there may be situations in which a user who owns a phone from the state of São Paulo is actually messaging from the state of Paraná, for example, these metadata still indicate important aspects for research with relational data. So, a phone number with the international dialing code +55 (Brazil) and the local dialing code 11 is classified as coming from the metropolitan region of São Paulo. The same procedure was used to identify users from other

regions of the country and from foreign countries. Based on a content analysis approach, groups were categorized according to a thematic-ideological typology, a functional typology, and a territorial typology.

In the first case, for the thematic typology, the groups were classified as (1) essentially supporting Bolsonaro, (2) integrated by followers of the self-proclaimed conservative philosopher Olavo de Carvalho, considered an intellectual pundit for Bolsonarism, (3) anticorruption activists, (4) defenders of the return of the monarchical regime, (5) defenders of the return of the military dictatorial regime, (6) anticommunist and conspiracy groups, (7) religious groups, notably neo-Pentecostal, or (8) other self-proclaimed right-wing supra-partisan groups.

The functional typology sought to identify groups according to their immediate purposes. In this case, the groups were classified as (1) misinformation broadcast groups and campaign agenda, (2) mobilization groups and call to action for campaign activities (such as marches, caravans, choreography organization, and the like), (3) debate and open political discussion groups, (4) apologetic and ideological propaganda groups, or (5) a single group with a gender focus (women who supported Bolsonaro), distinguished from the others because of its specificity. Regarding the misinformation patterns in Bolsonarist groups, not only can one rely on previous studies accounting for this as one of the most common activities in some of these groups (Chagas et al., 2019; Massuchin et al., 2021), but ethnographic observations also could verify that this was a prominent use in such cases. Most of these groups are characterized by an intense flow of links from far-right blogs and YouTube videos that attack political adversaries, spread hate speech against minorities, or disseminate conspiracy theories.

Finally, territorial typology was based on the enunciative description of each group, which often produced titles such as "Bolsonaro RJ" (RJ stands for the state of Rio de Janeiro) and "Bolsonaro–CityA" (CityA refers to either a metropolitan or inland urban place). Whenever a group was described as national, regional, or local, it was categorized this way in the content analysis. This inductive analysis was later compared with the deductive one, using the local dialing code metadata to recognize some patterns. Graphs presented in this study were created using Gephi software, arranged with Force Atlas 2 algorithm (Jacomy, Venturini, Heymann, & Bastian, 2014), and color-classified with content analysis classes. The distance between each node or each cluster was based on a statistical model simulating a physical system. Nodes and clusters repulse and attract each other according to their centrality in the whole network and the connections they establish (Jacomy et al., 2014). So, nodes and clusters presented as central mostly have high degrees and occupy a prominent position within the network, while peripheral ones can indicate marginal actors with fewer important connections. The results are shown in the following section.

Discussion

Data Overview

A first look at the data shows relevant aspects of the topology of this network. The snapshots, empirically oriented to cover the electoral period from three different time points, lead to variations in the number of members of the monitored groups, as well as in the relationship between admins and groups, an important aspect for astroturfing moderation dynamics. Table 2 shows that the number of group members

increased until the first voting round and then decreased as the end of the election approached. Elections in Brazil are held in two rounds when a candidate does not obtain more than half of the votes in the first. Therefore, if the first round has a large number of candidates, considering the Brazilian multiparty system, the election is limited to two candidates in the second round. It is usually during the second round that the negative campaigning and the rejection vote intensify. However, at this stage, it is common for voters to be less mobilized and for abstaining votes to generally increase. The data also show that some users are members of more than one group, and this average decreases over time.

Table 2. Overview of the Network (User Level).

	Campaign	After first round	After second round
Total users	12,352	12,794	10,092
Connections	15,284	15,537	12,230
Connections per user	1.237	1.214	1.212

Source: The author.

Table 3 presents data from the group level. Some groups became inactive, and others remained active during the analyzed period. Table 3 identifies these cases and indicates the average number of users per group on the whole network. A look at the time interval between the data extraction periods—20 days from September 17 to October 7, 2018, and 21 days from October 7 to October 28, 2018—indicates that the rate of the decrease in active groups was slower in the second half of the electoral process. On the other hand, the average number of users per group was considerably lower.

Table 3. Overview of the Network (Group Level).

	Campaign	After first round	After second round
Active groups at the time of data extraction	115	108	104
Users per group	99.6	103.2	81.4

Source: The author.

The first reports of mass messaging and irregular use of WhatsApp by the Bolsonaro campaign may have discouraged users and led to a larger exodus in the final period of the elections. But another piece of information, about the general structure of group administration, is also important because it sheds some light on a more specialized level of the network.

In the initial campaign period, there was an increase of 442 users on the network from one snapshot to another. These are voters who entered the groups through new invitations that circulated publicly. This increase is equivalent to approximately 3.5% of the number of users in the groups monitored on October 7. On the other hand, in the same period, the number of group admins jumped from 833 to 976, an increase of 14.7%. Briefly, the number of administrators increased more than four times in relation to the number of users. That is, the network became more controlled. In addition, some users managed more than one group, so *unique admins* refers to users who managed at least one group in the period. It should be noted that the proportion of number of total users to number of unique admins remained stable in the second analyzed time, as shown in Table 4. In addition, although there was an exodus of users after the first

electoral round, the number of admins who managed more than one group continued to increase. This suggests that the network was more hierarchical and less organic at the end of the elections, which may lead to an astroturfing model.

Table 4. Overview of the Network (Admin Level).

	Campaign	After first round	After second round
Admins	833	976	805
Unique admins	743	890	717
Unique admins per user	16.625	14.375	14.075

Source: The author.

It was also common for groups to present a reasonably organized visual structure. It was often possible to identify groups with identical cover images and standardized designations, as shown in Figure 1. This was because many of these groups were created and administered by the same users. For example, 99 users were responsible for creating the 124 monitored groups. Some alone created six groups. It also highlights, from a qualitative perspective, that the pace of creation of new groups each day doubled following the beginning of the official campaign period in Brazil, on August 16. In short, groups were often created in a series, managed by specialized users who moderated several and offered a customized visual identity. These characteristics point to some degree of coordination in the structure of the network and fit in an astroturfing strategy.

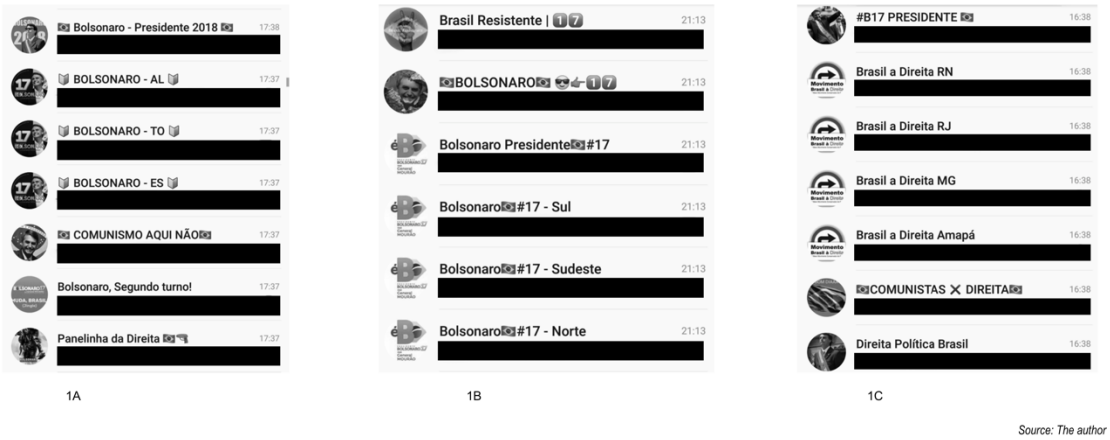

Figure 1. Visual structure of some groups. Source: The author.

Using Social Network Analysis Coupled With Content Analysis to Identify Different Ideological Nuances in the Network

SNA was coupled with content analysis to identify several trends in the sample. The first of these analyses sought to distinguish groups according to a thematic-ideological classification. The Bolsonarist

network is actually composed of different nuances, each representing very distinct and often conflicting interest groups.

While most groups (72.6%, *N* = 90) explicitly and exclusively supported Jair Bolsonaro's candidacy, other segments were equally relevant. The sample comprises 5.6% (*N* = 7) groups that identified themselves as conservative followers of self-proclaimed philosopher Olavo de Carvalho, 4.0% (*N* = 5) monarchist groups, 3.2% (*N* = 4) neo-Pentecostal religious groups, 2.4% (*N* = 3) militaristic groups, other 2.4% (*N* = 3) anticorruption groups, and 0.8% (*N* = 1) anticommunist and conspiracy groups; in addition, 8.9% (*N* = 11) identified with other right-wing trends. Recognizing how these nuances are articulated allows for better understanding of how the campaign coordination was structured.

Visually, these groups are arranged and color-classified in the graph shown in Figure 2. The monarchist groups form a small interconnected cluster at the bottom of the graph, while the Olavist groups are more spread out over the network, occupying intermediary strategic positions. Religious groups are typically more peripheral.

As noted earlier, not all groups perform the same functions. The next step was to classify these groups according to the functions they performed. In this case, an activist belt with several calls for participation, such as groups to organize demonstrations, caravans, and choreographies, to sell T-shirts, and to engage in many other activities, occupies a peripheral position compared with the campaign's hard core (Figure 3).

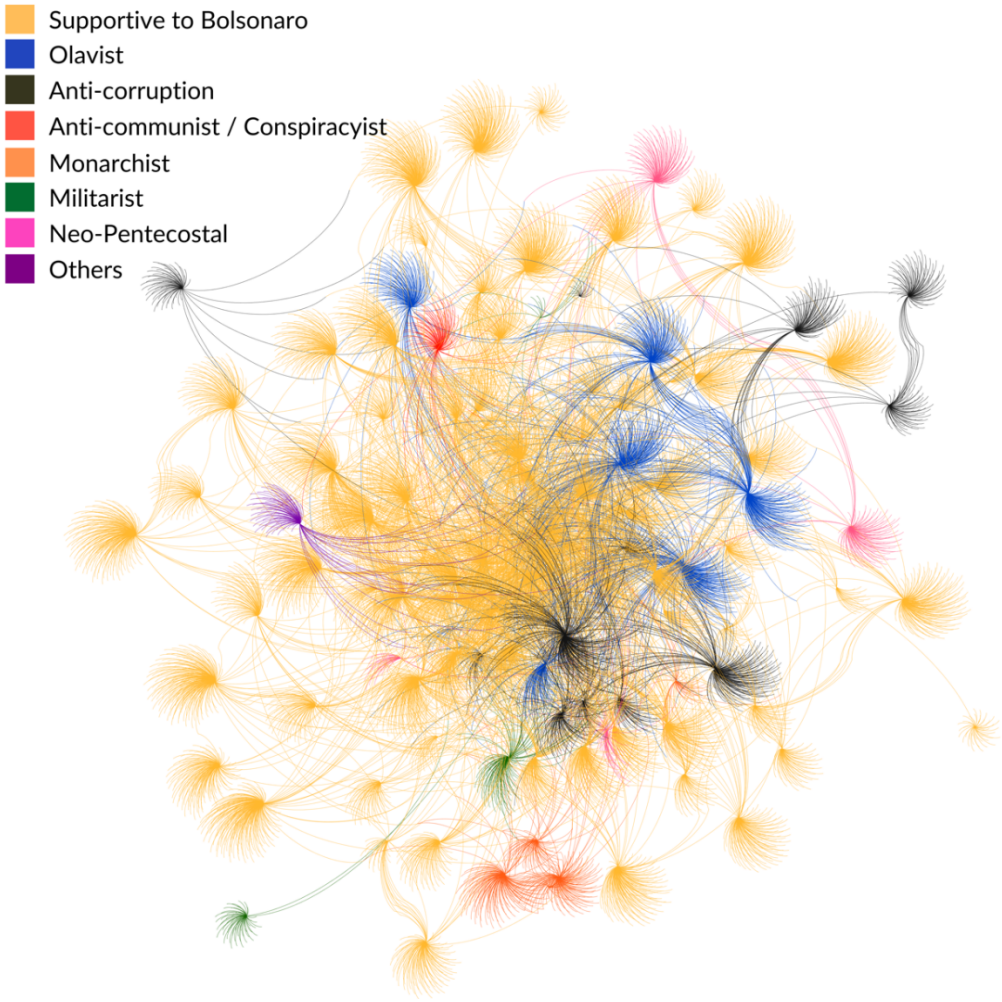

Figure 2. Network according to group theme. Source: The author.

Based on previous ethnographic observations, the research identified that groups spreading misinformation and engaging in political proselytism made up 70.2% (*N* = 87) of the sample, a proportion almost coincident with the core network of groups supportive of Bolsonaro. In parallel, the presence of 12.9% (*N* = 16) ideological propaganda groups can also be noted, as well as 12.1% (*N* = 15) debate and discussion groups, 4.0% (*N* = 5) groups for direct mobilization of voters, and, finally, 0.8% (*N* = 1) groups with a gender focus.

As Figure 3 shows, the direct mobilization groups are always marginal in relation to the others, while the misinformation groups occupy a highly central position. Again, there is an ideological belt around the most central groups, as if supporting their actions. They are monarchist, militaristic, and Olavist groups, which constitute, in a broad sense, the intellectual basis of Bolsonarism. One can assume the existence of

hardcore strategic planning, surrounded by an ideological circuit, and a sphere of action on the peripheries—a design that, once again, denotes a certain degree of coordination in the network.

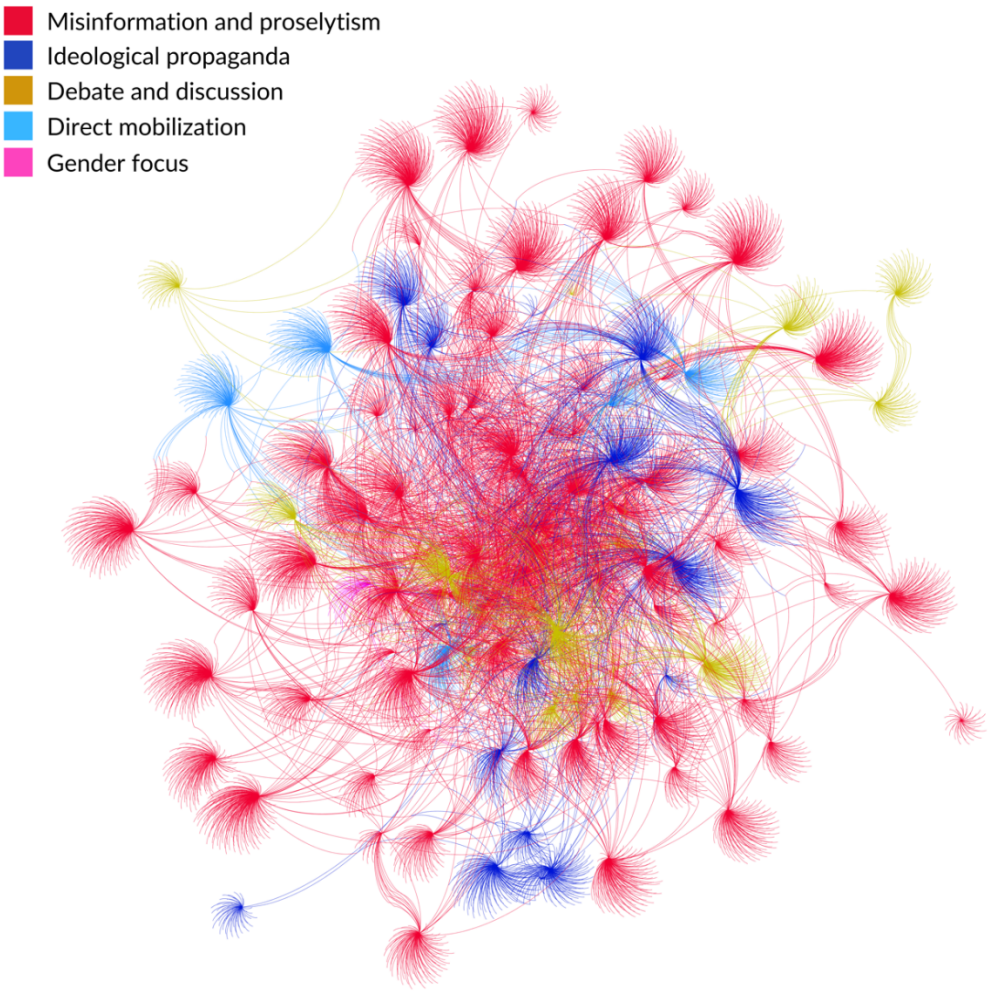

Figure 3. Network according to group function. Source: The author.

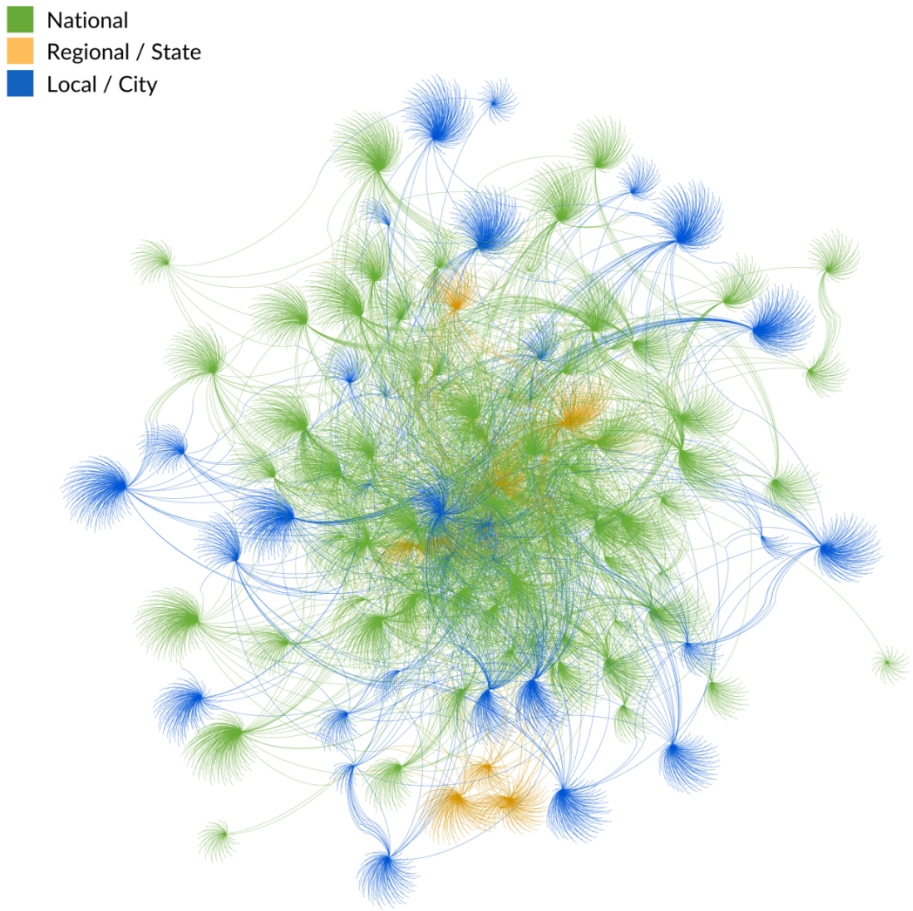

Figure 4. Network according to the aimed spatial coverage. Source: The author.

The territorial coverage of the groups was analyzed in two different ways. The first (Figure 4) considered the title and the description of the groups and their aimed coverage. The groups were distributed among national ("Bolsonaro Brasil"), regional ("Bolsonaro #17–Southeast"), or local ("Bolsomito Araguatins/TO") categories. National groups (65.3%, *N* = 81) occupied a greater centrality in the topological arrangement of the network. They were responsible for distributing information to groups of a regional (7.3%, *N* = 9) and local (27.4%, *N* = 34) character. The visual arrangement of these data in the graph presents a kind of hierarchical organization in which national groups carry the news to regional ones, and these to local groups. The investigation was unable to account for private WhatsApp groups, such as groups of family members and friends, but it is assumed that they occupy an even more external position in this chain.

The second procedure (Figure 5) took into account the phone dialing codes. Here, the local and regional groups present greater homogeneity among their members, whereas the national groups have a more "cosmopolitan" nature, including members with multiple origins.

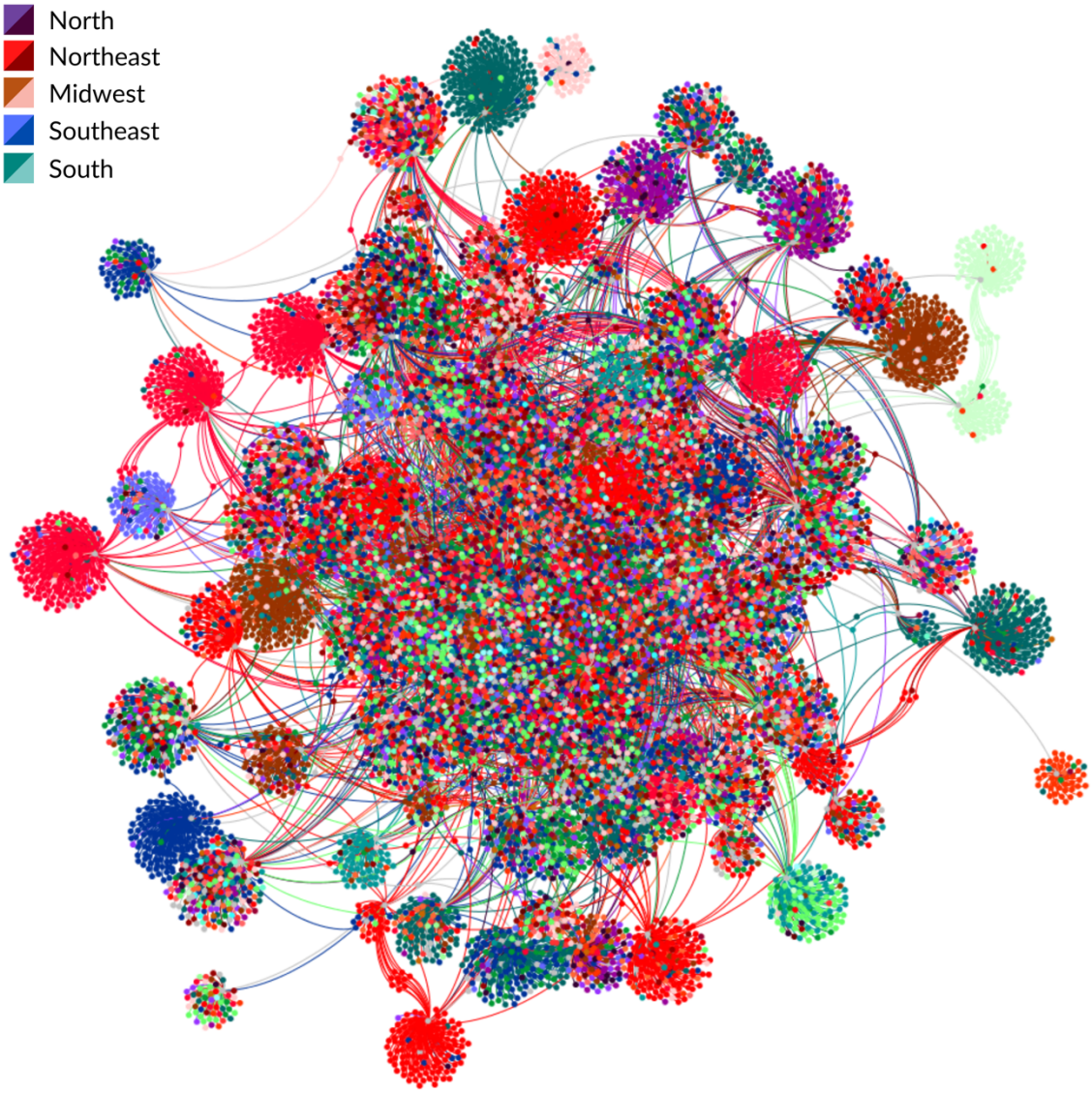

Figure 5. Network according to dialing codes. Source: The author.

The SNA, coupled with the content analysis, reinforces that, instead of having a single Bolsonarist network acting in unison, as common sense might predict, one can recognize a large set of trends of the conservative electorate (monarchists, militarists, anticommunists, etc.) and the different functions they perform. As suggested by H1, if applied to a relational database obtained empirically from a previous ethnographic effort, these methods provide a clear overview of the cleavages of these interest groups. It is also possible to identify patterns of network organization and coordination, typical of astroturfing actions. These data are further explored next.

Using Social Network Analysis Metrics to Identify Astroturfing Practices

In the last series of analyses based on SNA metrics, only the data obtained from the most intense period in the race—that is, after the first electoral round—were considered. For the sake of simplicity, only the October 7 snapshot was observed, which led to 890 unique admins—that is, users who, at that date, were administrators of at least one of the observed groups.

The first correlation concerns the number of groups that the same user managed and the number of groups in which the user participated. In the sample of 890 users, the range consisted of users who participated in only one group to those who were in as many as 22 groups simultaneously, and users who administrated between one and six groups simultaneously. The first metric is the so-called outdegree, usually taken in SNA, and the second metric is called the admin degree. As Figure 6 shows, although there were users who administrated a single group but participated in multiple groups, there was a tendency for the average number of groups that a user administrated to increase according to the number of groups in which he participated. This curve suggests that the influence of the administrators of political discussion groups was noticed by the community.

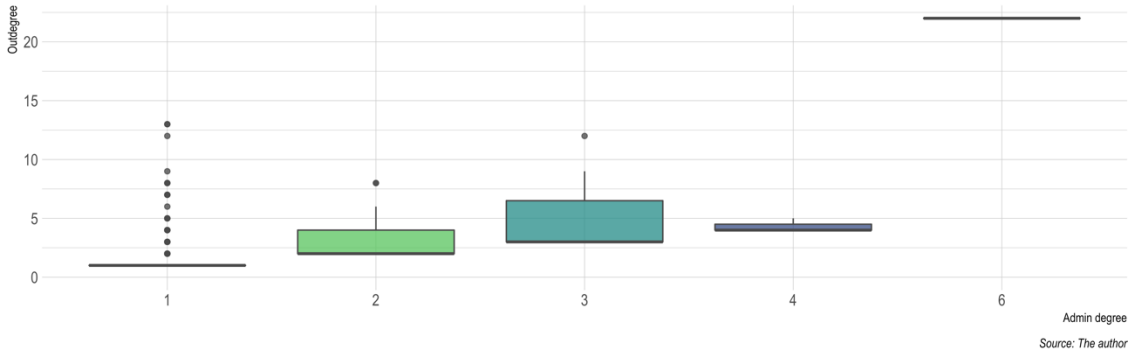

Figure 6. Admin degree versus outdegree. Source: The author.

But the correlation becomes stronger if one looks at the position that these users occupy visually in the SNA graphs, based on their eigenvector centrality. The eigenvector centrality is a measure of the influence of a node on a given network that takes into account that connections to more central nodes make the connected nodes more relevant to the whole network. Thus, in a given network, if node A occupies a central position and node B a more peripheral position, any connection to A tends to make a third node more central, and any connection to B tends to make it more peripheral. The result of this statistic is a linear scale that distributes the nodes in this network according to their relational importance.

As Figure 7 demonstrates, the number of groups in which a user participates is significantly correlated (R^2 = 0.8669) with his eigenvector centrality in the network. Thus, users who participate in multiple groups are potentially able to disseminate misinformation more quickly and efficiently because they occupy more strategic positions as central broadcasting actors. In the observed groups, the main senders can be represented by a

power law curve (Figure 8). That there is a remarkable discrepancy between a small set of users who distribute the most messages while the majority receive more than they send, and a similar discrepancy in the centrality occupied by some actors in this network, suggests again that it is far from a spontaneous and balanced structure in terms of engagement. Instead, it suggests a high degree of hierarchy and coordination.

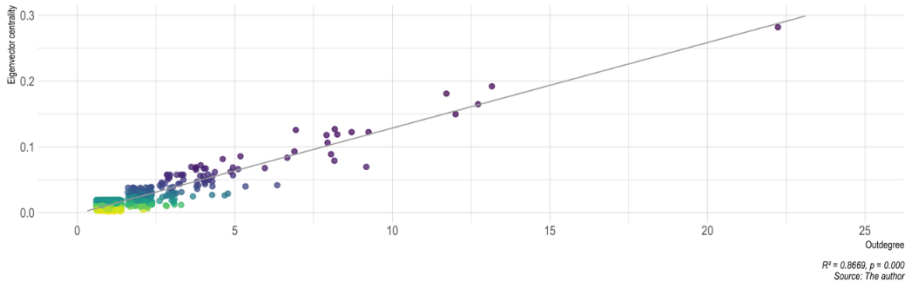

Figure 7. Eigenvector centrality versus outdegree. Source: The author.

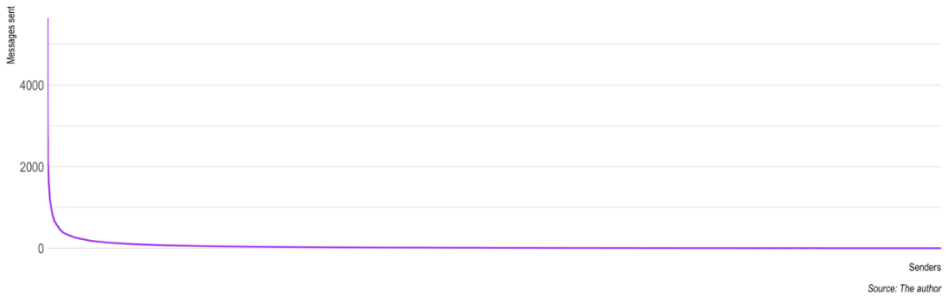

Figure 8. Messages sent per user. Source: The author.

Thus, corroborating H2, the SNA metrics can identify behaviors that deviate from the pattern of distributed networks. In the case of WhatsApp, participating simultaneously in 22 political discussion groups, for example, means that one can potentially reach more than 5,600 users in a few seconds.

It is also noteworthy that, in both Figures 6 and 7, there is an outlier who participated in 22 groups (17.7% of the sample) and administrated six of them. Although this investigation does not work with data on an individual scale, this is the only case that implies an exception; it is a public phone number, used and disclosed during the campaign on the official website of candidate Eduardo Bolsonaro, son of Jair Bolsonaro (see Figure 9). The data suggest that Eduardo Bolsonaro was one of the main organizers of the astroturfing campaign on WhatsApp during the 2018 Brazilian elections. The presence of his mobile phone in this sample fulfills the expectation that the Bolsonarist network on WhatsApp was also infiltrated by agents from the

professional field of politics despite repeated denials by the group itself. As seen, the attempt to hide this participation and the highly hierarchical structure of the network are typical elements of the conceptual definition of astroturfing.

Figure 9. Eduardo Bolsonaro's official website showing his WhatsApp number. Source: Bolsonaro (2018).

Final Remarks

Although a not irrelevant set of studies have already discussed digital astroturfing through a solely theoretical lens, little has been done for addressing astroturfing as an empirical phenomenon. The clandestinity that surrounds astroturfing actions is undoubtedly an important challenge. Especially with regard to WhatsApp, a platform deeply characterized by its environmental opacity, the absence of metadata is a great obstacle to realizing the development of coordinated, nonspontaneous, and politically oriented actions. SNA can unveil some aspects of this type of practice.

One of this article's research questions was how to recognize when astroturfing takes place on digital platforms. Its main hypotheses established that SNA and large-scale monitoring of inauthentic behavior on platforms such as WhatsApp can lead to a less opaque scenario in digital campaigns. Through the analysis of statistics obtained from relational data, this study sought to explore methods that allow for recognizing the development of astroturfing actions, focusing on the 2018 Brazilian elections as a case study. SNA was important to identify the functional character of some clusters and actors in the Bolsonarist network, to recognize super posters and super admins, and to understand how this digital campaign model combined an allegedly horizontal space with coordinated actions, often sponsored by professional agents from the field of politics.

The use of SNA and community detection algorithms, together with some categorical analysis, can help identify inauthentic behavior. Digital platforms, however, have focused their moderation efforts on content, without looking with the same efficiency at how some groups of users collude to carry out certain actions. This is where SNA, unlike other methods, can be particularly effective.

While WhatsApp does not have a wide audience in all countries, the use of a mobile instant messaging service in the context of an electoral race is absolutely relevant for understanding the effects of

these tools in the democratic arena. Political communication studies have neglected the role of WhatsApp, but relational data can present in detail how this app is used as a tool for mobilization and political engagement. Although it is a particularly opaque environment, one can access, through SNA metrics, the degree of relevance and influence of an actor in a given network, how content circulates, and how users are associated with each other. This article sheds some light on how SNA can, even with a scarcity of metadata, help in efforts to recognize ongoing astroturfing actions. It is also a contribution to building a connection between theoretical and empirical observations of digital astroturfing.

This research has some limitations. Among them is the fact that the data come from a nonprobabilistic sample scraped in a semiautomated way from the platform. In addition, this systematic monitoring was based on a covert research strategy, which, although endorsed by an ethics committee, presents, concerns regarding the private nature of extracted data. To circumvent these barriers, a research strategy that preserves users' privacy was undertaken. Understanding the dynamics of this kind of network, therefore, is also a methodological challenge.

An important issue to be highlighted is that SNA alone also presents some limitations related to astroturfing research. This method is useful for detecting communities, but one needs to combine it with other approaches. Thus, SNA seems to work better if combined with categorical content analysis and ethnographic observations, as is the case in this study.

On the other hand, digital platforms have largely based their policies on moderating harmful content. SNA allows one to look not only at the contents, but also at their economy of production and circulation, encompassing the actors and their network of influence.

This study may also lead to a more accurate understanding of the organizational model of the Bolsonarist campaign on WhatsApp in the 2018 Brazilian elections. In the future, it is important to understand how WhatsApp has been used by this same group as a tool to establish a permanent campaign mood.

References

Arnaudo, D. (2017). *Computational propaganda in Brazil: Social bots during elections*. Oxford, UK: Project on Computational Propaganda.

Arun, C. (2019). On WhatsApp, rumours, and lynchings. *Economic & Political Weekly, 54*(6), 30–35.

Barbosa, S., & Milan, S. (2019). Do not harm in private chat apps: Ethical issues for research on and with WhatsApp. *Westminster Papers in Communication and Culture, 14*(1), 49–65. doi:10.16997/wpcc.313

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford, UK: Oxford University Press.

Bolsonaro, E. (2018). *Eduardo Bolsonaro*. Retrieved October 28, 2018 from http://www.eduardobolsonarosp.com.br

Caetano, J., Oliveira, J., Lima, H., Marques-Neto, H., & Magno, G. (2018, February). *Analyzing and characterizing political discussions in WhatsApp public groups*. Paper presented at the Association for the Advancement of Artificial Intelligence Conference, New Orleans, LA. Retrieved from https://arxiv.org/abs/1804.00397

Chadwick, A. (2013). *The hybrid media system: Politics and power*. New York, NY: Oxford University Press.

Chagas, V. (2019). Digesting the indigestible: The political scatology of the Vomitaço. *Galáxia, 40*, 41–56. doi:10.1590/1982-25542019137545

Chagas, V., Mitozo, I., Barros, S., Santos, J. G. B., & Azevedo, D. (2021). The "new age" of political participation? WhatsApp and call to action on the Brazilian senate's consultations on the e-cidadania portal. *Journal of Information Technology & Politics*. Advance online publication. doi:10.1080/19331681.2021.1962779

Chagas, V., Modesto, M., & Magalhães, D. (2019). O Brasil vai virar Venezuela: Medo, memes e enquadramentos emocionais no WhatsApp pró-Bolsonaro [Brazil will become Venezuela: Fear, memes and emotional framings in pro-Bolsonaro WhatsApp]. *Esferas, 14*, 1–17. doi:10.31501/esf.v0i14.10374

Chakrabortty, A. (2020, December 9). Brexit was never a grassroots movement, but an elitist political takeover. *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2020/dec/09/brexit-grassroots-movement-elitist-takeover-2016-referendum

Christensen, H. S. (2011). Political activities on the Internet: Slacktivism or political participation by other means? *First Monday, 16*(2). doi:10.5210/fm.v16i2.3336

Datafolha: 6 em cada 10 eleitores de Bolsonaro se informam pelo WhatsApp [Datafolha: 6 out of 10 Bolsonaro voters get their information via WhatsApp]. (2018, October 3). *Veja*. Retrieved from https://veja.abril.com.br/politica/datafolha-eleitor-de-bolsonaro-e-o-que-mais-se-informa-por-redes-sociais

Friedberg, B., & Donovan, J. (2019). On the Internet, nobody knows you're a bot: Pseudoanonymous influence operations and networked social movements. *Journal of Design and Science, 6*.

Gomes, W., Fernandes, B., Reis, L., & Silva, T. (2009). "Politics 2.0": A campanha on-line de Barack Obama em 2008 ["Politics 2.0": The online campaign of Barack Obama in 2008]. *Revista de Sociologia e Política, 17*(34), 29–43. doi:10.1590/S0104-44782009000300004

Graham, T., & Wright, S. (2013). Discursive equality and everyday talk online: The impact of "Superparticipants." *Journal of Computer-Mediated Communication, 19*(3), 625–642. doi:10.1111/jcc4.12016

Henrie, K., & Gilde, C. (2019). An examination of the impact of astroturfing on nationalism: A persuasion knowledge perspective. *Social Sciences, 8*(38), 1–11. doi:10.3390/socsci8020038

Hous, D. (2018, October 20). WhatsApp bane centenas de milhares de contas por uso irregular do aplicativo [WhatsApp bans hundreds of thousands of accounts for irregular use of the app]. *Folha de S. Paulo*. Retrieved from https://www1.folha.uol.com.br/poder/2018/10/whatsapp-bane-100-mil-por-uso-irregular-do-aplicativo.shtml

Howard, P. (2006). *New media campaigns and the managed citizen*. New York, NY: Cambridge University Press.

Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi Software. *PLoS ONE, 9*(6), e98679. doi:10.1371/journal.pone.0098679

Kischinhevsky, M., Vieira, I. M., Santos, J. G. B., Chagas, V., Freitas, M., & Aldé, A. (2020). WhatsApp audios and the remediation of radio: Disinformation in Brazilian 2018 presidential election. *Radio Journal, 18*(2), 139–158. doi:10.1386/rjao_00021_1

Klang, M., & Madison, N. (2016). The domestication of online activism. *First Monday, 21*(6). doi:10.5210/fm.v21i6.6790

Kovic, M., Rauchfleisch, A., Sele, M., & Caspar, C. (2018). Digital astroturfing in politics: Definition, typology, and countermeasures. *Studies in Communication Sciences, 18*(1), 69–85. doi:10.24434/j.scoms.2018.01.005

Lee, C. (2010). The roots of astroturfing. *Contexts, 9*(1), 73–75. doi:10.1525%2Fctx.2010.9.1.73

Leiser, M. (2016). AstroTurfing, "cyberturfing" and other online persuasion campaigns. *European Journal of Law and Technology, 7*(1), 1–27.

Lewis, R. (2018). *Alternative influence: Broadcasting the reactionary right on YouTube*. Oxford, UK: Data & Society.

Massuchin, M., Tavares, C., Mitozo, I., & Chagas, V. (2021). Scientific institutions and the journalistic model when appropriate: The argumentative structure about discrediting science on WhatsApp during the COVID-19 pandemic. *Fronteiras, 23*(2), 160–174. doi:10.4013/fem.2021.232.11

Mello, P. C. (2018, October 18). Empresários bancam campanha contra o PT pelo WhatsApp [Entrepreneurs fund campaign against PT via WhatsApp]. *Folha de S. Paulo*. Retrieved from https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contra-o-pt-pelo-whatsapp.shtml

Moura, M., & Michelson, M. (2017). WhatsApp in Brazil: Mobilising voters through door-to-door and personal messages. *Internet Policy Review, 6*(4). doi:10.14763/2017.4.775

National Health Council (BR). (2016, April 7). *Resolution No. 510*. Retrieved from http://conselho.saude.gov.br/resolucoes/2016/Reso510.pdf

Pang, N., & Woo, Y. T. (2020). What about WhatsApp? A systematic review of WhatsApp and its role in civic and political engagement. *First Monday, 25*(12). doi:10.5210/fm.v25i12.10417

Passos, P. (2018, October 26). Metade dos usuários do WhatsApp diz acreditar em notícias compartilhadas [Half of WhatsApp users say they believe in shared news]. *Folha de S. Paulo*. Retrieved from https://www1.folha.uol.com.br/poder/2018/10/metade-acredita-em-noticias-compartilhadas-no-whatsapp.shtml

Piaia, V., & Alves, M. (2020). Opening the black box: Exploratory analysis of the Bolsonarista network on WhatsApp. *Intercom RBCC, 43*(3), 135–153.

Ratkiewicz, J., Conover, M., Meiss, M., Goncalves, B., Flammini, A., & Menczer, F. (2011). Detecting and tracking political abuse in social media. *Proceedings of the International AAAI Conference on Web and Social Media, 5*(1), 297–304. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14127

Resende, G., Melo, P., Sousa, H., Messias, J., Vasconcelos, M., Almeida, J., & Benevenuto, F. (2019). (Mis)Information dissemination in WhatsApp: Gathering, analyzing and countermeasures. In L. Liu & R. White (Eds.), *The Web conference: WWW '19* (pp. 818–828). San Francisco, CA: Association for Computing Machinery. doi:10.1145/3308558.3313688

Reuters Institute. (2019). *Digital news report 2019*. London, UK: Reuters. Retrieved from https://www.digitalnewsreport.org/survey/2019

Rossini, P., Stromer-Galley, J., Baptista, E., & Veiga de Oliveira, V. (2021). Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections. *New Media & Society, 23*(8), 2430–2451. doi:10.1177%2F1461444820928059

Santos, J. G. B. (2019). Mobile networks and the Brazilian 2018 presidential election: From technological design to social appropriation. *Brazilian Studies Program One Pager, 2*, 1–2.

Silva, D. R. (2015). *Astroturfing: Lógicas e dinâmicas de manifestações de públicos simulados* [Astroturfing: Logics and dynamics of simulated publics' demonstrations]. Belo Horizonte, Brazil: Selo PPGCOM.

Stromer-Galley, J. (2000). On-line interaction and why candidates avoid it. *Journal of Communication, 50*(4), 111–133. doi:10.1111/j.1460-2466.2000.tb02865.x

Treré, E. (2020). The banality of WhatsApp: On the everyday politics of backstage activism in Mexico and Spain. *First Monday, 25*(12). doi:10.5210/fm.v25i12.10404

Vie, S. (2014). In defense of "slacktivism": The Human Rights Campaign Facebook logo as digital activism. *First Monday, 19*(4). doi:10.5210/fm.v19i4.4961

Woolley, S., & Howard, P. (2018). Introduction: Computational propaganda worldwide. In S. Woolley & P. Howard (Eds.), *Computational propaganda: Political parties, politicians, and political manipulation on social media* (pp. 3–18). New York, NY: Oxford University Press.

Zhang, J., Carpenter, D., & Ko, M. (2013, August). *Online astroturfing: A theoretical perspective*. Paper presented at the 19th Americas Conference on Information Systems, Chicago, IL. Retrieved from https://www.researchgate.net/publication/286729041_Online_astroturfing_A_theoretical_perspective

Zúñiga, H. G., Arvèvol-Abreu, A., & Casero-Ripollés, C. (2019). WhatsApp political discussion, conventional participation and activism: Exploring direct, indirect and generational effects. *Information, Communication & Society, 24*(2), 201–218. doi:10.1080/1369118X.2019.1642933