EPJ Data Science
a SpringerOpen Journal

Multifaceted online coordinated behavior in the 2020 US presidential election

Serena Tardelli[1*] , Leonardo Nizzoli[1], Marco Avvenuti[2], Stefano Cresci[1] and Maurizio Tesconi[1]

*Correspondence:
serena.tardelli@iit.cnr.it
[1]Institute of Informatics and
Telematics, National Research
Council, Pisa, Italy
Full list of author information is
available at the end of the article

Abstract

Organized attempts to manipulate public opinion during election run-ups have dominated online debates in the last few years. Such attempts require numerous accounts to *act in coordination* to exert influence. Yet, the ways in which coordinated behavior surfaces during major online political debates is still largely unclear. This study sheds light on coordinated behaviors that took place on Twitter (now X) during the 2020 US Presidential Election. Utilizing state-of-the-art network science methods, we detect and characterize the coordinated communities that participated in the debate. Our approach goes beyond previous analyses by proposing a multifaceted characterization of the coordinated communities that allows obtaining nuanced results. In particular, we uncover three main categories of coordinated users: (*i*) moderate groups genuinely interested in the electoral debate, (*ii*) conspiratorial groups that spread false information and divisive narratives, and (*iii*) foreign influence networks that either sought to tamper with the debate or that exploited it to publicize their own agendas. We also reveal a large use of automation by far-right foreign influence and conspiratorial communities. Conversely, left-leaning supporters were overall less coordinated and engaged primarily in harmless, factual communication. Our results also showed that Twitter was effective at thwarting the activity of some coordinated groups, while it failed on some other equally suspicious ones. Overall, this study advances the understanding of online human interactions and contributes new knowledge to mitigate cyber social threats.

Keywords: Coordinated behavior; Disinformation; Social network

1 Introduction

In an era characterized by the widespread use of social media platforms, concerns have emerged regarding their role in shaping political discourse, influencing public opinion, and impacting democratic processes [1–3]. The 2016 United States (US) presidential election campaign, in particular, brought to light the potential for social media to be manipulated for various purposes, including spreading disinformation and polarization. This concern was exemplified by organized efforts, such as those orchestrated by the Internet Research Agency (IRA) on behalf of the Russian government, which aimed to disrupt online conversations. Simultaneously, the prevalence of automated social bots, responsible for a substantial portion of election-related tweets, raised questions about their influence

in amplifying content, including low-credibility information [3, 4]. Notably, beyond organized campaigns by specific actors, everyday social media users also contributed to the dissemination of divisive narratives, as observed in the realm of conspiracy theories [5].

Such deliberate attempts to shape the information landscape and influence public perception fall under the category of Information Operations (IOs). Previous research focused specific aspects of IOs, such as content features or account attributes [6, 7]. A paramount characteristic of these large-scale operations is the need to coordinate the activities of multiple accounts, as a tactic to effectively influence online discussions [8–11]. Consequently, the systematic and quantitative study of *coordinated online behavior* represents a powerful tool to unveil potential IOs, irrespective of their nature or objectives [12]. Coordinated behavior is defined as an unexpected or exceptional similarity, that is prolonged in time, between the activity of two or more users [9, 13, 14]. This definition of coordinated behavior implies that the concept extends beyond orchestrated campaigns, encompassing also other forms of coordination such as grassroots movements where genuine actors (e.g., political activists) collaborate to amplify their messages [15]. This study delves into the multifaceted landscape of coordinated online behavior, providing new insights into its various dimensions and their implications for online political discourse and democratic processes.

The multifaceted nature of online coordination can serve multiple purposes, which demands careful characterization of each detected instance of coordinated behavior [16]. Such nuanced understanding is instrumental towards discerning the true nature of the coordinated activities: malicious or otherwise. However, in spite of the multiple dimensions that allow to characterize online coordination, the existing literature has predominantly focused on a single one: inauthenticity. This choice stemmed from Facebook's initial focus on coordinated inauthentic behavior (CIB) [17], which influenced subsequent research in the area. However, many other orthogonal dimensions of coordinated behavior are also worth exploring [18]. Here, we investigate various alternative facets, extending the existing literature by broadening the spectrum of dimensions considered during the characterization of coordinated behavior.

1.1 Contributions

We present a systematic and comprehensive analysis of coordinated behaviors carried out on Twitter (now rebranded as X) during the US 2020 electoral debate. To reach our goal, we collect and analyze a large dataset of 260M tweets shared by 15M distinct users. Our study complements the large existing body of work on the US 2020 election – which includes studies on the spread of fake news [19–21] and conspiracy theories [22, 23], and on the activity of social bots [24, 25] – by investigating the important, yet so far overlooked, dimension of online coordination.

In detail, we first *detect*, and subsequently *characterize*, the main coordinated communities that took part in the Twitter electoral debate. In the detection step we apply a state-of-the-art methodology for coordination detection [9]. Then, in the characterization step we significantly advance state-of-the-art approaches – which mainly focus on (in)authenticity [24, 26] – by investigating coordinated behavior from different angles, including: the degree of coordination of the different communities, their use of automation, their political partisanship, the degree of factuality of the content they shared, and platform suspensions of coordinated accounts. We demonstrate that simultaneously examining all these aspects

provides rich insights into the intent, activity, and harmfulness of the different types of co-ordinated communities.

Our results reveal that not all coordinated communities are equally malicious or harmful. In addition, our nuanced characterization reveals that Twitter managed to detect and remove only some malicious coordinated communities, while others remained active on the platform. As such, our results are not only useful for better understanding the US 2020 online electoral debate, but also for identifying the areas and directions where online platforms might need to intervene more effectively in the future. Our results are thus particularly relevant as we progress through 2024 – a year with a multitude of elections where nearly half of the world population will be impacted by an electoral process.[1]

Our main contributions can be summarized as follows:

- We collect and publicly share a vast dataset of tweets about the 2020 USA Presidential Election.[2] For the sake of reproducibility we also provide the code used to detect coordinated communities.[3]
- We analyze such a dataset, uncovering instances of coordinated behavior previously unexplored in literature. Our results highlight networks of *foreign influence* that exhibit signs of activism and social movements (e.g., Hong Kong and Nigeria protests), along with conspiracy theorists in the online debate. Moreover, we identify right-wing conspiracy communities as particularly detrimental to the discourse.
- We apply a state-of-the-art technique to detect coordinated behaviors [9], also extending it by introducing a novel indicator of coordination that simultaneously considers both the level of coordination and the size of the community. Our results demonstrate the usefulness of the indicator towards identifying harmful communities.
- We characterize each coordinated community along multiple dimensions, going beyond existing works that are mainly focused on inauthenticity. Our nuanced characterization contributes to identifying a wide spectrum of diverse coordinated communities, spanning from modest grassroots activists to riots, and culminating in highly detrimental instances of inauthentic behavior.

1.2 Significance

This research offers valuable insights into the dynamics of coordination, inauthenticity, and harmfulness within the US 2020 landscape of online behaviors. It highlights the tendency of conservative users to engage with like-minded individuals and to share non-factual, hyper-partisan content. These distortions in the online discourse may, in part, be attributed to both domestic and foreign malicious information operations facilitated through coordinated behavior.

The significance of this study extends beyond these findings, underscoring the need to understand and mitigate these distortions to maintain the integrity of the public discourse and democratic processes. By bringing attention to the role of foreign influence alongside domestic factors, this research contributes to safeguarding the trustworthiness and reliability of information ecosystems. It provides valuable insights into the complex dynamics

---

[1] Ray, S. (2024, January 3). "2024 Is The Biggest Election Year In History – Here Are The Countries Going To The Polls This Year". *The Economist*. https://www.economist.com/interactive/the-world-ahead/2023/11/13/2024-is-the-biggest-election-year-in-history.

[2] https://doi.org/10.5281/zenodo.7358386.

[3] github.com/srntrd/framework-to-identify-online-coordinated-behavior.

of online behavior and its impact on political and social landscapes, emphasizing the importance of a comprehensive approach to addressing these challenges.

1.3 Roadmap

The remainder of this paper is organized as follows. Section 2 examines existing research, focusing on two main areas: the analysis of the 2020 US Election Twitter debates and the identification of coordinated behaviors. Section 3 details the dataset related to the 2020 USA Presidential Election online debate. Section 4 details the our methodology used to identify and characterize coordinated communities, which introduces enhancements to the characterization process and examines specific facets to offer a nuanced understanding of coordinated behavior. Section 5 presents our findings, showing the narratives of the identified communities and offering an individual and comprehensive analysis of the various facets to understand the depth and breadth of coordinated activities. Section 6 discusses our findings and methodological limitations. Lastly, Sect. 7 wraps up our study, highlighting potential directions for future research.

2 Related work

We first review prior relevant work on the Twitter online debate about the 2020 US election. Next, we discuss state-of-the-art approaches for detecting and characterizing coordinated behavior, highlighting the advancements that this study offers.

2.1 The 2020 US election Twitter debate

To the best of our knowledge, the only work that analyzed coordinated behavior in the context of the US 2020 election is [27]. They focused on the debate related to the spread fraud allegations and revealed structured communities with strong evidence of coordination to promote (i.e., retweet) such claims. Then, they characterized the communities in terms of suspensions and content and find that bans imposed by the platform did not have effectively stopped the spread of minformation. Our study broadens the focus to the entire debate, uncovering coordinated communities and characterizing a wider range of aspects beyond just topics and suspensions. Other studies analyzed online coordination across different major events [28] 2020 US Primaries, yet coordination patterns during the 2020 US election remain largely unexplored.

   The majority of studies examined the 2020 US election focusing on automation. For instance, the authors in [29] uncovered discussion groups and analyzed their topics and levels of automation to understand how they contributed to the spread of fraud claims. The authors in [24] provided insights into automated accounts and narrative distortions within the debate. Similarly, [26] investigated the community of Trump supporters, presenting a characterization centered on the inauthenticity aspect. Instead, [30] analyzed fraud claims by spotlighting patterns that led to user suspensions, while the authors of [22] analyzed targeted disinformation topics and engagements with QAnon conspiracies, focusing on political leanings to understand manipulation attempts. Other studies explored the influence of fake news [1], misinformation [31], and conspiracy theories [32] within the debate. However, existing works primarily focused on a single deception aspect, especially (in)authenticity. In this work we complement the existing analyses on the US 2020 election, by investigating coordinated communities that took part in the online debate from multiple perspectives, including: the degree of coordination of the different communities, their

use of automation, their political partisanship, the degree of factuality of the content they shared, and platform suspensions of coordinated accounts. We demonstrate that simultaneously examining all these aspects provides rich insights into the intent, activity, and harmfulness of the different types of coordinated communities.

2.2 Detection and characterization of coordinated behavior

For the detection of coordinated behavior we adopt the methodology that we proposed in [9]. Our approach revolves around building a Term Frequency-Inverse Document (TF-IDF) co-retweet network filtered using a backbone strategy [33], on which we perform community detection at different coordination levels, so as to detect coordinated communities. The peculiarity of our methodology with respect to others in literature [10, 12] is that it provides an estimation of the extent of coordination between users, rather than a binary distinction between coordinated and non-coordinated ones. However, our method has the drawback of demanding manual and subjective assessments of the extent of coordination of each detected community [9]. In addition, it focuses on the detection task and deals only marginally with the characterization of the detected communities. Our present work overcomes both these limitations.

Besides our previous work, the authors of [10, 12] adopted similar approaches by constructing user networks based on common behavioral traces and by performing community detection. However, their methods involve imposing arbitrary thresholds to remove low-weight edges, which discards significant portions of the network as well as many potentially interesting traces of coordination. By applying fixed thresholds, these and other approaches [34, 35] cast the coordination detection task as a binary problem, classifying users and communities as either coordinated or non-coordinated [8]. However, coordination within a community expresses *the degree* to which the actions and behaviors of the community members contribute towards a common goal [5, 9]. Our research aims to capture this nuanced complexity. As such, we move beyond the binary classification that is predominant in many state-of-the-art methods, allowing a deeper understanding of the multifaceted nature of online coordination [5, 16].

Recent studies focused on developing multilayer network frameworks to detect coordination. For instance, the authors in [14] leveraged three standard action types (i.e., co-hashtag, co-URL, and co-mention) and constructed a three-view network based on these interactions. They considered highly synchronized actions occurring within a 5-minute time frame and identified the clusters with the highest density, which then visually inspected. Similar works explored cross-platform coordinated behavior through multilayer networks, where each layer represented a different social media platforms [36, 37]. Other studies [11, 38] proposed multilayer networks with a temporal dimension, with each layer representing a time frame, allowing to analyze the evolution of coordinated communities. In summary, a growing number of studies are considering the wider complexities involved in coordinated behavior within their analyses. This suggests a trend towards a more comprehensive approach in understanding these dynamics. Our study aligns with this trend, thoroughly evaluating numerous aspects to characterize coordinated communities.
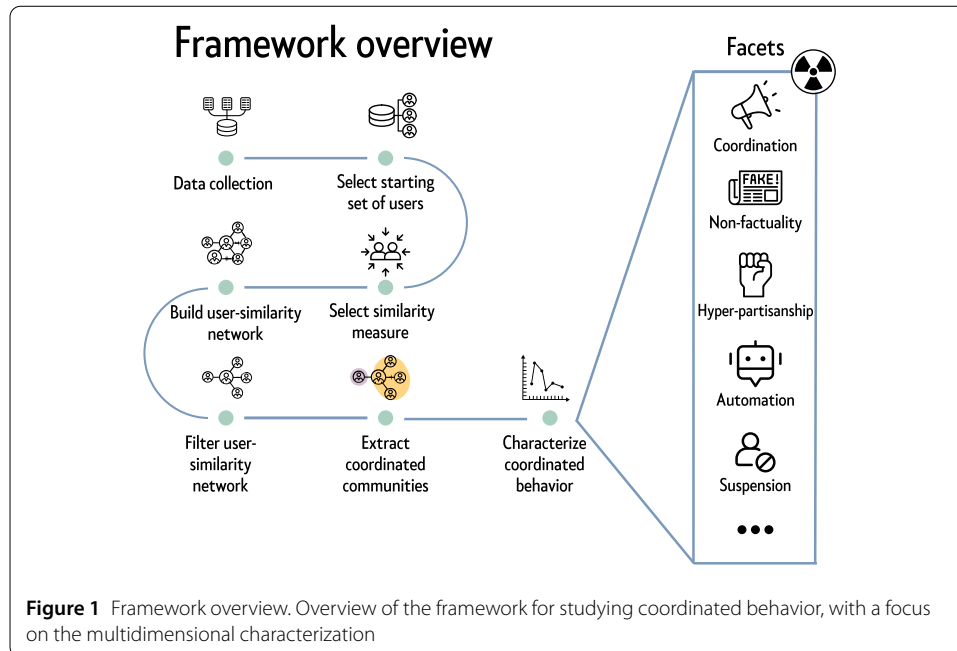
3 Dataset

We collected a large dataset of tweets related to the 2020 US Presidential Election using the Twitter Streaming Application Programming Interfaces (APIs). The data collection

Table 1 Statistics about data collected via keywords. Dataset statistics are related to the dataset after the filter applied to retain the activity of superspreaders – the most influential spreaders of information. The total number of tweets and superspreaders refers to the dataset, while, since a single tweet can contain one or more keywords, the sum of the single aggregate values does not reflect the real dataset size

Hashtag	Lean	Users	Tweets
#usa2020elections	N	3	4
#usa2020	N	988	4263
#Election2020	N	86,305	2,560,391
#ElectionDay	N	78,100	584,095
#Debates2020	N	80,267	703,307
#Vote	N	85,642	2,694,442
#VoteEarly	N	52,497	349,301
#Ivoted	N	24,189	45,957
#JoeBiden	N	81,193	743,605
#biden	N	84,480	1,106,452
#DonaldTrump	N	41,513	118,199
#trump	N	86,245	1,976,369
#VoteBlue	D	27,007	153,856
#VoteBlueToSaveAmerica	D	23,464	157,107
#Biden2020	D	48,163	141,351
#BidenHarris2020	D	71,302	979,039
#joebiden2020	D	13,797	19,789
#NeverTrump	D	21,534	32,740
#WakeUpAmerica	D	34,900	122,378
#VoteRedToSaveAmerica	R	43,324	380,114
#VoteRed	R	40,748	245,643
#trump2020	R	73,901	3,032,268
#trumppence2020	R	58,679	512,048
#donaldtrump2020	R	2891	3808
#MAGA	R	85,216	4,863,113
#KAG	R	65,332	859,569
#NeverBiden	R	21,293	44,045

Mentions	Lean	Users	Tweets
@JoeBiden	D	87,101	12,499,125
@DrBiden	D	57,442	288,255
@KamalaHarris	D	86,100	2,971,072
@SenKamalaHarris	D	61,019	260,468
@TheDemocrats	D	67,105	324,688
@POTUS	R	86,482	3,670,176
@realDonaldTrump	R	87,115	43,131,119
@Mike_Pence	R	81,164	1,022,082
@VP	R	82,898	853,809
@MELANIATRUMP	R	24,366	45,442
@FLOTUS	R	83,908	1,260,890
@GOP	R	85,669	1,727,397
total	–	87,410	71,506,397

period covered one month before and one month after Election Day, from October 2 to December 2, 2020. During this period, we collected all the tweets mentioning at least one hashtag from a set of predefined neutral and politically polarized hashtags related to the debate. Additionally, we collected all tweets mentioning official party accounts or other relevant official political accounts using predefined mentions as keywords. Our dataset includes a total of 263,518,037 tweets, posted by 15,288,527 distinct users. We focus our analysis on retweets only, corresponding to 53% (140,911,519) of the total tweets, posted by 57% (8,731,107) of the total users. In particular, as we explain in Sect. 4, we target the most influential information spreader users – superspreaders – which we define as the top

Figure 1 Framework overview. Overview of the framework for studying coordinated behavior, with a focus on the multidimensional characterization

1% of users that shared more retweets. This definition resulted in the selection of 87,410 users, responsible for a total of 71,506,397 retweets. This allows us to analyze the 27% of the total tweets, and the 51% of retweets in the dataset as a good compromise between computational effort and dataset size. Table 1 provides statistics on the data collected per hashtag and mention related to the activity of the superspreaders. The table lists all the hashtags and mentions used during this phase, their corresponding political leaning (N: Neutral, D: Democrat, R: Republican), and the related data collected. We publicly share our dataset for research purposes and to ensure reproducibility.[4]

4 Methods

In the first subsection we describe the state-of-the-art methodology employed for detecting coordinated communities. In the second subsection we outline our approach to the characterization of each detected community. This involves a comprehensive exploration of various facets that we integrate and analyze simultaneously. Our multi-dimensional analysis offers a rich understanding of the coordinated behaviors in the context of the 2020 US Presidential Election discourse. The overall workflow of our methodology is shown in Fig. 1.

4.1 Detection

As previously stated, we detect the coordinated communities that took part in the online debate by applying the state-of-the-art framework proposed by Nizzoli et al. in [9]. The framework integrates a variety of advanced techniques and unfolds through a series of steps. For completeness, in the following we briefly describe each step, by also specifying all choices where our approach diverges with respect to the original method.

---

[4]https://github.com/srntrd/framework-to-identify-online-coordinated-behavior.

4.1.1 Selecting the starting set of users

The first step of the method by Nizzoli et al. involves selecting an initial set of users. This is useful for filtering out irrelevant users, so as to streamline and improve the computational feasibility of the subsequent analyses [9]. In our study of the US 2020 election data, we focus on highly engaged users who are the primary contributors to the online discussion [39]. In particular, we target the most influential disseminators of information, whom we refer to as *superspreaders*. Superspreaders are identified as the top 1% of users with the most retweets. Notably, they represent a significant portion of our dataset, accounting for 27% of all tweets and 51% of all retweets. By adopting this approach, we are able to closely examine the activities of the most active users and analyze a substantial part of the overall communication within the network. Notably, the volume of retweets has been recognized a pertinent metric for identifying influential users, a strategy adopted by other studies as well [29, 40].

4.1.2 Selecting the similarity measure

Next, Nizzoli et al.'s framework identifies coordination by detecting repeated similarities in user behaviors. Similarities between users can be observed across many behavioral dimensions. Among them, retweeting behavior similarity is by far the most widely used [12, 29, 41]. Similarly to previous work, we measure user similarities based on their retweeting activity. In detail, each user is defined by the TF-IDF weighted vector of the tweets they retweeted. Pairwise similarities between all users are computed as the cosine similarity of their user vectors. The TF-IDF weighting scheme diminishes the influence of widely popular (i.e., viral) tweets and amplifies the contribution of retweets targeting less popular tweets. This approach is particularly effective at identifying suspicious behaviors and coordination that manifest through the retweeting of not-so-popular posts [9]. At the same time, the methodology accounts for the fact that sporadic retweets of the same content do not imply coordination. On the contrary, this method contributes to uncovering deeper, non-obvious patterns of behavior that might also be indicative of genuine coordination, rather than mere homophily or the virality of popular tweets.

4.1.3 Building the user similarity network

Pairwise similarities are converted into a weighted undirected user similarity network, which encodes behavioral and interaction patterns between users. Edge weights in this network are proxies for the degree of coordination [9, 12, 29].

4.1.4 Filtering the user similarity network

Being based on pairwise similarities, user similarity networks are typically too big to be analyzed straight away. Therefore, all network-based coordination detection frameworks include a filtering step to prune irrelevant edges and disconnected nodes from the network [9, 12]. Nizzoli et al. propose extracting the multiscale backbone [33] of the user similarity network as the filtering technique. This approach allows retaining all statistically significant network structures, independently of their scale. As a result of the multiscale backbone extraction, our filtered user similarity network consists of 9911 nodes and 148,038 edges.

4.1.5 Detecting coordinated communities

The advantage of network-based representations of coordination lies in the possibility to easily detect groups of users that are strongly coordinated with one another. All network-based coordination detection frameworks identify coordinated groups of users via community detection [9, 10, 12]. Although many community detection algorithms can be used for this task, the vast majority of existing works leverage the Louvain algorithm [42]. In particular, Nizzoli et al. proposed to perform Louvain community detection within an iterative network dismantling process [9]. This approach allows to analyze the structure and behavior of each coordinated community at all degrees of coordination. As such, it provides more nuanced results with respect to the methods that apply fixed and arbitrary edge weight thresholds, which allow analyzing only a single snapshot of the user similarity network. Here, we apply Nizzoli et al.'s approach to the US 2020 user similarity network, detecting the main coordinated communities that took part in the electoral debate. For practical reasons related to computational efficiency and clarity of visualizations, we provide detailed results for the top 10 largest coordinated communities, whose members account for 79% of all retweets.

4.1.6 Characterizing coordinated communities

In the final step of the analysis, each detected coordinated community is investigated so as to make sense of its intent and actions. The majority of existing literature performed the characterization task via qualitative manual analyses [9, 27]. Here instead, we propose and compute a number of quantitative indicators based on established and state-of-the-art methodologies. Moreover, previous works mainly focused on analyzing either the topics discussed by each community or their network structure and topology [12, 27]. By computing many relevant indicators, here we go beyond existing works by encompassing multiple facets and perspectives. Overall, our new approach to the characterization of coordinated communities offers a comprehensive and nuanced depiction of online coordination. In the following subsection we describe the facets according to which we characterize each coordinated community, the methodologies behind their computation at both user and community level, and their importance in the broader context of our analysis.

4.2 Characterization

We characterize each coordinated community according to five important facets of coordinated online behavior: (*i*) the degree of coordination, (*ii*) the degree to which members of the community make use of automation, (*iii*) account suspensions by the platform, (*iv*) the extent of political partisanship of the community, and (*v*) the degree to which the community shares non-factual news. Each facet offers insights into the intent, behavior, and structure of the coordinated communities, allowing to make sense of their activity and involvement in the US 2020 online debate. Then, in addition to computing and discussing these indicators individually, we also compute a composite measure for these facets, which enables comparative analyses across varied behaviors and forms of coordination. Thus, a key contribution of our work lies in analyzing these facets both separately and in conjunction, providing a holistic view that is instrumental towards evaluating the nature (e.g., harmful and/or inauthentic) and potential impact of coordinated communities.

4.2.1 Coordination

Most state-of-the-art methods cast the coordination detection task as a binary problem, according to which users and communities are either coordinated or non-coordinated [12]. However, coordination within a community expresses *the degree* to which the actions and behaviors of the community members contribute towards a common goal. As such, multiple degrees of coordination exist, which demand nuanced modeling and analyses that go beyond the simplistic binary definition used so far. Differently from the majority of previous work, our approach for detecting coordination assigns a coordination score on a continuous scale $\in [0, 1]$ to each user in a community. In [9], this coordination score is used to build a coordination curve for each community that quantifies the fraction of community members at any given coordination score. It is then left to the analyst to manually interpret the coordination curves and the behavior of the detected coordinated communities [9].

Here we take a step forward by computing an aggregated coordination score for each community, as the area under the coordination curve. Similarly to the user-level coordination score, also the community coordination score $\in [0, 1]$ and quantifies the degree to which the members of the community are coordinated between one another. The advantage of this approach lies in the possibility to have an objective and quantitative indicator of the degree of coordination within a community, rather than a subjective assessment derived from manual inspection. In addition, the aggregated coordination score lends itself to useful comparisons with other indicators that can be computed for the same communities, as described in the following.

4.2.2 Automation

Automation of a number of social media accounts is one of the ways in which online coordination can be achieved. For example, social botnets exploit the synchronization allowed by the use of automation to effectively push narratives or support certain public characters [43]. As such, social botnets that tamper with the regular flow of information are a paramount example of inauthentic and harmful coordination [10, 35, 44]. However, not all inauthentic and harmful coordination is carried out via automated means,[5] let alone authentic and harmless one. This makes the analysis of automation a useful complement, rather than a substitute, to the analysis of coordination.

Many bot detection methods have been proposed in the last few years [45–47], with Botometer being the *de facto* standard in the field. Here we resort to Botometer v4 [48] to obtain an estimate of the degree of automation in Twitter accounts. Botometer assesses an account's similarity to known social bot characteristics by analyzing a comprehensive range of features, including: account metadata, network of friends, social network structure, patterns of activity, and the use of language and sentiment expressions. For users who communicate in English, we use Botometer English language score, whereas for all other users we utilize the universal score. Similarly to our coordination score, also Botometer automation scores $\in [0, 1]$, reflecting a continuum of automation rather than a simple binary classification. To assess the degree of automation within a community, we compute

---

[5]Y. Roth, and N. Pickles (2020) "Bot or not? The facts about platform manipulation on Twitter." *Twitter*. https://blog.twitter.com/en_us/topics/company/2020/bot-or-not.html.

the mean automation score of its members. Scores $\simeq 0$ indicate communities with minimal automation, while scores $\simeq 1$ indicate communities whose members make large use of automation.

4.2.3 Suspensions

Account suspensions are the actions taken by social media platforms to temporarily or permanently deactivate accounts following repeated policy violations. Among the reasons for account suspensions is engaging in antisocial behavior, spamming, or using hateful and toxic speech [49]. Notably, Twitter's suspension practices have often been associated with political activities [50]. Analyzing account suspensions is thus relevant because each suspension provides evidence that the suspended account behaved in violation to the platform's policies. Suspensions are thus a proxy for harmful behaviors.

We obtain the suspension status of a Twitter account directly from Twitter's APIs,[6] which provide a distinct error code when an account is suspended [51–53]. The suspensions score for a coordinated community is then computed as the fraction of the members of that community that have been suspended for policy violations. This facet sheds light on the prevalence of harmful behavior within a community, enriching our analysis into the community's conduct and its repercussions on the broader ecosystem.

4.2.4 Political partisanship

In this study, assessing the degree of political partisanship of coordinated communities is particularly relevant given that our analysis is focused on the US 2020 presidential elections. Multiple methods have been recently proposed to compute social media scores of political partisanship, bias, or leaning [54–56]. Among them, here we compute a community score of political partisanship based on the political bias of the news shared by community members, which is a widely-used approach in this field [1, 57]. In detail, we first extract all URLs contained in the tweets shared by the members of each coordinated community. Then, we standardize each URL by resolving redirections from shortening services, by extracting the domain from the URL, and by grouping together URLs pointing to the same domain. We subsequently filter out all domains that do not correspond to news outlets. Finally, we assign a political partisanship score to each URL based on the political bias of the corresponding news outlet. We rely on Media Bias/Fact Check (MBFC)[7] to obtain an extensive list of news outlets and the corresponding political bias scores. MBFC is a well-known resource that provides expert-driven ratings of the political bias and factual accuracy of many news outlets. Being primarily targeted at US outlets, this resource is particularly suitable for our analysis. MBFC categorizes each news outlet into one of six ordinal political bias categories, ranging from extreme-left to extreme-right. Here we convert each category to a numeric score by respecting their order, such that a score of −1 corresponds to extreme-left news outlets, and a score of +1 corresponds to extreme-right ones, with all other values in between. We then compute the political partisanship score for each member of a coordinated community, which represents the average political bias of the news outlets shared by that member. Next, the political partisanship of a coordinated community is computed as the mean of the political partisanship

---

[6] https://developer.twitter.com/en/docs/twitter-api.

[7] mediabiasfactcheck.com.

scores of the members of that community. Finally, in those analyses where we are interested in the extent of political polarization independently of the political direction of such polarization, we consider the absolute value of the community-level political partisanship score, such that scores $\simeq 0$ indicate minimal polarization while scores $\simeq 1$ indicate strong polarization (either towards the left or right). Notably, being defined $\in [0, 1]$, this latter indicator is suitable to be compared to all other indicators used to characterize coordinated communities.

4.2.5 Nonfactuality

In the context of information disorder, coordination is often exploited to boost the spread of false or misleading information, with the ultimate goal of manipulating public opinion and sowing discord [10]. Here, we account for this possible use of coordinated behavior by measuring the (non)factuality of the news shared by members of the coordinated communities. Nonfactuality refers to the extent to which information shared within a community deviates from accurate and reliable news, thus implying the spread of mis- or disinformation.

Similarly to our approach for measuring political partisanship, we measure the degree of nonfactuality of each coordinated community by relying on MBFC, which evaluates news outlets based on their commitment to factual reporting. Specifically, MBFC assigns factual ratings to each news outlet according to six ordinal values that range from very high, to very low factuality. Here we convert these ratings to a numerical nonfactuality score $\in [0, 1]$, such that scores $\simeq 0$ indicate high factuality, while scores $\simeq 1$ indicate high nonfactuality (i.e., low factuality). User-level nonfactuality scores are obtained by averaging the scores of all news outlets shared by each user. Community-level scores are then obtained by averaging the scores of all members of each community.

5 Results

The application of our methodology brought to the detection and characterization of multiple coordinated communities, revealing a complex landscape of coordination. Figure 2 shows the filtered user similarity network related to the Twitter debate on US 2020 Presidential Election. The main coordinated communities are highlighted in figure with labels and different colors. In addition, different shades of color are applied to the nodes (i.e., users) in each community, so that strongly coordinated users are dark-colored and weakly coordinated ones are light-colored.

The analysis of the user similarity network reveals three large communities consisting of thousands of coordinated accounts, and seven medium- and small-sized communities. The largest community (REP) features a mix of highly coordinated users (dark-colored) surrounded by those with weaker coordination ties (light-colored). Other communities, such as DEM, PCO, IRN, and QCO, predominantly display weak coordination. In contrast, the BFR community exhibits a strong and relatively uniform coordination among its members, highlighted by the prevalence of dark-colored nodes. In the following we shed light on the intent, activity, and potential harmfulness of each coordinated community with a threefold analysis. To provide context for the detected coordinated communities, we first analyze the main narratives discussed by members of each community. Then, we further characterize each community from different angles, including: their degree of coordination, their use of automation, their political partisanship, the degree of factuality of the

Figure 2 User similarity network. Communities of coordinated accounts emerging within Twitter online debates related to the 2020 USA presidential election. Communities are color-coded. Accounts exhibiting higher coordination with other accounts are visualized with darker shades of colors
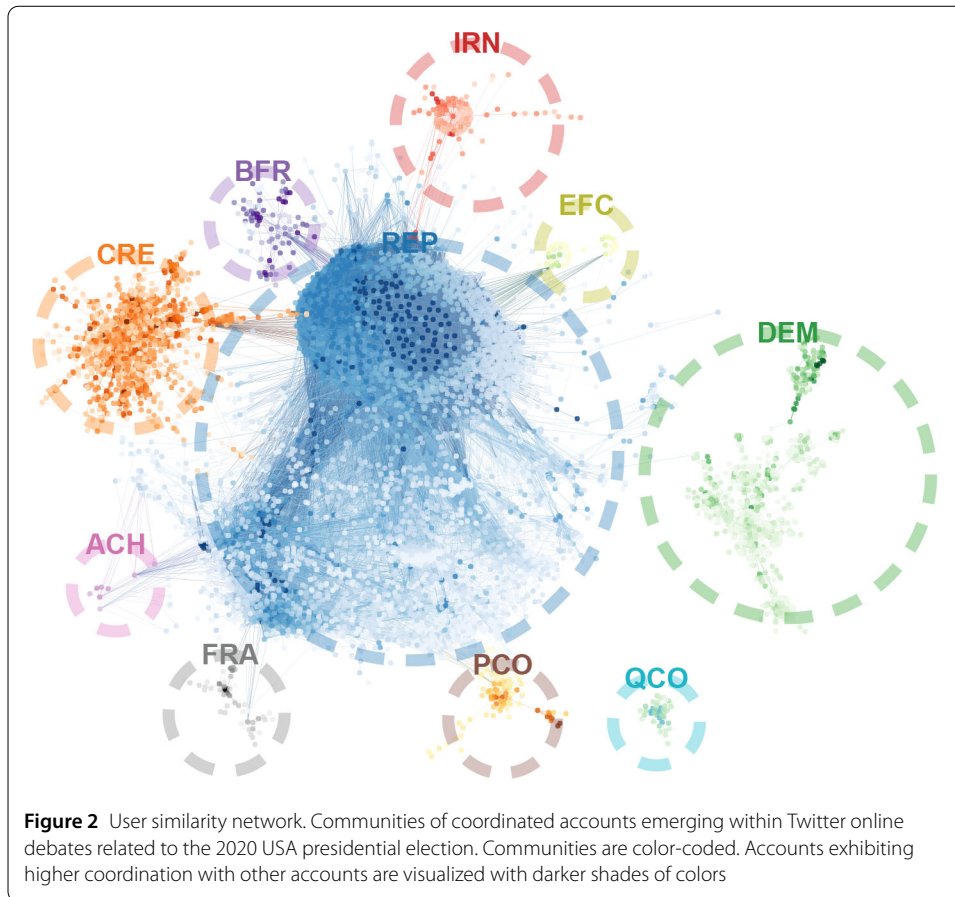
content they shared, and platform suspensions of coordinated accounts. Lastly, we simultaneously examine all these aspects, providing rich insights into the dynamics of coordinated communities.

5.1 Narratives

Here we present summary information for each coordinated community, together with the main narratives discussed by their members. This initial analysis allows us to label the communities and to provide context for their participation in the US 2020 online debate. The word clouds shown in Fig. 3 are obtained from the TF-IDF weighted hashtags used by the members of each community. As such, they provide information on the distinctive themes and narrative of each community. The main coordinated communities that took part in the online debate are presented in the following, ordered by decreasing number of members:

1. 🔵 REP: *Republicans (4522 users).* The largest coordinated community discussed hashtags supportive of the Republican candidate and former President Donald Trump, such as trump2020, americafirst, and maga. Despite some engagement with extremist and controversial hashtags, the majority of their activity leaned towards moderate conservatism.

2. 🟠 CRE: *Conspirative Republicans (1553 users).* This community also supported the Republican party. However, differently from REP, its members were predominantly

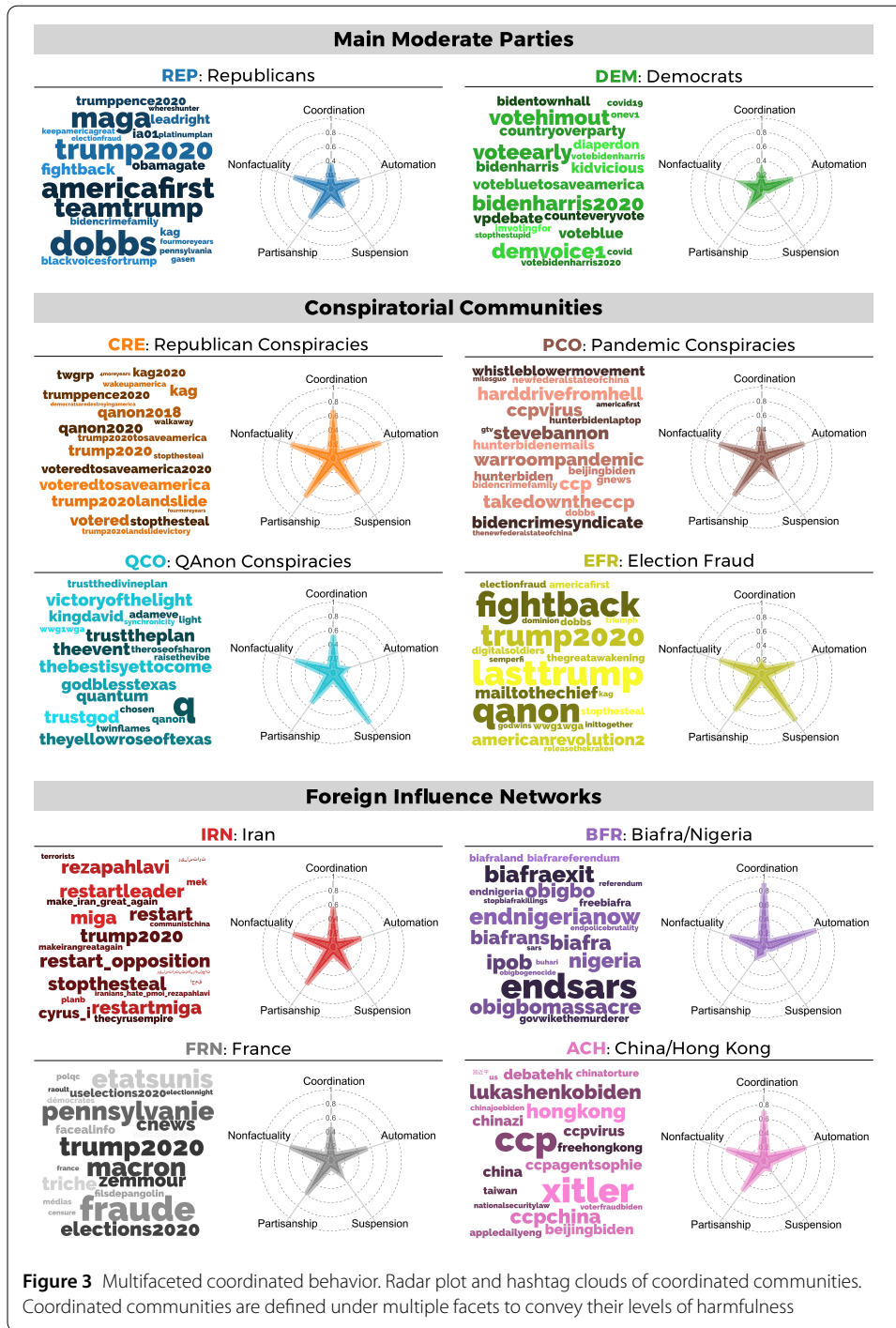**Figure 3** Multifaceted coordinated behavior. Radar plot and hashtag clouds of coordinated communities. Coordinated communities are defined under multiple facets to convey their levels of harmfulness

interested in multiple conspiracy theories, as shown in Fig. 3 by the presence of hashtags such as qanon2018 and qanon2020.

3. 🟢 DEM: *Democrats (1039 users).* Representing Biden and Democratic supporters, this community discussed topics aligned with Democratic ideals, encouraging early voting and a proactive attitude in the electoral process.

4. 🔴 IRN: *Iranians (96 users).* Narratives within this coordinated community are centered around the *Restart* movement, with hashtags such as restart_opposition and

miga (i.e., Make Iran Great Again). Restart was an Iranian political opposition movement, in support of Donald Trump due to his stance against the Iranian regime. By participating in the online electoral debate, Restart aimed to gain visibility and influence within the US political discourse.[8]

5. ● PCO: *Pandemic conspiracies (84 users).* Centered around a far-right news media platform that spread controversial narratives, this community engaged in various debated topics, ranging from conspiracy theories related to COVID-19 to political narratives critical of the Chinese government.

6. ● BFR: *Biafrans (69 users).* A dense and strongly-coordinated community advocating for Biafra's independence from Nigeria. Its members aimed to exploit Trump's presidency and the upcoming election to attract international support. They promoted their separatist goals and sought broader recognition for their movement.[9]

7. ● FRA: *French (49 users).* This small coordinated group is mostly composed of French accounts supporting Trump and the election fraud narrative. This community also sought to influence domestic French politics, including the endorsement of far-right presidential candidates [58].

8. ● QCO: *QAnon conspiracies (43 users).* A community focused on the QAnon conspiracy theory, portraying a battle between Donald Trump and the alleged deep state, with narratives encompassing the fight between good and evil [38, 59].

9. ● EFC: *Election fraud (34 users).* Focused on advancing claims of election fraud, and particularly those involving postal ballots, this coordinated group engaged heavily with the stopthesteal movement[10] and the corresponding narrative [38].

10. ● ACH: *Anti China (31 users).* A small community of Hong Kong protesters mobilized against the Chinese regime. They were showing support for Donald Trump while criticizing Joe Biden and the Democrats for their alleged complicity with authoritarian regimes.[11]

Overall, the above results on the role of the main coordinated communities in the US 2020 online debate resonate with the political climate of the time and corroborate previous research findings [22, 24, 29]. The analysis of the main coordinated communities revealed a striking predominance of communities aligned with Republican or far-right ideologies, and a single community associated with the Democrats. This distribution mirrors the political polarization observed during the election and highlights the better strategic use of social platforms by conservative groups [60, 61].
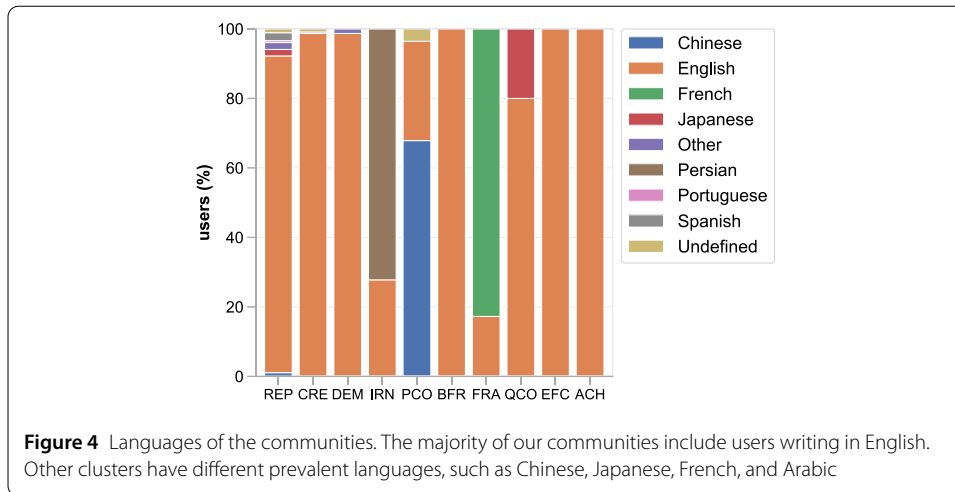
The previous analysis of the narratives of each detected community also highlights three distinct groups:

[8]Tabatabai, A. (2020, July 15)., "QAnon Goes to Iran", 15 July. *Foreign Policy*. https://foreignpolicy.com/2020/07/15/qanon-goes-to-iran/.

[9]Akinwotu, E. (2020, October 31). "'He just says it as it is': why many Nigerians support Donald Trump". *The Guardian*. https://www.theguardian.com/world/2020/oct/31/he-just-says-it-as-it-is-why-many-nigerians-support-donald-trump.

[10]Atlantic Council's DFRLab (2021, February 10). "#StopTheSteal: Timeline of Social Media and Extremist Activities Leading to 1/6 Insurrection". *Just Security*. https://www.justsecurity.org/74622/stopthesteal-timeline-of-social-media-and-extremist-activities-leading-to-1-6-insurrection/.

[11]Davidson, H. (2020, November 12). "Why are some Hong Kong democracy activists supporting Trump's bid to cling to power?". *The Guardian*. https://www.theguardian.com/us-news/2020/nov/13/trump-presidency-hong-kong-pro-democracy-movement.

Figure 4 Languages of the communities. The majority of our communities include users writing in English. Other clusters have different prevalent languages, such as Chinese, Japanese, French, and Arabic
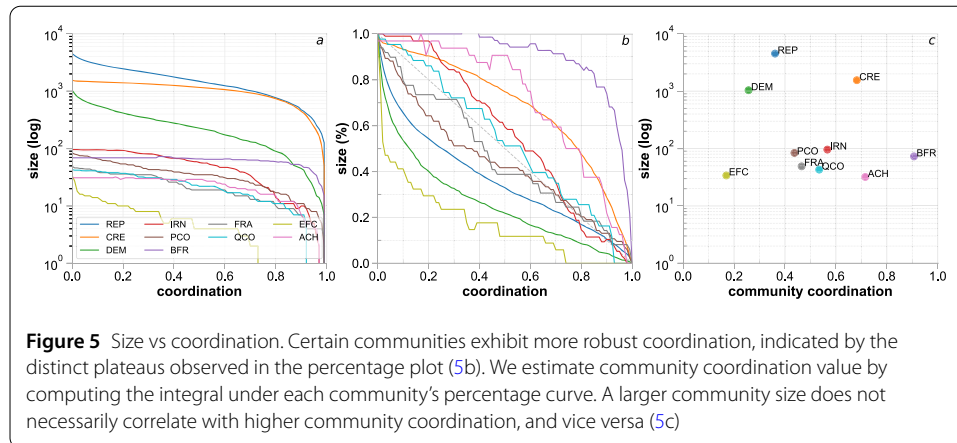
- *Moderate parties*: this group includes coordinated communities representing the two major political parties involved in the election (i.e., REP and DEM). Users from the communities in this group engaged in political discourse with moderate tones.
- *Conspiratorial communities*: this group includes coordinated communities who supported various conspiracy theories, including those related to QAnon (i.e., CRE, QCO), election fraud claims (i.e., EFC), and the pandemic (i.e., PCO). The activity of these types of communities on social platforms typically contributes to the spread of misinformation, shaping false narratives and potentially influencing public opinion.
- *Foreign influence networks*: coordinated communities in this group have ties to foreign politics (i.e., IRN, BFR, FRA, ACH). Their involvement in the debate was either aimed at undermining public trust and tampering with the democratic process, or to draw international attention to local affairs. Similar coordinated initiatives were observed for the US 2018 midterm elections [62]. Figure 4 shows the predominant language used by the members of each coordinated community. For the communities in this group, the analysis indicates a user base speaking both the native language and English. This hints at strategic involvement with geopolitical agendas within the US 2020 political debate.

5.2 Multifaceted characterization

Our approach to understanding coordinated behavior is multi-dimensional, recognizing various facets that reveal motivations and potential impacts on the debate. In this section, we measure and discuss the various indicators introduced in Sect. 4.2, for each coordinated community.

5.2.1 Coordination

Each community showcases unique narratives and goals, alongside a spectrum of coordination intensities. Figure 5 shows the relationship between the absolute and relative size of each community and the coordination of their members, offering insights into the collective behavior of the coordinated communities. When analyzing Fig. 5a, the DEM and REP communities representing the main moderate parties, show a downward trend in size at increasing levels of coordination. This suggests a more dispersed and possibly grassroots
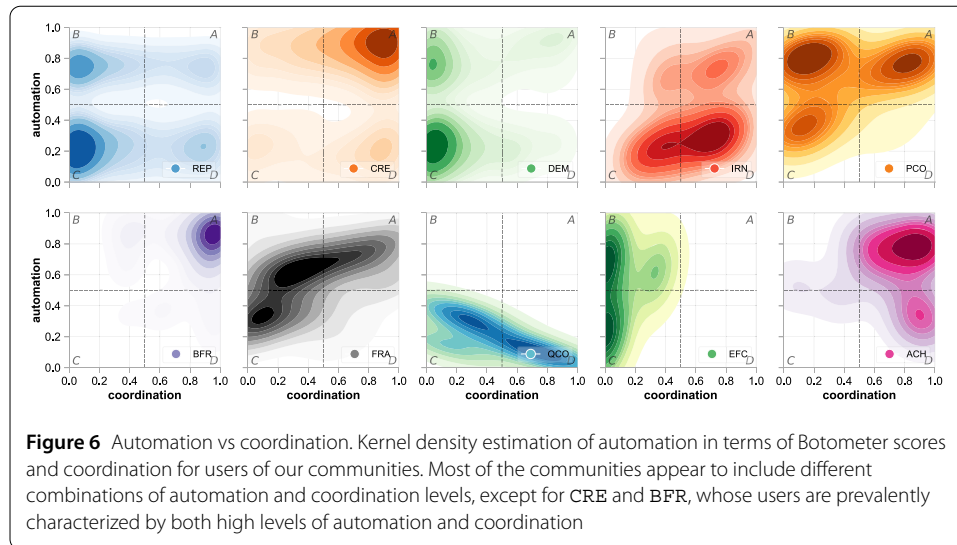
Figure 5 Size vs coordination. Certain communities exhibit more robust coordination, indicated by the distinct plateaus observed in the percentage plot (5b). We estimate community coordination value by computing the integral under each community's percentage curve. A larger community size does not necessarily correlate with higher community coordination, and vice versa (5c)

level of engagement, where individual actors engage in political discourse without a centralized or strongly coordinated strategy. The presence of a small core of strongly coordinated users might drive the agenda of the group, while the broader base reflects a more organic, diverse political conversation. For communities characterized by conspiracy theories we observe diverse patterns of coordination. For instance, CRE shows a higher level of sustained coordination, as implied by the plateau in its curve. This reflects a strong, cohesive group possibly employing centralized strategies to spread their narratives. In contrast, QCO lacks a pronounced plateau, suggesting the presence of loosely connected conspiracy theorists. The foreign influence networks, particularly BFR and ACH, stand out with their plateaued coordination, reflecting a consistent and stable level of high coordination levels. These communities might employ a strategic organized approach, with possibly centralized coordination to effectively influence or engage with the broader political conversation.

Figure 5b repeats the analysis in relative terms, removing possible biases deriving from the different absolute sizes of the various communities. Results indicate that a community's size does not necessarily correlate with the strength of coordination. This characteristic is even more visible in Fig. 5c. The scatterplot describes the relationship between the absolute size of each community and an aggregated indicator of community coordination, computed for each community as the area the respective curve in Fig. 5b. As shown, both large and small communities can achieve large community coordination scores, emphasizing the non-linear relationship between community size and coordination.
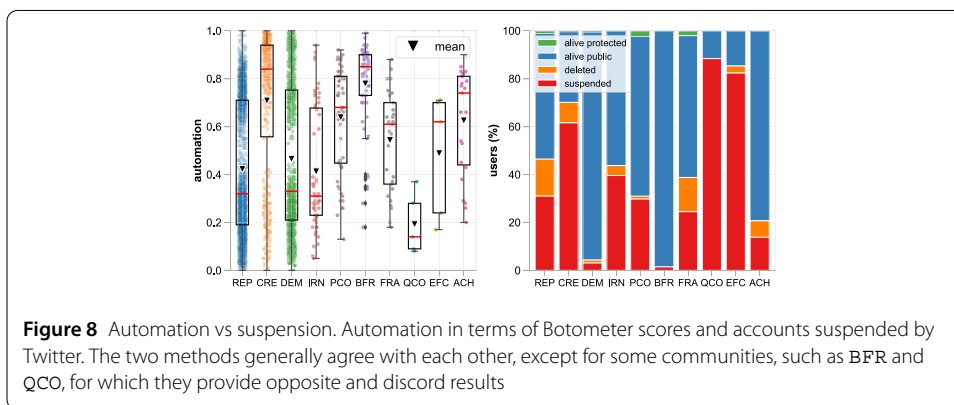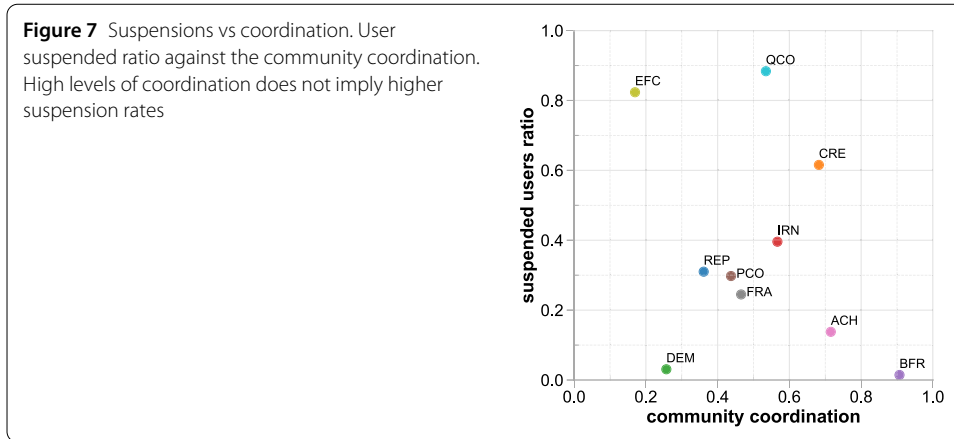
5.2.2 Automation

Figure 6 shows a Kernel Density Estimations (KDE) to examine the Botometer scores relative to community coordination. The KDE plots are divided into four quadrants, each representing different combinations of automation and coordination levels among members of the coordinated communities. Quadrant A (top-right) highlights users with both high automation and high coordination. This area is indicative of communities where automated accounts are prevalent and engage in highly synchronized behaviors. Quadrant B (top-left) captures users with high automation but low coordination. These accounts, while automated, do not exhibit coordination patterns, suggesting isolated or sporadic engagement. Quadrant C (bottom-left) includes users with low levels of both automation and coordination, likely consisting of genuine users engaging in organic interactions. Quad-

Figure 6 Automation vs coordination. Kernel density estimation of automation in terms of Botometer scores and coordination for users of our communities. Most of the communities appear to include different combinations of automation and coordination levels, except for `CRE` and `BFR`, whose users are prevalently characterized by both high levels of automation and coordination

rant D (bottom-right) shows users with a high coordination achieved without the use of automation. Results show that communities categorized as moderate parties typically display lower automation levels, as shown by the prevalent densities in Quadrant C. This suggests that their online activity may be driven by genuine user engagement rather than automated tools. These communities prioritize organic interactions. Conversely, conspiracy-oriented communities show markedly different results. The `CRE` community exhibits a strong user presence in Quadrant A, indicating high levels of both automation and coordination. This suggests a reliance on automated accounts to disseminate conspiratorial narratives, potentially to bolster their visibility and influence. For conspiracy communities such as `QCO` and `EFC`, which do not prominently feature in Quadrant A, the relationship between automation and coordination is more nuanced. These groups may not heavily rely on automated accounts for their activities, suggesting a different, more organic, kind of engagement strategy within the conspiracy theory domain. Overall, these results suggests the lack of a one-size-fits-all relationship between automation and coordination within communities of conspiracy users. Finally, foreign influence network accounts, represented by communities like `BFR` and `ACH`, are also heavily clustered in Quadrant A, indicating high automation alongside strong coordination. These communities likely involve automated accounts to support their agendas, whether for geopolitical influence or to amplify their messages. In summary, while moderate parties tend toward more organic methods, conspiracy and foreign influence networks appear to strategically leverage more automation to magnify their impact and shape the discourse, albeit with interesting differences.
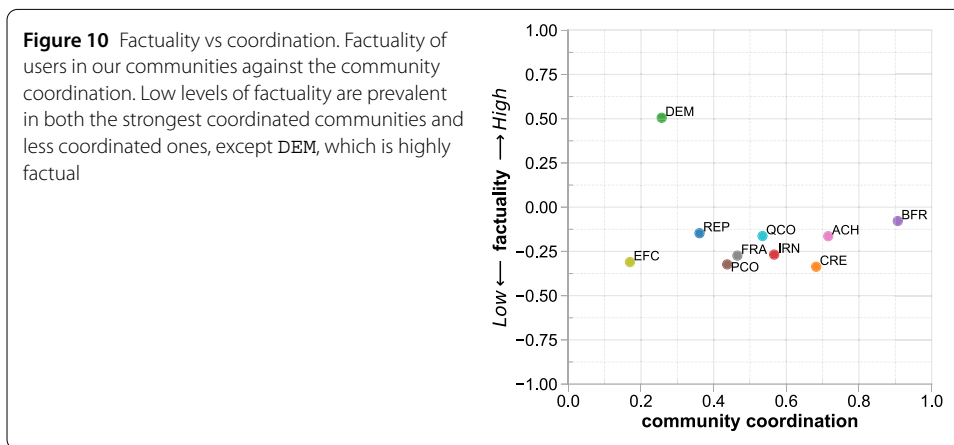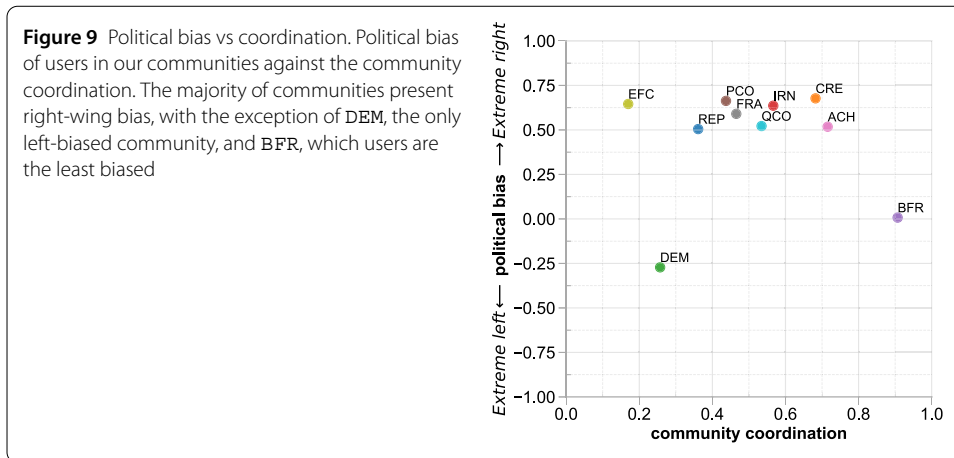
5.2.3 Suspensions

In Fig. 7, we analyze the interplay between community coordination and account suspensions. We compare the proportion of suspended users with each community's coordination value. Results indicate that moderate parties display lower suspension rates, suggesting their alignment with platform rules. Conspiracy communities, particularly `CRE` and `QCO`, often show higher suspension rates, which may correlate with the controversial nature of the content they share. Their higher suspension rates might be a result of systematic platform policy violations, including the spread of false narratives or other forms of

Figure 7 Suspensions vs coordination. User suspended ratio against the community coordination. High levels of coordination does not imply higher suspension rates

Figure 8 Automation vs suspension. Automation in terms of Botometer scores and accounts suspended by Twitter. The two methods generally agree with each other, except for some communities, such as `BFR` and `QCO`, for which they provide opposite and discord results

disruptive behavior. Instead, foreign influence networks like `BFR`, `ACH`, and `IRN` present diverse suspension rates. Some communities, such as `BFR`, maintain a high level of coordination with very few suspensions, indicating an organic or a more strategic approach that carefully avoids crossing platform rules. Others, like `ACH`, experience more suspensions, perhaps due to more aggressive tactics or content. These differences may reflect the diverse objectives and strategies within influence networks, from overt political activism to more covert influence operations. Alternatively, they might indicate the limited effectiveness of Twitter at moderating a subset of misbehaviors. Finally, we compare the levels of automation with suspension rates. Figure 8a shows that communities like `CRE` and `BFR` show high automation scores. Yet, as depicted in Fig. 8b, Twitter's enforces suspensions towards user of `CRE`, but not towards `BFR`. In contrast, the `QCO` community, while showing lower automation, has a high rate of user suspensions, suggesting that factors beyond automation contribute to Twitter's suspension decisions. These findings suggest complex, non-linear relationships between automation and suspensions: high automation doesn't always result in suspensions, and vice versa.

5.2.4 Political partisanship

In Fig. 9, we illustrate the political bias of the communities in relation to their levels of coordination. The results formalize our previous insights into partisanship. In fact, the majority of communities align with right-leaning ideologies. In detail, the `REP` community exhibits a right-leaning bias, albeit slightly more centrist and moderate compared to the

Figure 9 Political bias vs coordination. Political bias of users in our communities against the community coordination. The majority of communities present right-wing bias, with the exception of DEM, the only left-biased community, and BFR, which users are the least biased

Figure 10 Factuality vs coordination. Factuality of users in our communities against the community coordination. Low levels of factuality are prevalent in both the strongest coordinated communities and less coordinated ones, except DEM, which is highly factual

other right-leaning communities. Instead, the DEM community displays a left-leaning bias with lower coordination values. Conspiracy-oriented communities show a right-leaning bias. The CRE community, in particular, shows a high coordination value, suggesting a co-ordinated effort to share a strong partisan viewpoint. Lastly, foreign influence networks show varying degrees of bias. In particular, the BFR community presents an interesting combination of a politically-centered bias and high coordination. This possibly indicates that while their political messaging may be more moderate or diverse, their approach to spreading these messages is highly organized and possibly strategic, aiming to influence or sway public opinion in a focused manner. In summary, the prevalence of far-right ideologies identified through our community detection criteria aligns with observed political polarization during the election, indicating more strategic utilization of social platforms by conservative groups [60, 61].

5.2.5 Nonfactuality

In Fig. 10, we show the relationship between the community coordination values and the factuality scores of the content shared within each community. The majority of our communities are associated with low factuality levels, indicating a propensity for sharing low quality information. When focusing on communities of moderate parties, we find varied factuality levels: the DEM community shows a commitment to higher factuality and lower levels of coordination. In contrast, the REP community shows lower levels of credi-

bility. Conversely, conspiracy-oriented communities generally display low degrees of factuality. Indeed, conspiracy narratives frequently exhibit lower levels of factual reporting and tend to prioritize the reinforcement of ideological beliefs. Remarkably, the CRE community emerges as the least factual, indicating a high propensity for deceitful behaviors. Finally, foreign influence networks show diverse balance between factual and non-factual content as a means of legitimizing their cause while promoting their strategic objectives. The use of factuality may be instrumental in gathering support and credibility, especially when reaching out beyond their immediate community to influence broader discussions. In summary, each category utilizes a distinct mix of factuality and coordination in their online behavior, reflecting their unique goals and strategies.

5.3 Composite characterization

So far, we have analyzed each facet separately. However, our previous results suggest that these facets form a complex tapestry that requires a comprehensive approach to fully understand online community coordination dynamics. Here, we weave these facets into a multifaceted analysis to provide rich insights into the inauthenticity and harmfulness of the different types of coordinated communities. Figure 3 shows a radar chart for each coordinated community that aggregates all dimensions of our multifaceted analysis. The size and shape of each plot reflect the collective behavior and the high-level characteristics of each community, offering insights into their potential impact on the online discourse. With the exception of the *coordination* dimension that is in theory neutral – albeit frequently exploited for nefarious purposes – all other dimensions are defined in such a way that large scores (i.e, $\simeq 1$) correspond to malicious behaviors. Consequently, coordinated communities whose radar chart spans a larger area in Fig. 3 can be considered as more malicious than those spanning a smaller area. To this end, the radar charts in Fig. 3 serve as compact summaries of the complex and multifaceted coordinated behavior of the different communities. They are capable of highlighting differences between the communities, as well as strikingly suspicious coordinated behaviors.

As an example, the analysis of the radar charts of Fig. 3 reveals that communities of moderate parties tend to exhibit smaller plots featuring moderate or low levels in the majority of facets. This pattern suggests organic, harmless, and authentic forms of engagement, potentially indicative of genuine activism rather than deceptive or manipulative practices. Conversely, conspiratorial communities display markedly larger radar charts, each with pronounced spikes in multiple dimensions. In particular, all conspiratorial communities features high levels of nonfactuality and partisanship. Some feature strong reliance on automation, and all but PCO feature massive account suspensions. Out of all the coordinated communities that we investigated, CRE emerges as the most harmful one. Finally, foreign influence networks show overall similar radar charts to those of the conspiratorial communities. Indeed, all foreign influence networks feature high nonfactuality, partisanship, and automation. However, contrarily to conspiratorial communities, all feature relatively low levels of suspensions. This last finding might indicate a reduced effectiveness of Twitter – or a reduced interest – in detecting and moderating these specific types of coordinated groups. Overall, our nuanced results highlight the challenges of distinguish genuine, organic activism from orchestrated influence campaigns. In summary, each community's radar chart serves as a visual fingerprint of the multifaceted characteristics of that community, combining indicators across multiple dimensions. The analysis of the size and

shape of these fingerprints contributes to the identification of genuine grassroots movements and to the understanding of harmful and inauthentic coordinated behavior, thus contributing to the ongoing efforts aimed at safeguarding the health and integrity of the online political discourse.

6 Discussion

We investigated the Twitter debate surrounding the US 2020 Presidential Election. In general, our evidence points towards a predominance of right-wing coordination, aligning with previous findings of right-leaning communities leveraging social media effectively to disseminate their content [60, 61].

6.1 The heterogeneous landscape of online coordination

While prior studies have largely focused on English-speaking clusters, our inclusion of non-English language communities provides broader insights. In fact, the presence of foreign influence networks that engage with US political debates has profound implications on political communication and the democratic processes [62]. On one hand, the existence of genuine, issue-focused groups suggests a healthy engagement with political issues. On the other hand, the identification of foreign influence networks highlights the intricacy of online debates and the need for a nuanced strategy to comprehend and counteract potentially manipulative activities [5]. Similarly, the presence of conspiratorial communities highlights the potential for polarizing, divisive conversations that could damage social cohesion. The intricate nature of online coordination, coupled with the diverse range of heterogeneous coordinated behaviors that we observed, presents inherent challenges in their detection and disambiguation. For example, our results underscore the necessity for nuanced methodological approaches to accurately study these phenomena and effectively distinguish between malicious forms of coordination and neutral or benign ones. To this end, our work provides contributions towards a nuanced and detailed characterization of coordinated communities. However, the same heterogeneity also presents an opportunity, as the systematic and scientifically-grounded study of coordinated behavior could represent a unified framework for investigating and potentially mitigating the various afflictions within our online ecosystems.

6.2 Offline consequences of online coordination

Our results also shed some light on the thin boundary between online and offline events, which opens up the possibility to discern potential real-world consequences via the analysis of online activities. For instance, communities marked by pervasive harmful behavior, such as the Conspirative Republicans (CRE), could set the stage for real-world turmoil, similar to what led to the Capitol Hill insurrection [38]. In contrast, communities with high coordination but low partisanship and nonfactuality may reflect organic activism that ignites or reinforces public demonstrations and societal change, reminiscent of the pro-democracy protests in Hong Kong [63–65] and similar movements in Nigeria [66–68]. Similarly, a community characterized by mild political opposition and low levels of harmfulness (e.g., IRN) could hint at hidden discontent that may eventually lead to real actions of dissent [69, 70]. The previous examples highlight the importance of studying online coordination, which holds the potential to influence, and at times also anticipate, significant real-world events ranging from political protests and social movements to market

fluctuations and public health crises. Therefore, unraveling the dynamics of online coordination – particularly via nuanced and multifaceted analyses – allows gaining insights that enhance the comprehension of contemporary societal trends, both from a descriptive and possibly even from a predictive standpoint.

6.3 Securing the integrity of online political discourse

This study comes in the wake of the European Commission's decision to open formal proceedings against Twitter for possible violations of the Digital Services Act (DSA), including failure to take effective measures to combat information manipulation on the platform [71]. Interestingly, our study possibly provides some insights in this regard, as part of our findings indicate a diminished efficacy – or attention – from Twitter in detecting and moderating certain specific categories of coordinated groups. Paramount among them are the foreign influence networks that tampered with the US 2020 online debate. These results underscore the importance of post-hoc analyses of coordinated behavior as auditing tools for evaluating the moderation practices of large social media platforms [72]. By independently assessing the discrepancies and effectiveness of moderation efforts, particularly in the context of impending major electoral events that are bound to be replete with manipulation attempts, our study provides valuable insights into the areas where platforms may need to bolster their detection and intervention strategies. This proactive approach to auditing moderation practices might serves as an added mechanism for fostering the integrity and transparency of online discourse, ultimately contributing to safeguarding the democratic process from undue influence and manipulation.

6.4 Limitations

Our approach, while insightful, faces some limitations. First, the restricted accessibility of Twitter data, following Twitter's transition to X.com, limits content retrieval and study replicability [73]. Additionally, methods such as Botometer may not only suffer from these restrictions but also come with other intrinsic limitations [74, 75]. Another limitation relates to the user selection process. While leveraging the retweet volume as a metric for selecting the top 1% of users to analyze is recognized as an adequate approach for identifying influential users [29, 40], it may also introduce a selection bias. However, political bias in our dataset likely mirrors the differing behaviors of political communities on social media, with research showing that right-leaning users are more active and utilize these platforms more effectively than their left-leaning counterparts [60, 61]. Finally, the use of MBFC to evaluate news source bias and credibility may introduce subjectivity due to the evaluators' personal biases and the variability in their flagging criteria, which may not be universally accepted. Moreover, limitations in MBFC's methodology transparency, potential inaccuracies in assessments, and incomplete coverage of smaller news outlets further complicate the reliability of their evaluations. However, the framework employed in this study is flexible, allowing the use of alternative tools for news source evaluation (e.g., Newsguard), user selection (e.g., top 1% of active users), and bot detection [46, 76–78]. Likewise, our proposed framework can also incorporate additional relevant facets, like toxicity levels and psycholinguistic traits. Furthermore, our characterization approach is applicable beyond political debates to other information operations datasets, including those about conspiracy theories [79, 80] and infodemics [81, 82].

7 Conclusions

We carried out a thorough investigation of coordinated behavior occurred on Twitter (now X) in the run up to the US 2020 Presidential Election. Our approach is based on an established state-of-the-art coordination detection methodology, followed by a novel and nuanced approach to the characterization of coordinated communities that simultaneously examines the degree of coordination of the different communities, their use of automation, their political partisanship, their factuality, and platform suspensions of coordinated accounts. We demonstrate that the conjoint analysis of these aspects provides rich insights into the intent, activity, and harmfulness of the different coordinated communities. Among the notable findings of our study is the identification of three distinct groups of coordinated communities that participated in the online electoral debate: moderate political communities mostly involved in organic and harmless discourse, conspiratorial communities spreading misinformation and fueling political polarization, and foreign influence networks aiming to promote their local agendas.

By offering a nuanced view of coordinated behavior, this study offers orthogonal perspectives to those of existing studies. This research paves the way for future explorations into the multifaceted aspects of online coordinated behavior, allowing for the incorporation of additional analytical dimensions. Such comprehensive analyses are essential for gaining deeper insights into the impact of coordinated behavior on political discourse and its potential to both benefit and harm the democratic process.

Declarations

Author details
[1]Institute of Informatics and Telematics, National Research Council, Pisa, Italy. [2]Department of Information Engineering, University of Pisa, Pisa, Italy.

References

 1. Bovet A, Makse HA (2019) Influence of fake news in Twitter during the 2016 US presidential election. Nat Commun 10(1):1–14
 2. Grinberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D (2019) Fake news on Twitter during the 2016 US presidential election. Science 363(6425):374–378
 3. Bessi A, Ferrara E (2016) Social bots distort the 2016 US presidential election online discussion. First Monday 21(11–7)
 4. Shao C, Ciampaglia GL, Varol O, Yang K-C, Flammini A, Menczer F (2018) The spread of low-credibility content by social bots. Nat Commun 9(1):1–9
 5. Starbird K, Arif A, Wilson T (2019) Disinformation as collaborative work: surfacing the participatory nature of strategic information operations. In: Proceedings of the ACM on human-computer interaction 3(CSCW), pp 1–26
 6. Alizadeh M, Shapiro JN, Buntain C, Tucker JA (2020) Content-based features predict social media influence operations. Sci Adv 6(30)
 7. Alam F, Cresci S, Chakraborty T, Silvestri F, Dimitrov D, Da San Martino G, Shaar S, Firooz H, Nakov P (2022) A survey on multimodal disinformation detection. In: The 29th international conference on computational linguistics (COLING), pp 6625–6643
 8. Giglietto F, Righetti N, Rossi L, Marino G (2020) It takes a village to manipulate the media: coordinated link sharing behavior during 2018 and 2019 Italian elections. Inf Commun Soc 23(6):867–891
 9. Nizzoli L, Tardelli S, Avvenuti M, Cresci S, Tesconi M (2021) Coordinated behavior on social media in 2019 UK general election. In: Proceedings of the international AAAI conference on web and social media, vol 15, pp 443–454
10. Weber D, Neumann F (2021) Amplifying influence through coordinated behaviour in social networks. Soc Netw Anal Min 11(1):1–42
11. Tardelli S, Nizzoli L, Tesconi M, Conti M, Nakov P, Martino GDS, Cresci S (2023) Temporal dynamics of coordinated online behavior: stability, archetypes, and influence. arXiv preprint. arXiv:2301.06774
12. Pacheco D, Hui P-M, Torres-Lugo C, Truong BT, Flammini A, Menczer F (2021) Uncovering coordinated networks on social media: methods and case studies. In: Proceedings of the 15th international AAAI conference on web and social media (ICWSM'21), vol 15, pp 455–466
13. Weber D, Neumann F (2020) Who's in the gang? Revealing coordinating communities in social media. In: The 2020 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM'20), pp 89–93
14. Magelinski T, Ng L, Carley K (2022) A synchronized action framework for detection of coordination on social media. J Online Trust Saf 1(2)
15. Steinert-Threlkeld ZC, Mocanu D, Vespignani A, Fowler J (2015) Online social networks and offline protest. EPJ Data Sci 4(1):19
16. Starbird K (2019) Disinformation's spread: bots, trolls and all of us. Nature 571(7766):449–450
17. Gleicher N (2018) Coordinated inauthentic behavior explained. https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/
18. Starbird K, DiResta R, DeButts M (2023) Influence and improvisation: participatory disinformation during the 2020 US election. Soc Media Soc 9(2):20563051231177943
19. Pennycook G, Rand DG (2021) Examining false beliefs about voter fraud in the wake of the 2020 presidential election. The Harvard Kennedy School Misinformation Review
20. Calvillo DP, Rutchick AM, Garcia RJ (2021) Individual differences in belief in fake news about election fraud after the 2020 US election. Behav Sci 11(12):175
21. Rossini P, Stromer-Galley J, Korsunska A (2021) More than "fake news"? The media as a malicious gatekeeper and a bully in the discourse of candidates in the 2020 US presidential election. J Lang Polit 20(5):676–695
22. Sharma K, Ferrara E, Liu Y (2022) Characterizing online engagement with disinformation and conspiracies in the 2020 US presidential election. In: Proceedings of the international AAAI conference on web and social media, vol 16, pp 908–919
23. Ebermann D (2022) Conspiracy theories on Twitter in the wake of the US presidential election 2020. From Uncertainty to Confidence and Trust 18:51
24. Ferrara E, Chang H, Chen E, Muric G, Patel J (2020) Characterizing social media manipulation in the 2020 US presidential election. First Monday
25. Shevtsov A, Tzagkarakis C, Antonakaki D, Ioannidis S (2022) Identification of Twitter bots based on an explainable machine learning framework: the US 2020 elections case study. In: Proceedings of the international AAAI conference on web and social media, vol 16, pp 956–967
26. Abilov A, Hua Y, Matatov H, Amir O, Naaman M (2021) Voterfraud2020: a multi-modal dataset of election fraud claims on Twitter. In: Proceedings of the international AAAI conference on web and social media, vol 15, pp 901–912
27. Linhares RS, Rosa JM, Ferreira CH, Murai F, Nobre G, Almeida J (2022) Uncovering coordinated communities on Twitter during the 2020 US election. In: 2022 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM). IEEE, pp 80–87
28. Ng LHX, Carley KM (2022) Online coordination: methods and comparative case studies of coordinated groups across four events in the United States. In: 14th ACM web science conference 2022, pp 12–21
29. Da Rosa JM, Linhares RS, Ferreira CHG, Nobre GP, Murai F, Almeida JM (2022) Uncovering discussion groups on claims of election fraud from Twitter. In: International conference on social informatics. Springer, Berlin, pp 320–336
30. Tran HD (2021) Studying the community of trump supporters on Twitter during the 2020 US presidential election via hashtags #maga and #trump2020. Journal Media 2(4):709–731
31. Green J, Hobbs W, McCabe S, Lazer D (2022) Online engagement with 2020 election misinformation and turnout in the 2021 Georgia runoff election. Proc Natl Acad Sci 119(34):2115900119
32. Wang EL, Luceri L, Pierri F, Ferrara E (2023) Identifying and characterizing behavioral classes of radicalization within the qanon conspiracy on Twitter. In: Proceedings of the international AAAI conference on web and social media, vol 17, pp 890–901
33. Serrano MÁ, Boguná M, Vespignani A (2009) Extracting the multiscale backbone of complex weighted networks. Proc Natl Acad Sci 106(16):6483–6488

34. Giglietto F, Righetti N, Rossi L, Marino G (2020) Coordinated link sharing behavior as a signal to surface sources of problematic information on Facebook. In: International conference on social media and society, pp 85–91

35. Fazil M, Abulaish M (2020) A socialbots analysis-driven graph-based approach for identifying coordinated campaigns in Twitter. J Intell Fuzzy Syst 38(3):2961–2977

36. Ng LHX, Cruickshank IJ, Carley KM (2022) Cross-platform information spread during the January 6th capitol riots. Soc Netw Anal Min 12(1):133

37. Ng LHX, Cruickshank IJ, Carley KM (2023) Coordinating narratives framework for cross-platform analysis in the 2021 US Capitol riots. Comput Math Organ Theory 29(3):470–486

38. Vishnuprasad PS, Nogara G, Cardoso F, Giordano S, Cresci S, Luceri L (2024) Tracking fringe and coordinated activity on Twitter leading up to the US Capitol attack. In: The 18th international AAAI conference on web and social media (ICWSM'24)

39. Pei S, Muchnik L, Andrade JS Jr, Zheng Z, Makse HA (2014) Searching for superspreaders of information in real-world social media. Sci Rep 4:5547

40. Flamino J, Galeazzi A, Feldman S, Macy MW, Cross B, Zhou Z, Serafino M, Bovet A, Makse HA, Szymanski BK (2023) Political polarization of news media and influencers on Twitter in the 2016 and 2020 US presidential elections. Nat Hum Behav 7:904–916

41. Firdaus SN, Ding C, Sadeghian A (2021) Retweet prediction based on topic, emotion and personality. Online Soc Netw Media 25:100165

42. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech Theory Exp 10008:1–12

43. Cresci S (2020) A decade of social bot detection. Commun ACM 63(10):72–83

44. Khaund T, Kirdemir B, Agarwal N, Liu H, Morstatter F (2021) Social bots and their coordination during online campaigns: a survey. IEEE Trans Comput Soc Syst 9(2):530–545

45. Kudugunta S, Ferrara E (2018) Deep neural networks for bot detection. Inf Sci 467:312–322

46. Mannocci L, Cresci S, Monreale A, Vakali A, Tesconi M (2022) MulBot: unsupervised bot detection based on multivariate time series. In: The 10th IEEE international conference on big data (BigData'22). IEEE, pp 1485–1494

47. Dimitriadis I, Dialektakis G, Vakali A (2024) Caleb: a conditional adversarial learning framework to enhance bot detection. Data Knowl Eng 149:102245

48. Sayyadiharikandeh M, Varol O, Yang K-C, Flammini A, Menczer F (2020) Detection of novel social bots by ensembles of specialized classifiers. In: Proceedings of the 29th ACM international conference on information & knowledge management, pp 2725–2732

49. Ali S, Saeed MH, Aldreabi E, Blackburn J, De Cristofaro E, Zannettou S, Stringhini G (2021) Understanding the effect of deplatforming on social networks. In: 13th ACM web science conference 2021, pp 187–195

50. Chowdhury FA, Allen L, Yousuf M, Mueen A (2020) On Twitter purge: a retrospective analysis of suspended users. In: Companion proceedings of the web conference 2020, pp 371–378

51. Ferrara E (2017) Contagion dynamics of extremist propaganda in social networks. Inf Sci 418:1–12

52. Luceri L, Deb A, Badawy A, Ferrara E (2019) Red bots do it better: comparative analysis of social bot partisan behavior. In: Companion proceedings of the 2019 world wide web conference, pp 1007–1012

53. Bruno M, Lambiotte R, Saracco F (2022) Brexit and bots: characterizing the behaviour of automated accounts on Twitter during the uk election. EPJ Data Sci 11(1):17

54. Fagni T, Cresci S (2022) Fine-grained prediction of political leaning on social media with unsupervised deep learning. J Artif Intell Res 73:633–672

55. Stefanov P, Darwish K, Atanasov A, Nakov P (2020) Predicting the topical stance and political leaning of media using tweets. In: Proceedings of the 58th annual meeting of the association for computational linguistics (ACL'20), pp 527–537

56. Jiang J, Ren X, Ferrara E (2023) Retweet-BERT: political leaning detection using language features and information diffusion on social networks. In: Proceedings of the 17th international AAAI conference on web and social media (ICWSM'23), vol 17, pp 459–469

57. Chipidza W, Krewson C, Gatto N, Akbaripourdibazar E, Gwanzura T (2022) Ideological variation in preferred content and source credibility on reddit during the COVID-19 pandemic. Big Data Soc 9(1):20539517221076486

58. Abdine H, Guo Y, Rennard V, Vazirgiannis M (2022) Political communities on Twitter: case study of the 2022 French presidential election. arXiv preprint. arXiv:2204.07436

59. Tollefson J (2021) Tracking qanon: how Trump turned conspiracy-theory research upside down. Nature 590(7845):192–194

60. Pew Research Center (2014) Political polarization & media habits. https://www.pewresearch.org/journalism/2014/10/21/political-polarization-media-habits/. Accessed: 2024-02-01

61. Benkler Y, Faris R, Roberts H (2018) Network propaganda: manipulation, disinformation, and radicalization in American politics. Oxford University Press, London

62. Henschke A, Sussex M, O'Connor C (2020) Countering foreign interference: election integrity lessons for liberal democracies. J Cyber Policy 5(2):180–198

63. Lee F, Chan M, Chen H-T (2020) Social media and protest attitudes during movement abeyance: a study of Hong Kong university students. Int J Commun 14:4932–4951

64. Siu L (2019) Social media empower social movement. https://varsity.com.cuhk.edu.hk/index.php/2019/11/social-media-empower-social-movement/

65. Zervopoulos A, Alvanou AG, Bezas K, Papamichail A, Maragoudakis M, Kermanidis K (2022) Deep learning for fake news detection on Twitter regarding the 2019 Hong Kong protests. Neural Comput Appl 34(2):969–982

66. Uwazuruike AR (2020) #endsars: the movement against police brutality in Nigeria. Harvard Hum Rights J

67. Soladoye AA, Ojo AO (2020) Impact of social media on police brutality awareness in Nigeria. Int J Manag Mod Educ 1(1):14–27

68. Akerele-Popoola OE, Azeez AL, Adeniyi A (2022) Twitter, civil activisms and endsars protest in Nigeria as a developing democracy. Cogent Soc Sci 8(1):2095744

69. Shahi A, Abdoh-Tabrizi E (2020) Iran's 2019–2020 demonstrations: the changing dynamics of political protests in Iran. Asian Aff 51(1):1–41. https://doi.org/10.1080/03068374.2020.1712889

70. Dehghan E, Moon B, Keller T, Graham T, Bruns A, Angus D (2020) Investigating bots and coordinated influnce campaigns in Twitter discussions of the 2019-20 Iran protests. AoIR Selected Papers of Internet Research 2020. https://doi.org/10.5210/spir.v2020i0.11199

71. Commission E (2023) Commission opens formal proceedings against X under the Digital Services Act. https://ec.europa.eu/commission/presscorner/detail/en/IP_23_6709. Accessed: 2024-01-30

72. Trujillo A, Fagni T, Cresci S (2023) The dsa transparency database: auditing self-reported moderation actions by social media. arXiv:2312.10269

73. Tardelli S, Avvenuti M, Cola G, Cresci S, Fagni T, Gambini M, Mannocci L, Mazza M, Senette C, Tesconi M (2022) Cyber intelligence and social media analytics: current research trends and challenges

74. Grimme C, Assenmacher D, Adam L (2018) Changing perspectives: is it sufficient to detect social bots? In: Social computing and social media. User experience and behavior: 10th international conference, SCSM 2018, held as part of HCI international 2018, Las Vegas, NV, USA, July 15–20, 2018, proceedings, part I 10. Springer, Berlin, pp 445–461

75. Rauchfleisch A, Kaiser J (2020) The false positive problem of automatic bot detection in social science research. PLoS ONE 15(10):e0241045

76. Rodríguez-Ruiz J, Mata-Sánchez JI, Monroy R, Loyola-Gonzalez O, López-Cuevas A (2020) A one-class classification approach for bot detection on Twitter. Comput Secur 91:101715

77. Luceri L, Giordano S, Ferrara E (2020) Detecting troll behavior via inverse reinforcement learning: a case study of Russian trolls in the 2016 US election. In: Proceedings of the international AAAI conference on web and social media, vol 14, pp 417–427

78. Ezzeddine F, Luceri L, Ayoub O, Sbeity I, Nogara G, Ferrara E, Giordano S (2022) How "troll" are you? Measuring and detecting troll behavior in online social networks. arXiv preprint. arXiv:2210.08786

79. Amarasingam A, Argentino M-A (2020) The qanon conspiracy theory: a security threat in the making. CTC Sentinel 13(7):37–44

80. Gambini M, Tardelli S, Tesconi M (2024) The anatomy of conspiracy theorists: unveiling traits using a comprehensive twitter dataset. Comput Commun

81. Calamusa A, Tardelli S, Avvenuti M, Cresci S, Federigi I, Tesconi M, Verani M, Carducci A (2020) Twitter monitoring evidence of COVID-19 infodemic in Italy. Eur J Public Health 30(Supplement_5):ckaa165.066

82. Batzdorfer V, Steinmetz H, Biella M, Alizadeh M (2022) Conspiracy theories on Twitter: emerging motifs and temporal dynamics during the COVID-19 pandemic. Int J Data Sci Anal 13(4):315–333

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.