Coordinated campaigns on Twitter during the coronavirus health crisis in Mexico

C. A. Piña-García a and A. Espinoza b

aLaboratorio para el Análisis de Información Generada a través de las Redes Sociales en Internet (LARSI), Centro de Estudios de Opinión y Análisis, Universidad Veracruzana, Xalapa, México; bColegio de Estudios Avanzados de Iberoamérica, Xalapa, México

ABSTRACT
Social media is fast becoming a key instrument to manipulate or influence social perception. Digital platforms are having a serious effect on the manipulation of public opinion through the spread of political propaganda and message amplification via coordinated campaigns. As one of the most used social platforms among politicians and democratic governments, Twitter plays a critical role in how information flows through trending topics. The main purpose of this study is to explore how coordinated campaigns, in this case astroturfing, were used to influence and manipulate public opinion during the coronavirus health crisis in Mexico. Our research provides new insights into the early detection of astroturfing and artificial amplification, in order to expose the efforts to manipulate online discourse in Mexico. In the pages that follow, it will be argued that Mexico is currently experiencing online manipulation through malicious strategies that may threaten its democracy. The following hashtags were used to explore and compare our approach in Mexico: #GatellOrgulloMexicano (Gatell Mexican Pride) and #AMLOPresidenteDeLaSalud (AMLO President of Health). This study intends to build awareness and to improve the public’s understanding coordinated behavior on Twitter.

RESUMO
A mídia social está rapidamente se tornando um instrumento-chave para manipular ou influenciar a percepcão social. As plataformas digitais têm um forte efeito na manipulação da opinião pública, por meio da disseminação de propaganda política e ampliação de mensagens, administradas por campanhas governamentais coordenadas. O Twitter é considerado uma das plataformas sociais mais utilizadas entre políticos e governos democráticos, portanto desempenha um papel crítico em termos de como a informação flui através das tendências. O objetivo principal deste manuscrito é explorar como campanhas coordenadas (astroturfing) foram usadas para influenciar e manipular a opinião

KEYWORDS
Astroturfing; Twitter; COVID-19; coordinated campaigns; Mexico

PALAVRAS-CHAVE
Astroturfing; Twitter; COVID-19; campanhas coordenadas; México

PALABRAS CLAVE
Astroturfing; Twitter; COVID-19; campañas coordinadas; México

CONTACT
C. A. Piña-García cpina@uv.mx
1. Introduction

During the recent coronavirus pandemic (Taylor 2021), rapid access to information has been a key factor in communication. The coronavirus disease (COVID-19) has received considerable attention due to its spread to nearly every country in the world since it was first identified in China at the end of 2019. According to the World Health Organization (WHO), one of the best ways to prevent and slow down transmission is to be well...
informed. Therefore, the access to reliable sources of information becomes crucial to this health crisis (WHO 2021).

The changes experienced by social media over the past year have seen increasingly rapid developments in the ways we interact socially. It is important to note that lockdown measures increased the use of social platforms such as WhatsApp and Facebook, with increased usage of around 50% around the world; streaming services like Netflix grew 66% in Italy and 35% in Spain in terms of new downloads (Aral 2020); the short-video app TikTok tripled its user base in Mexico from 4.6 million in 2019 to 17 million in 2021 (Dolan 2021); and in Brazil 40% of Internet users rely on social media for information on COVID-19 (Dolan 2020).

The coronavirus pandemic has also heightened the need for information searches through social media. Users’ engagement such as reactions, sharing information, and content consumption related to the COVID-19 topic rose to a high point in mainstream platforms, e.g. Twitter (30 January 2020), YouTube (31 January 2020), and Instagram (5 February 2020) (Cinelli et al. 2020). In this regard, Twitter has become a popular platform for sharing information rapidly.

However, according to Badawy et al. (2019), it is becoming extremely difficult to ignore the existence of malicious strategies that mislead and promote fake news. Lazer et al. (2018) define fake news as fabricated information that simulates news media content. Moreover, according to Varol et al. (2017), it is estimated that between 9% and 15% of active Twitter accounts are “bot” accounts that manipulate retweets and mentions to endorse a specific hashtag. Similarly, Stella, Ferrara, and De Domenico (2018) have established that 19% of non-human interactions in Twitter are directed from bots to humans, mainly through retweets (74%) and mentions (25%).

Since it was reported in 2016, fake news on Twitter during presidential elections has attracted significant interest (Grinberg et al. 2019). Fake news and coordinated campaigns have seen a growing inclination towards adopting misinformed beliefs. Thus, the main challenge faced by social network research is uncovering such malicious behavior (Pacheco et al. 2020). Previous studies (Che et al. 2018; Herik and Dimov 2012) have indicated that a significant amount of information shared on Twitter emerges from users through cybercascades.

A cybercascade refers to a massive snowball-effect of information that gets spread on the Twitter network via retweets triggered by opinion leaders or “influencers” (Kurka, Godoy, and Zuben 2016). A correlation between strategies encouraging online participation and anomalous accounts or social bots has been found (Savage, Monroy-Hernández, and Höllerer 2016). Social bots are an important issue among social networks researchers; traditionally, these accounts are defined as artificial agents – most of the time only reactive agents – that continuously interact and promote a specific trend. Their main task is to pull or persuade organic users into information or opinions that endorse their ideological homogeneity (echo chambers).

It is necessary here to clarify terms such as misinformation, disinformation, propaganda, manipulation, astroturfing, and coordinated campaigns. Misinformation refers to a false, unverified, or inaccurate information that is deliberately created and is intentionally or unintentionally propagated (Wu et al. 2019); disinformation can be defined as a false information that is spread with the sole intention to mislead (Shu et al. 2020). According to DeVito (1986, 239); propaganda is described as an ideology spread to an
audience with the purpose of manipulating it with controlled information. In this regard, different definitions of manipulation have been proposed. For instance, Barnhill (2014, 51) states that defining manipulation must include several varieties and exclude non-manipulative forms, such as coercion and persuasion. Manipulation could be understood as “influencing the individual’s choice to accomplish the outcome sought by the manipulator” (Wood 2014, 31), while a person is coerced when they do “something that it does not want to, and when it does choose to do it, it does it because it has no acceptable alternative” (Wood 2014, 21); persuasion is “the process by which a person’s attitudes or behavior are, without duress, influenced by communications from other people” (Brittanica 2015). Therefore, the term manipulation will be used here to refer to the promotion of misleading narratives and false information via social media and the influence on the users on Twitter.

Astroturfing is defined as the attempt to create an impression of widespread genuine grassroot movement for a policy, individual, or product. Multiple online identities and fake pressure groups are used to mislead the public into believing that the position of the astroturfer is the commonly held view. Additionally, it refers to a coordinated campaign, particularly through social media, with the aim to promote and amplify propaganda, fake news, and misinformation (Bienkov 2012; Keller et al. 2017). One of the most interesting patterns found in this type of strategy is based on the artificial amplification of a hashtag (#) by retweeting several times a few tweets in a short time. The key challenge of astroturfing is to emulate organic growth and redirect public attention.

Likewise, a social media campaign is described as the planning, execution, and analysis of strategies and actions to pursue the main objective across multiple social networks (Winterer 2021). Another significant aspect of coordinated campaigns is that some of them are based on astroturfing: this type of campaign is designed to manipulate trends and conversations with the aim to create the perception that an idea or opinion is widely shared, supported, and accepted among a large group of users. From here onwards, we will refer to astroturfing as a type of coordinated campaign.

Zerback, Töpfl, and Knöpfle (2021) state that a false sense of widespread public support for particular issues may be established by using fake accounts or profiles known as “sock puppets,” or coordinated social bots that “push” the online strategy. In addition, according to Lits (2020; 2021), the main issue with astroturfing is that the sources of communications remain hidden and pretend to mimic ordinary user behavior, misleading by pretending they have public support. Twitter has been used to manipulate information using astroturfing, and in this context the astroturfer relies on genuine users to propagate a misleading message (Ratkiewicz et al. 2011). This can manipulate and influence public opinion, a phenomenon particularly relevant when social bots send and reproduce (retweets) with a political message (Keller and Klinger 2019).

During 2009 H1N1 pandemic, Mexican authorities were praised for their effective handling of that health crisis (Sánchez-Talanquer et al. 2021), unlike the handling of the coronavirus pandemic, with the first COVID-19 case officially confirmed on February 28th in Mexico (González 2021). There has been significant concern about the inadequacy of the information related to the number of coronavirus cases and the true risk level in Mexico. Despite maintaining a significant flow of information, systematic failures such as low testing rates and wrong decisions by political leaders on face mask wearing,
have contributed to the devastating impact of the pandemic in Mexico (Sánchez-Talanquer et al. 2021).

The face mask issue has been particularly addressed by Mexican president, Andrés Manuel López Obrador, and by Deputy Minister of Health, Hugo López Gatell, who at the beginning of the pandemic stated that the use of a face mask was only a preventive measure that should not be considered mandatory (Nolasco 2020). In examining this topic, the main context in this study is how digitally coordinated campaigns have intended to undermine and disrupt public opinion about the health crisis in Mexico.

1.1. Political Twitter context

Persuasion can play an important role in addressing the issue of manipulation in social media: a persuader can mislead an audience and when this happens, the audience will be manipulated (without knowing it) with the propaganda provided by the persuader (Jowett and O’Donnell 2019, 36). According to López and Piña-García (2017), political participation through Twitter is generally accomplished by growing negative sentiments towards political actors. On the other hand, Enli and Skogerbo (2013) find that Twitter can be a space used for dialogue among political actors and users. Moreover, Twitter is mainly used to debate and spread information via a well-coordinated retweet function and some studies have attempted to assess its impact on manipulation, disruption, and influence (Brachten et al. 2017). Artificial support is frequently used to amplify a political message, and in some cases, bot-like behavior is found in organized retweet rooms. Abokhodair, Yoo, and McDonald (2015) have shown that bots are programmed to flood the Twitter timeline with unrelated hashtags, but with related content (smoke screening) or by driving users’ attention away from a current topic (misdirection).

It is widely acknowledged that different governments and political actors make use of coordinated campaigns with the aim to manipulate public opinion, promote political messages, and disrupt political conversations. Forelle et al. (2015) found that astroturfing has been applied in different countries, e.g. in Venezuela, bots generated a small portion of the political discussion in social media operated by Venezuela’s radical opposition. Another example was found in South Korea, where political astroturfing has been used to spread propaganda (Keller et al. 2017). Elmas et al. (2019) detected astroturfing in Turkey that attempted to manipulate local political trends on Twitter. In 2017, social bots were used to potentially influence and manipulate social media users during the German Bundestag elections (Brachten et al. 2017). Finally, King, Pan, and Roberts (2017) reported that the Chinese regime has used astroturfing strategies to artificially promote inauthentic behavior to appear popular.

Very little attention has been paid to the role of coordinated campaigns in Mexico. This lack of information indicates a need to understand how these campaigns are used to influence and manipulate social media platforms. A few Mexican cases related to social media manipulation can be found in Gutiérrez (2016), who found that in 2016 the Tribunal Electoral del Poder Judicial de la Federación, (TEPJF – Federal Electoral Court) fined the Partido Verde (Green Party) for sponsoring coordinated campaigns with the following hashtags: #VamosVerdes (Go Greens) and #ApyemosALSVerdes (Let’s Support the Greens). Similarly, Fregoso (2018) reported one of the most important examples of how astroturfing is used to distort political debate during elections.
In 2018, the Mexican Ministry of Economy coordinated the campaign #YaSabesQuién (You Know Who) with the main goal of shaping the conversation to influence doubtful voters (Navarro 2019). In May 2020 Article 19, reported that a coordinated campaign was launched using personal Twitter accounts of Mexican news Agency Notimex staff (Article 19 2020). Finally, in August of 2021, after being appointed cultural attaché at the Embassy of Mexico in Spain, the writer Brenda Lozano was the target of astroturfing (harassment) with the hashtag #RenunciaBrendaLozano (Quit Brenda Lozano) (Tello 2021).

Our research presents two cases where astroturfing was used to influence and manipulate the public opinion. We seek to bring to public light how these hashtags can be artificially amplified to disrupt online perceptions. The rest of the paper is structured as follows; the Material and Methods section will give an account of how Twitter data were collected, and the methodology used to uncover coordinated campaigns via astroturfing. The Results and Discussion section will describe in a systematic and detailed way the outcomes in terms of Twitter activity with the aim to gain insights about coordinated behavior. Finally, a summary explaining the significance of our findings will be highlighted in the Conclusions.

2. Material and methods

2.1. Twitter data collection

One of the most well-known tools for assessing Twitter data is through its application program interface (API). Previous studies have based their results on this interface (Miller et al. 2014; Pacheco et al. 2020; Piña-García and Ramírez-Ramírez 2019). First, we collected an initial sample of tweets that triggered both hashtags, but it is important to note that a coordinated campaign always pretends to cover its temporal patterns when it is introduced.

The use of quantitative case studies is a well-established approach in terms of astroturfing (Kovic et al. 2018; Elmas et al. 2019). In this regard, we propose to explore those trends that achieved a notable degree of political engagement and exposure after a year of the coronavirus pandemic in Mexico. We chose the following hashtags: #GatellOrgulloMexicano (Gatell Mexican Pride), triggered on 24 May 2020 (Story Wrangler n.d.a), and #AMLOPresidenteDeLaSalud (AMLO President of Health), triggered on February 6, 2021 (Story Wrangler n.d.b). These hashtags were examined specifically during the second wave of the pandemic, when the number of deaths reached a peak and over 400,000 excess deaths had already accumulated. In this context, many Mexicans blamed the government for overwhelmed hospitals and the number of deaths related to coronavirus.

Only those tweets that contained the hashtags #GatellOrgulloMexicano or #AMLOPresidenteDeLaSalud were included in our dataset. Tweets were classified into two groups according to their hashtags and dates of creation related to the second wave of the pandemic. The main criterion for selecting those hashtags was based on the “pump and dump” behavior described by Pacheco et al. (2020). We hypothesize that coordinated pump and dump campaigns are highly vulnerable to central coordination: a few controlled accounts (sock puppets) with a significant number of followers “pump” a
specific hashtag and then the rest of their followers immediately use the retweet and like functions to amplify it (dump), producing a cyber-cascade effect. In this study, we have explored how suspicious groups of accounts were able to coordinate different posts with highly similar sequences of hashtags across multiple tweets. Therefore, this sort of pattern denotes a first temporal feature that needs to be explored due to its accelerated growth in terms of retweets.

Table 1 shows the collected information related to the hashtags: the number of tweets, the date of collection, its language, and geolocation data. It should be highlighted that the number of collected tweets corresponds only to the early detection of a coordinated campaign. Regarding the pump and dump behavior, the number of tweets could be considered low. However, the key concept of this method is the potential use of multiple accounts to post messages in close temporal proximity, since “the shorter the time interval in which two tweets are posted, the less likely they are to be coincidental,” Pacheco et al. (2020).

To corroborate whether a hashtag was coordinated, we carried out a preliminary inspection to check unusual behavior in terms of the number of tweets and retweets in a short period. Essentially, the explored hashtags showed an interesting level of astroturfing (King, Pan, and Roberts 2017); i.e. a significant number of tweets were masked as genuine opinions.

This odd behavior allowed us to find orchestrated attempts to artificially promote these hashtags on Twitter. Each sample was representative with respect to their “peak window” (Lin et al. 2013). Moreover, metadata obtained from these tweets contain information such as: user id, date, time zone, username, text, mentions, retweets count, hashtags, and user-defined location. However, most of this metadata is messy and requires a filtering process. First, we performed data cleansing aimed to detect, filter, and remove corrupt or inaccurate records. We removed obvious errors e.g. null fields, empty sets, and incomplete data, using a free and open-source tool. On the other hand, regarding the data collection, we developed a “social explorer” in Python programming language (Piña-García, Gershenson, and Siqueros-García 2016 & Python n.d.). All analyses were carried out using the R programming language, a free software environment for statistical computing and graphics (The R Foundation n.d.).

Given the nature of this study, it is worth briefly discussing the ethical and legal implications of using Twitter data to conduct this research. Tweets that were collected through the public Twitter API are subject to the Twitter terms and conditions (Twitter n.d.a), and to the developer agreement and policy as well (Twitter n.d.b). The privacy policy indicates that users consent to the collection, transfer, manipulation, storage, and disclosure of data that are public. Therefore, this research explored and gathered only those tweets that were public (i.e. no privacy settings were selected by the user) (Twitter n.d.c). It is important to note that with the aim to comply with Twitter terms of service, data cannot be

| Dataset                     | Date of collection | Number of collected Tweets | Language | Geolocated |
|-----------------------------|-------------------|----------------------------|----------|------------|
| #GatellOrgulloMexicano      | Jan/05/2021       | 3200                       | Spanish  | N/A        |
| #AMLOPresidenteDeLaSalud    | Feb/06/2021       | 1838                       | Spanish  | N/A        |
publicly shared. Interested researchers may repeat and reproduce the experiments by following the procedure described in this manuscript.

2.2. Early detection of astroturfing

Currently, astroturfing has become one of the most efficient strategies for many organizations to manipulate conversations. In addition, astroturfing aims to create an impression on Twitter of widespread grassroots support or attack for a policy or an individual. However, it is not easily detectable and may depend on the viewpoint of common users. This research aims to contribute to the development of a new knowledge framework to improve detection of this malicious use of automation to undermine and disrupt the public conversation, such as trying to get something to trend.

Early detection of astroturfing is essential for exposing the presence of efforts to manipulate online discourse. To detect this spammy tactic, a massive collection of tweets is not strictly necessary. Mainly, once a hashtag has achieved a top position, it is highly likely that public conversation has been polluted and polarized. A well-known example of natural coordinated behavior can be found during an earthquake: when people are exposed to the motion, they tend to share the experience on social media (see Figure 1(c)). This could be considered coordinated behavior, but in this case, coordination is determined by a natural phenomenon. During astroturfing time windows, patterns emerge over time causing a burst of tweets. According to Mendoza, Poblete, and Valderrama (2019), a burst is defined as a considerable number of occurrences of tweets within a short time window. We thus analyze how a burst of tweets is triggered in a short time window \( w \). In this regard, we define a threshold in terms of the number of tweets per minute as follows: \( 50 \text{tw} \leq \text{Thr} \leq 100 \text{tw} \), where these boundaries were empirically determined. However, different thresholds may be set to reflect a burst (Zhang and Shasha 2006).

Coordination in this context implies that the same set of accounts may signal coordinated behavior through synchronical “retweet,” “reply” or “like” of the same tweet. Figure 1 shows the burst detection or cybercascading within a certain time window for hashtags such as #GatellOrgulloMexicano, #AMLOPresidenteDeLaSalud, #GraciasLopezGatell, #SomosMillonesConAMLO, #INECorrupto, #AmloLiderMundial, #VictoriaDelPueblo and #VotaTodoMorena. Signals of astroturfing were found in these cases, so that in a short period a significant number of interactions was observed. Basically, this rapid growth may be considered the first sign of an atypical behavior of a hashtag. Once the burst or cybercascader is confirmed empirically, we carried out a preliminary inspection to detect the cumulative number of retweets. When a hashtag reveals suspicious coordinated behavior, we collected a data sample with the aim to catch the very first tweets and retweets of this trend. With this data sample, it is possible to accurately analyze the temporal behavior in terms of retweet frequency.

Figure 1 compares different temporal patterns of Twitter hashtags. From the chart it can be seen that most of the hashtags showed a rapid growth in the number of tweets. However, it is possible to observe that unlike Figure 1(b, c) the rest of the time series present the same shape. This pattern reveals that there has been a marked burst
of tweets. What is interesting in these hashtags is that all of them are associated with a political context between the year 2020 and 2021.

Figure 1(a) shows that there has been a marked increase in the hashtag #GatellOrgulloMexicano between 10:36am and 12:08pm Central Daylight Time (CDT). Regarding #AMLOPresidenteDeLaSalud a time series analysis was carried out to examine its temporal behavior. Figure 1(b) shows variability at first glance. However, a closer inspection revealed that this hashtag reached four peaks: (1) 16 February 2021 16:34pm, (2) 16 February 2021 21:07pm, (3) 17 February 2021, 00:30am and (4) 17 February 2021, 01:54am. The number of tweets in this case is likely to be related to several attempts to boost social engagement in Twitter. In this case, there was a significant effort to maintain the dominance of this hashtag.
3. Results and discussion

The present study was designed to expose the use of coordinated campaigns during coronavirus health crisis in Mexico; prior studies that have noted the importance of astroturfing on social media have been reported in the literature (Albadi, Kurdi, and Mishra 2019; Davis et al. 2016 and Geiger 2016).

Statistical analysis was used to determine the cumulative behavior. Figure 2 shows the number of tweets and retweets over time; the difference between the number of tweets and retweets was considerable, due to retweeting being the simplest way in which a coordinated campaign can be amplified. We consider that the majority presence of retweets over original tweets is a highly likely signal of astroturfing. In both hashtags, we note that the number of retweets rapidly increased, while the number of tweets was dropping. Shares and retweets can be thought as a canonical measure of amplification on Twitter.

Tweets with the highest engagement are temporarily summarized in Figures 3 and 4. These tweets are revealing in several ways. Firstly, Figure 3 shows those tweets that were promoting the use of the hashtag #GatellOrgulloMexicano; in most cases for supporting and in others for blaming. A closer inspection of these tweets revealed that their online activity was part of a coordinated campaign. On the other hand, in the hashtag #AMLOPresidenteDeLaSalud it is possible to examine those tweets that were related to political propaganda, suggesting the presence of efforts to manipulate online discourse.

Another significant aspect of coordinated campaigns based on astroturfing is the type of content included in tweets e.g. emojis and words. In this context, emojis are pictorial symbols that are typically presented in a colorful form and used inline in text. They correspond to characteristics such as attitude, mood, and sentiment (Unicode n.d.). Additionally, emojis have become popular for clarifying online communication and reducing ambiguity in the written discourse (Kaye, Malone, and Wall 2017).

Firstly, a content analysis was carried out with the aim to identify the most frequent emojis. In Figure 5(a), we can see that the face with tears of joy 😂 was the most used emoji followed by clapping hands 👏. Nevertheless, it could be suggested that the face with tears of joy is widely used to show irony.

![Figure 2. Cumulative behavior by Hashtag: #GatellOrgulloMexicano &#AMLOPresidenteDeLaSalud.](image-url)
Analogous to the previously discussed hashtag, Figure 5(b) shows that the face with tears of joy 😊 was the most used emoji. Moreover, the most interesting aspect of this figure is that we found supporting emojis such as two hands clapping 👏 and a decorative heart ❤️, but also some discontent emojis such as an angry red face indicating it is swearing or being vulgar 🗣 and a pile of poo 💩 that may be used to represent feces and other bathroom related topics.

Secondly, a content analysis was carried out with the aim to identify the most frequent words. we can observe on the word count in Figure 6, that there is some degree of

Figure 3. Tweets with the highest engagement #GatellOrgulloMexicano.
polarization at a text level, it is possible to find support words such as: “gracias” (thanks) and “vacunas” (vaccines); but similarly, it is possible to find disapproval words such as “#gatellasesino” (Gatell assassin) and “#lopezdestruyendomexico” (Lopez destroying Mexico).

The most interesting finding to emerge from the analysis is that astroturfing was aimed to support President Andrés Manuel López Obrador (AMLO) and Mexico’s coronavirus czar Hugo Lopez-Gatell. It should be highlighted that these two political figures have

Note: To protect the identity of the users, the tweets were anonymized.

Figure 4. Tweets with the highest engagement #AMLOPresidenteDeLaSalud.
been heavily criticized by the local and international press for downplaying pandemic seriousness (González 2021; Kitroeff 2020 and Malkin 2020). This situation has contributed to strategically amplifying pro-Government supporting content through coordinated and inauthentic activity.

Coordinated campaigns via astroturfing exert an influential effect upon public opinion through state-sponsored propaganda. It is important to bear in mind possible bias in this

![Figure 5. Emoji occurrences.](image1)

![Figure 6. Word frequency.](image2)
type of observational study. Consequently, although the number of collected tweets seems to be relatively small, it needs to be considered that astroturfing is triggered in a short time window.

These findings cannot be extrapolated to all hashtags; a note of caution is due here since systematic research into astroturfing is lacking. However, we argue that even though no ground truth on Mexican astroturfing is available, our early-stage approach may help to detect such artificially coordinated campaigns.

4. Conclusions

We have found that, during the coronavirus health crisis in Mexico, there was a highly likely presence of coordinated campaigns in its astroturfing version on Twitter. Preliminary results suggest that astroturfing was used to amplify hashtags such as #GateoOrgulloMexicano, and #AMLOPresidenteDeLaSalud. Although this study focused on these trends, the findings may well have a bearing on other coordinated hashtags that were triggered during 2020 (see Figure 1). This study contributes to existing knowledge of social media manipulation by providing insights about efforts on the strategic distraction of information in Mexico. In addition, it establishes an original approach for the early detection of astroturfing in Mexico. This paper contributes to the development of a new framework that improves the early detection of coordinated disinformation campaigns aimed to undermine and disrupt the public opinion by pretending to be ordinary citizens acting independently.

The outcomes reported here shed new light on the importance of preventing platform manipulation via fabricating social media comments as if they were the genuine opinions (grassroots). This approach should help build awareness and improve the public’s understanding of coordinated behavior. Notwithstanding the relatively limited sample, this research offers insights into how certain hashtags are promoted in a coordinated fashion with the aim to amplify political messages on Twitter.

Our research makes headway towards enhancing our understanding of how political propaganda is artificially promoted. Our exploration found that there are anomalous users (astro-puppets) that are likely used to distract public attention. Additionally, our content analysis revealed that most of the text contains some ‘copy paste’ content aimed to cheerlead or redirect public interest.

Further research should be undertaken to investigate the misuse of digital misinformation on Twitter. A natural progression of this study is to analyze reply tweets, with the aim to detect targeted harassment, abusive and hateful behavior. In addition, further studies could assess how perceptions are influenced and distorted by being exposed to echo chambers. Considerably, more research will be needed to determine the mid-term effects of intolerance towards groups who hold differing opinions in Twitter.

Finally, the current challenge is to uncover those accounts that repetitively use the reply, retweet, or like functions to spam and pollute hashtags. Greater efforts are still needed to reduce polarization and misinformation on the Twitter platform.

ORCID

C. A. Piña-García http://orcid.org/0000-0001-8741-4860
References

Abokhodair, N., D. Yoo, and D. W. McDonald. 2015. “Dissecting a Social Botnet.” Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing – CSCW ’15. 839–851. doi:10.1145/2675133.2675208

Albadi, N., M. Kurdi, and S. Mishra. 2019. “Hateful People or Hateful Bots?: Detection and Characterization of Bots Spreading Religious Hatred in Arabic Social Media.” Proceedings of the ACM on Human-Computer Interaction 3 (61): 1–25. doi:10.1145/3359163.

Aral, S. 2020. The Hype Machine: How Social Media Disrupts Our Elections, Our Economy, and Our Health – and How We Must Adapt. New York, NY: Currency.

Badawy, A., A. Addawood, K. Lerman, and E. Ferrara. 2019. “Characterizing the 2016 Russian IRA Influence Campaign.” Social Network Analysis and Mining 9: 1–11. doi:10.1007/s13278-019-0578-6.

Barnhill, A. 2014. “What Is Manipulation?” In Manipulation: Theory and Practice, edited by C. Coons and M. Weber, 51–72. E.U.A. New York, NY: Oxford University Press.

Bienkov, A. 2012, February 8. “Astroturfing: What Is It and Why Does It Matter?” The Guardian. https://www.theguardian.com/commentisfree/2012/feb/08/what-is-astroturfing

Brachten, F., S. Stieglitz, L. Hofeditz, K. Kloppenborg, and A. Reimann. 2017. Strategies and Influence of Social Bots in a 2017 German State Election – A Case Study on Twitter. https://arxiv.org/abs/1710.07562

Brittanica, T. 2015, October 7. “Editors of Encyclopaedia.” Persuasion. Encyclopedia Britannica. https://www.britannica.com/science/persuasion-psychology

Che, L., X. Liu, Z. Shuai, Z. Li, and Y. Wen. 2018. “Cybercascades Screening Considering the Impacts of False Data Injection Attacks.” IEEE Transactions on Power Systems 33 (6): 6545–6556. doi:10.1109/TPWRS.2018.2827060.

Cinelli, M., W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, and A. Scala. 2020. “The COVID-19 Social Media Infodemic.” Scientific Reports 10: 1–10. doi:10.1038/s41598-020-73510-5.

Davis, C. A., O. Varol, E. Ferrara, A. Flammini, and F. Menczer. 2016. “BotOrNot: A System to Evaluate Social Bots.” Proceedings of the 25th International Conference Companion on World Wide Web, (WWW ’16 Companion). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE. 273–274. doi:10.1145/2872518.2889302.

DeVito, J. A. 1986. The Communication Handbook: A Dictionary. New York, NY: Harper & Row.

Dolan, E. 2020, April 14. “Snapshot: Changing Behaviors in Latin America.” Blog Global Web Index. https://blog.gwi.com/trends/snapshot-coronavirus-attitudes-in-latam/

Dolan, S. 2021, July 13. “Mexico’s TikTok User Base More than Tripled in 2020.” eMarketer. https://www.emarketer.com/content/mexico-tiktok-user-base-more-than-tripled-2020

Elmas, T., R. Overdorf, A. F. Özkalay, and K. Aberer. 2019. Ephemerall Astroturfing Attacks: The Case of Fake Twitter Trends. https://arxiv.org/abs/1910.07783.

Enli, G. S., and E. Skogerbo. 2013. “Personalized Campaigns in Party-Centred Politics.” Information, Communication & Society 16 (5): 757–774.

Forelle, M., P. Howard, A. Monroy-Hernández, and S. Savage. 2015. Political Bots and the Manipulation of Public Opinion in Venezuela. https://arxiv.org/abs/1507.07109.

Fregoso, J. 2018. #Mexico2018 | “Fake News” and Social Media: The New Heads of the Hydra. Journalist Fellows Papers. Reuters Institute for the Research of Journalism. University of Oxford. https://reutersinstitute.politics.ox.ac.uk/our-research/mexico2018fake-news-and-social-media-new-heads-hydra

Geiger, R. S. 2016. “Bot-Based Collective Blocklists in Twitter: The Counterpublic Moderation of Harassment in a Networked Public Space.” Information, Communication & Society 19 (6): 787–803. doi:10.1080/1369118X.2016.1153700.
González, M. 2021, March 2. “Coronavirus: ‘En México no es que no hayan sabido qué hacer con la pandemia, es que tomaron la decisión de no hacerlo.’” BBC News Mundo Mundo México. https://www.bbc.com/mundo/noticias-america-latina-56220420

Grinberg, N., K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer. 2019. “Fake News on Twitter During the 2016 U.S. Presidential Election.” Science 363 (6425): 374–378. doi:10.1126/science.aau2706.

Gutiérrez, H. 2016, November 18. “Multa TEPJF con 7 mdp a PVEM por tuits.” Reforma. https://www.reforma.com/aplicacioneslibre/articuloamp/default.aspx?id = 987850

Herik, J., and D. Dimov. 2012. “Towards Crowdsourced Online Dispute Resolution.” Journal of International Commercial Law and Technology 7 (2): 99–111. https://citeseerx.ist.psu.edu/viewdoc/download?doi = 10.1.1.828.1681&rep = rep1&type = pdf

Jowett, G. S., and V. O’Donnell. 2019. Propaganda & Persuasion. New York, NY: SAGE Publishing.

Kaye, L. K., S. A. Malone, and H. J. Wall. 2017. “Emojis: Insights, Affordances, and Possibilities for Psychological Science.” Trends in Cognitive Sciences 21 (2): 66–68. doi:10.1016/j.tics.2016.10.007.

Keller, F. B., D. Schoch, S. Stier, and J. Yang. 2017. “How to Manipulate Social Media: Analyzing Political Astroturfing Using Ground Truth Data from South Korea.” Proceedings of the Eleventh International AAAI Conference on Web and Social Media. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/viewPaper/15638

Keller, T. R., and U. Klinger. 2019. “Social Bots in Election Campaigns: Theoretical, Empirical, and Methodological Implications.” Political Communication 36 (1): 171–189. doi:10.1080/10584609.2018.1526238.

Kitty, D., A. Godoy, and F. V. Zuben. 2016. “Online Social Network Analysis: A Survey of Research Applications in Computer Science.” ArXiv, 1–45. http://arxiv.org/abs/1504.05655

Lazer, D. M. J., M. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. Metzger, et al. 2018. “The Science of Fake News.” Science 359 (6380): 1094–1096. doi:10.1126/science.aao2998.

Lin, Y.-R., D. Margolin, B. Keegan, A. Baronchelli, and D. Lazer. 2013. “#Bigbirds Never Die: Understanding Social Dynamics of Emergent Hashtags.” Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013. 370-379. https://arxiv.org/abs/1303.7144

Lits, B. 2020. “Detecting Astroturf Lobbying Movements.” Communication and the Public 5 (3–4): 164–177. doi:10.1177/2057047320969435.

Lits, B. 2021. “Exploring Astroturf Lobbying in the EU: The Case of Responsible Energy Citizen Coalition.” European Policy Analysis 7 (1): 226–239. doi:10.1007/epa2.1086.

López, J. C. A. D., and C. A. Piña-García. 2017. “Political Participation in Mexico Through Twitter.” In Complex Networks & Their Applications V. COMPLEX NETWORKS 2016 2016. Studies in Computational Intelligence, edited by H. Cherifi, S. Gaito, W. Quattrociocchi, and A. Sala. 693. Cham: Springer. doi:10.1007/978-3-319-50901-3_48.

Malkin, E. 2020. “Covid Fatalities Soar in Mexico as President Condemned for Inaction.” The Guardian, December 23. https://www.theguardian.com/world/2020/dec/23/mexico-coronavirus-amlo-cases-deaths

Mendoza, M., B. Poblete, and I. Valderrama. 2019. “Nowcasting Earthquake Damages with Twitter.” EPJ Data Science 8 (3), doi:10.1140/epjds/s13688-019-0181-0.

Miller, Z., B. Dickinson, W. Deitrick, W. Hu, and A. H. Wang. 2014. “Twitter Spammer Detection Using Data Stream Clustering.” Information Sciences 260: 64–73. doi:10.1016/j.ins.2013.11.016.
Navarro, M. F. 2019, January 1. “Selección 2018, Ellos son los creativos que dieron un vuelco a la campaña de AMLO.” Forbes. https://www.forbes.com.mx/ellos-son-los-creativos-que-dieron-un-vuelco-a-la-campana-de-amlo/

Nolasco, S. 2020, April 29. “Uso de cubrebocas no es la medida principal: López-Gatell.” El Economista. https://www.eleconomista.com.mx/arteseideas/Uso-de-cubrebocas-contra-Covid-19-no-es-la-medida-principal-Lopez-Gatell-20200429-0135.html

Pacheco, D., P.-M. Hui, C. Torres-Lugo, B. T. Truong, A. Flammini, and F. Menczer. 2020. “Uncovering Coordinated Networks on Social Media.” http://arxiv.org/abs/2001.05658.

Piña-García, C. A., C. Gershenson, and J. M. Siqueiros-García. 2016. “Towards a Standard Sampling Methodology on Online Social Networks: Collecting Global Trends on Twitter.” Applied Network Science 1 (3). https://arxiv.org/abs/1507.01489

Piña-García, C. A., and L. Ramírez-Ramírez. 2019. “Exploring Crime Patterns in Mexico City.” Journal of Big Data 6: 1–21. doi:10.1186/s40537-019-0228-x.

Python. n.d. Python 3.6.0. Accessed 27 April 2020. https://www.python.org/downloads/release/python-360/.

The R Foundation. n.d. R: The R Project for Statistical Computing. Accessed 14 April 2020. https://www.r-project.org/.

Ratkiewicz, J., M. Conover, M. Meiss, B. Gonçalves, S. Patil, A. Flammini, and F. Menczer. 2011. “Truthy: Mapping the Spread of Astroturf in Microblog Streams.” Proceedings of the 20th International Conference Companion on World wide web (WWW ‘11). Association for Computing Machinery. 249–252. doi:10.1145/1963192.1963301.

Savage, S., A. Monroy-Hernández, and T. Höllerer. 2016. “Botivist: Calling Volunteers to Action Using Online Bots.” Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW ’16). Association for Computing Machinery. 813–822. doi:10.1145/2818048.2819985.

Sánchez-Talanquer, M., E. González-Pier, J. Sepúlveda, L. Abascal-Miguel, J. Fieldhouse, C. del Río, and S. Gallalee. 2021. “La respuesta de México al COVID-19: Estudio de caso. Institute for Global Health Science.” https://globalhealthsciences.ucsf.edu/sites/globalhealthsciences.ucsf.edu/files/la_resposta_de_mexico_al_covid_esp.pdf

Shu, K., S. Wang, D. Lee, and H. Liu. 2020. “Mining Disinformation and Fake News: Concepts, Methods, and Recent Advancements.” ArXiv, abs/2001.00623. https://arxiv.org/abs/2001.00623.

Stella, M., E. Ferrara, and M. De Domenico. 2018. “Bots Increase Exposure to Negative and Inflammatory Content in Online Social Systems.” Proceedings of the National Academy of Sciences of the United States of America 115 (49): 12435–12440. doi:10.1073/pnas.1803470115.

Story Wrangler. n.d.a. Usage of #GatellOrgulloMexicano. Accessed 11 October 2021. https://storywrangling.org/ngrams?ngrams=#GatellOrgulloMexicano&language=es&rt=true&scale=lin&metric=rank&start=2009-04-16&end=2022-06-25

Story Wrangler. n.d.b. Usage of #AMLOPresidenteDeLaSalud. Accessed 11 October 2021. https://storywrangling.org/ngrams?ngrams=#AMLOPresidenteDeLaSalud&language=es&rt=true&scale=lin&metric=rank&start=2009-04-16&end=2022-06-25

Taylor, D. B. 2021, March 21. “A Timeline of the Coronavirus Pandemic.” The New York Times. https://www.nytimes.com/article/coronavirus-timeline.html

Tello, I. 2021, August 20. “Astroturfing y los ataques a Brenda Lozano.” El Universal. https://www.eluniversal.com.mx/opinion/Irene-Tello-Arista/astroturfing-y-los-ataques-brenda-lozano

Twitter. n.d.a. Terms of Service. https://twitter.com/en/tos

Twitter. n.d.b. Developer Agreement and Policy – Twitter Developers. Accessed 13 April 2020. https://developer.twitter.com/en/developer-terms/agreement-and-policy

Twitter. n.d.c. Privacy Policy. Accessed 13 April 2020. https://twitter.com/en/privacy

Unicode. n.d. About Emoji. https://home.unicode.org/emoji/about-emoji/

Varol, O., E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini. 2017. Online Human-Bot Interactions: Detection, Estimation and Characterization. https://arxiv.org/abs/1703.03107

Winterer, S. 2021. “What is a Social Media Campaign?” Digital Logic, March 8. https://www.digitallogic.co/blog/what-is-a-social-media-campaign/
Wood, A. W. 2014. “Coercion, Manipulation, Exploitation.” In *Manipulation: Theory and Practice*, edited by C. Coons and M. Weber, 17–49. E.U.A. New York, NY: Oxford University Press.

World Health Organization [WHO]. 2021, June 14 Considerations for Implementing and Adjusting Public Health and Social Measures in the Context of COVID-19. Interim Guidance. https://www.who.int/publications/i/item/considerations-in-adjusting-public-health-and-social-measures-in-the-context-of-covid-19-interim-guidance

Wu, L., F. Morstatter, K. M. Carley, and H. Liu. 2019. “Misinformation in Social Media: Definition, Manipulation, and Detection.” *ACM SIGKDD Explorations Newsletter* 21 (2): 80–90.

Zerback, T., F. Töpfl, and M. Knöpfe. 2021. “The Disconcerting Potential of Online Disinformation: Persuasive Effects of Astroturfing Comments and Three Strategies for Inoculation Against Them.” *New Media & Society* 23 (5): 1080–1098. doi:10.1177/1461444820908530.

Zhang, X., and D. Shasha. 2006. “Better Burst Detection.” 22nd International Conference on Data Engineering (ICDE’06). doi:10.1109/ICDE.2006.30.