It Takes a Village to Manipulate the Media:

Coordinated Link Sharing Behaviour During 2018 and 2019 Italian elections

Fabio Giglietto (corresponding author)

University of Urbino Carlo Bo

fabio.giglietto@uniurb.it

Nicola Righetti

University of Urbino Carlo Bo

nicola.righetti@uniurb.it

Luca Rossi

Networks Data & Society research group

IT University of Copenhagen

lucr@itu.dk

Giada Marino

University of Urbino Carlo Bo

giada.marino@uniurb.it

Abstract

Over the last few years, a proliferation of attempts to define, understand and fight the spread of problematic information in contemporary media ecosystems emerged. Most of these attempts focus on false content and/or bad actors detection. In this paper, we argue for a wider ecological focus. Using the frame of *media manipulation* and a revised version of the "coordinated inauthentic behavior" original definition, the paper presents a study based on an unprecedented combination of Facebook data, accessed through the CrowdTangle API, and two datasets of Italian political news stories published in the run-up to the 2018 Italian general election and 2019 European election. By focusing on actors' collective behavior, we identified several networks of pages, groups, and verified public profiles ("entities"), that shared the same political news articles on Facebook within a very short period of time. Some entities in our networks were openly political, while others, despite sharing political content too, deceptively presented themselves as entertainment venues. The proportion of inauthentic entities in a network affects the wideness of the range of news media sources they shared, thus pointing to different strategies and possible motivations. The paper has both theoretical and empirical implications: it frames the concept of "coordinated inauthentic behavior" in existing literature, introduces a method to detect coordinated link sharing behavior and points out different strategies and methods employed by networks of actors willing to manipulate the media and public opinion.

Keywords*: political news, authenticity, coordinated inauthentic behavior, Facebook, CrowdTangle, elections, Italy*

Introduction

Citizens' exposure to online disinformation has become a major concern all over the world for a while now. The fear that malicious actors could sow the seeds of discord and distrust among digitally connected citizens, feeding polarization and stirring up insurmountable divisions so as to undermine the democratic process, has filled the unceasing flow of reports and news articles that have been published on the topic. Especially since 2016, when Brexit referendum in UK and US Presidential elections marked a turning point in the history of the relations between the Internet, social media, public opinion, and politics. From that moment on, it has become clear that the antagonist online participatory practices of sharing, collaborating and organizing collective actions (Jenkins, 2006; Shirky, 2008), which used to be considered the prerogative of democratizing forces fighting established powers, could be just as effective to support the spread of extremisms, hate speech, violence and false news (Marwick & Lewis, 2017).

Since the stakes are so high, researchers, governments and supranational institutions have clearly put a lot of effort into clarifying disinformation-related concepts, unravelling the complex, intertwined dimensions of the phenomenon, studying its empirical manifestations and trying to find solutions. Unfortunately, despite all efforts, stopping disinformation has proved harder than expected. A serious obstacle in fighting the problem effectively has certainly been the difficulty to mark a clear boundary between problematic and non-problematic information.

More recently, a new approach has proposed to circumvent this definitional obstacle by shifting the focus from content to dynamics of information spreading within online networks. Online content, indeed, benefits from a multitude of actors that amplify its reach, with a magnitude proportional to the popularity of online actors, the budget they can invest in social

media ads, and the activation of platform algorithms that prioritize better-performing images, videos, and posts, making popular content spread faster. It follows that "bad actors" may attempt to coordinate their efforts to get the initial plug which, once detected by the algorithm, may ignite the propagation machine and even attract the attention of mainstream media (Phillips, 2018) on the content they spread for profit or propaganda. Although this is not a new phenomenon (boyd, 2017), during the last few years we observed similar practices applied with the aim of enhancing the spread of political news stories.

Despite this new approach seems promising and also a leading and socially impactful social media company as Facebook currently employs it to fight disinformation by targeting what it has called "coordinated inauthentic behavior" (Gleicher, 2018), there are still few attempts to substantiate it through independent empirical studies. To start to fill this gap, the research we are going to present framed the ill-defined concept of coordinated inauthentic behavior in the existing scientific literature and tested the related "action-based" approach to disinformation detection through an unprecedented combination of Facebook data and two datasets of political news stories shared on Facebook in the six months before the 2018 general election and 2019 European election in Italy.

A key contribution of this work is the introduction of a method for identifying networks of pages and groups that coordinately shared the same news items on Facebook, a phenomenon we call "coordinated link sharing behavior". By employing this method – which is easily applicable to other national and political contexts – we were able to test whether the coordinated social media activity was associated with the spread of problematic information, that is, whether shifting the focus from online content to patterns of actions represents a fruitful approach to problematic information detection.

The paper is structured as follows: the first section frames the main challenges of problematic information detection, media manipulation and "coordinated inauthentic behavior" in the existing scientific literature. Then, the research questions are formulated and the datasets and methods used to answer them are detailed. Afterwards, the limitations of the research are enucleated and its results discussed to draw general conclusions and provide hints for future research.

Literature review and Research Questions

The widely recognized risks of misinformed citizens for healthy democracies brought a cohort of scholars to tackle this issue from a range of different perspectives. The deeply and still undergoing transformations of contemporary media ecologies led to a renewed interest in this topic resulting in a rapidly growing body of interdisciplinary scholarly works published during the last few years. Rather than attempting to provide a systematic review of these studies, the following paragraphs highlight the essential literature that frames our approach, clarify the terminology used and lead to our research questions.

The first paragraph highlights the results and limits of content-based and actor-based approaches to detect bad information and malicious actors. The second describes the media manipulation frame and the concepts of amplification and problematic information. The third and the fourth analyze the concept of "coordinated inauthentic behavior" and pinpoint its potential roots in the existing literature.

Challenges of content-based and actor-based problematic information detection

While unanimously recognizing misinformation as detrimental for healthy democracies, the existing literature is fragmented when it comes to defining the object of study. The lack of a shared, consistent and operationalizable definitions undermines both the attempts to estimate

its prevalence over legitimate information (Lazer et al., 2018) and measure the impact of misinformation on citizens' opinions or behaviour (Weeks & Gil de Zúñiga, 2019). While the issue of definition is widely recognized by scholars, the prevalent effect of the countless attempts to formally address it by way of new definitions and taxonomies (Molina, Sundar, Le, & Lee, 2019; Silverman, 2017; Wardle & Derakhshan, 2017) seems to have mostly dragged the scientific community deeper into the epistemological rabbit hole of "fake news" (Caplan, Hanson, & Donovan, 2018; HLEG EU Commission, 2018). To avoid falling in this trap, throughout this article we adopt the umbrella terminology of problematic information (Jack, 2017) to reference the whole spectrum of contents that range from deliberately or mistakenly false news to propaganda, gaslighting and satire.

Even when avoiding to differentiate the phenomenon based on the motivations of who creates and distributes problematic information, the simple basic choice of flagging the content as true or false is not always possible or advisable (Giglietto, Iannelli, Valeriani, & Rossi, 2019; Marwick, 2018). On the one hand, drawing such a clear distinction requires a significant amount of time, skills and resources for each content. On the other hand, due to the lack of commonly accepted definitions, making such calls is a great responsibility that deeply affects the outcomes of the study. For all these reasons, a large body of studies tend to delegate this crucial process to established external bodies (e.g. list of false content flagged by fact-checkers) (Allcott, Gentzkow, & Yu, 2019; Fletcher & Nielsen, 2017; Guess, Nagler, & Tucker, 2019) or adopting narrow definitions (Khan, Khondaker, Iqbal, & Afroz, 2019) or features based detections (Reis, Correia, Murai, Veloso, & Benevenuto, 2019) that are strict or vague enough to be operationalized in an algorithm (in a certain sense another form of delegation).

Both approaches come however with their own well-known limits. Due to the amount of work required to fact-check a single content, the quota of content flagged as false tends to be a small fraction of the overall false content circulating online. The algorithmic approach has also its shortcomings due to the need for narrowing down the definition enough to make it possible to operationalize the concept and minimize false positives. In both cases, there is a high risk of underestimating the real prevalence of existing false content[1].

Along the same line of delegations, it is possible to shift the attention from single content to actors. The idea is that a user/page/account/news outlet that repeatedly published or shared content that was flagged as false, can be deemed as problematic as a whole. This idea, often based on "black lists" provided by professional fact-checkers, streamline the whole process and allow for raw but automatic estimates of the prevalence of content created by problematic actors. While widely used, this approach carry some risks of biased estimates. Bad and malicious actors tend to disappear and reappear quickly in different forms (new accounts, domains, pages, etc) (Bastos & Mercea, 2018). For this reason, black lists risks to become quickly outdated as well.

Beside using black-lists, malicious actors may also be automatically detected (Shu, Zhou, Wang, Zafarani, & Liu, 2019). This approach has been frequently used for bots – agents that, with a varying degree of automatization, communicate on social media. Features such as the account creation date, clearly recognizable account behavior patterns, the absence of customization such as profile image or covers are often weighed by algorithms aimed at automatically detecting such actors (Yang et al., 2019). Bots have been shown to play active roles in campaigns aimed at artificially boost the reach of certain content (Bessi & Ferrara,

[1] Training the algorithm with content flagged as false by professional fact-checker (supervised machine learning) sounds like a promising compromise. However, even computationally and financially resourceful companies such as Google, Facebook or Twitter are still experimenting with this approach when it comes to misinformation.

2016; Howard, Woolley, & Calo, 2018). While offering better results than automatic means to identify false content, bot detection is also far from perfect (Morstatter, Wu, Nazer, Carley, & Liu, 2016). Sophisticated bots may be difficult and sometimes impossible to detect (Luceri, Deb, Giordano, & Ferrara, 2019).

Media manipulation and false content

To make things even more tricky, creating and distributing false content only represents the tip of the iceberg of the strategies employed by malicious actors to manipulate social media, mainstream media and the public debate (Marwick & Lewis, 2017). Sometimes, even a suitably titled legitimate news stories, opportunely and artificially amplified, can be weaponized to skew the public narrative around certain issues. Both false and real content benefit from a multitude of actors that amplify (intentionally or not) its reach. Depending on the popularity of each actor in the network and the budget they can invest in social media ads, the magnitude of this amplification may change drastically. Furthermore, popular content tends to spread faster on social media due to the effect of algorithms that prioritize better-performing links, images, videos, and posts. These performances depend on an estimate of popularity based on the analysis of quantified attention metrics provided by each platform (likes, reactions, retweets, views, shares, etc). Beside the effect of this "rich will get richer" feedback loop, popular social media content and highly discussed topics are often featured in traditional media, thus benefiting from a significant further spin. The centrality of these metrics offers big rewards to those interested in increasing the visibility of certain content (Y. Zhang, Wells, Wang, & Rohe, 2017). For these reasons, different actors may attempt to coordinate their efforts to get the initial spin which, once detected by the algorithm, may ignite the propagation machine and even attract the attention of mainstream

media (Phillips, 2018). This is not at all a new phenomenon. Fans' attempts to coordinate their behavior to push certain hashtags into Twitter trending topics date back to 2011 at least (boyd, 2017). During the last few years, we observed similar practices applied with the aim of enhancing the spread of political news stories. The practice of *hacking the attention economy* can be driven by a range of motivations from ideology to commercial, to gain status or attention or simply for fun (Marwick & Lewis, 2017; J. Zhang, Carpenter, & Ko, 2013).

Similar campaigns can be identified by looking at the veracity of content or actors involved. For this reason, in this paper we argue for a broader ecological approach that primarily takes into account the collective behavior of malicious actors. While it is in fact perfectly possible that problematic content are published and distributed without any form of attempts of amplifying its reach, it is highly probable that the spread of harmful content are supported by these operations. Detecting the coordinated attempts of multiple actors to increase the visibility of certain content may thus lead to identify networks of potentially inauthentic actors aimed at amplify problematic content. Following a terminology introduced by Facebook, we describe this type of operations as "Coordinated Inauthentic Behavior". "Coordinated Inauthentic Behavior" have been defined in a brief two-minutes explanatory video by Nathaniel Gleicher, Head of Cybersecurity Policy of Facebook as a case when "groups of pages or people work together to mislead others about who they are or what they are doing" (Gleicher, 2018). By shifting the attention to deceptive behaviors, the definition deliberately avoids to fall in the trap of judging the truthfulness of content: "The posts themselves may not be false". In the same video, Gleicher also provides an example: "We may take a network down for making it look like it's being run from one part of the world when in fact it's being run from another. This could be done for ideological purposes or can be financially motivated." Beside the operations undertaken by foreign or local governments,

the policy also applies to "non-state actors, domestic groups and commercial companies" (Gleicher, 2019).

In other terms, the definition comprises of coordination and inauthenticity. Both concepts have been widely studied, albeit rarely in conjunction. In the next paragraphs, we summarize these studies with the aim of grounding the definition of coordinated inauthentic behavior in the existing literature.

Coordination

Coordination can be defined as the act of making people and/or things involved in an organized cooperation. Several authors argued that it is a distinctive mark of users' participation within online spaces (Bruns, Highfield, & Burgess, 2013; Jenkins, 2006; Rotman et al., 2011; Shirky, 2008). Such coordination plays a key role in the online participatory culture described by Henry Jenkins in "Convergence Culture" (2008). Online fandom, for instance, proved to be capable to organize collective actions with different purposes, as inflate social media attention metrics (likes, retweets, etc.) on a specific topic or to influence the plot of a narrative or the trade of an item.

Online activism benefited from the opportunity of building online communities and coordinating their collective actions allowed by the Internet (Bennett & Segerberg, 2012). While most of the early accounts and scholarly work focuses on the beneficial outcomes of digital mediated forms of collaboration as they empower protest movements to fight established and sometimes oppressive powers (Coleman, 2015; Freelon, McIlwain, & Clark, 2018; Loader & Mercea, 2011), the same infrastructure and organization techniques can be employed by a range of diversely motivated malicious actors (Jenkins, Ito, & boyd, 2015; Marwick & Lewis, 2017).

Authenticity

While authenticity has become an increasingly relevant topic for social media companies (Salisbury & Pooley, 2017), malicious actors used whatever websites and social media opportunity to propagate ideas while hiding their real identities and intentions (Bastos & Farkas, 2019; Daniels, 2009; Donovan & Friedberg, 2019). Several scholars provided a range of examples of these activities, from anti-abortion sites masked under the pro-choice tag (Daniels, 2014) to false Islamist Facebook pages spreading anti-muslims content (Farkas, Schou, & Neumayer, 2018).

Besides such "cloaked websites", a well-known type of inauthentic online behavior is that of bots and fake accounts, key tool for spreading computational propaganda (Woolley & Howard, 2016). Bots are widely exploited to manipulate online political discussion and boost politicians' followers to generate false impressions of popularity (Bastos & Mercea, 2017; Bessi & Ferrara, 2016; Ratkiewicz, Conover, Meiss, Flammini, & Menczer, 2011; Woolley & Howard, 2016). Paid users are also employed to impersonate fake social media accounts to undermine online public discourse and distract the public from controversial issues (King, Pan, & Roberts, 2017).

In the seminal work "The people's choice", Lazarsfeld and colleagues (1944) inquieried the role played by personal influence (exposure to casual conversations about politics as opposed to the role played by mass media) on the formation of political opinions, finding that personal influence, compared with traditional media, is able to reach more frequently undecided voters and catch the audience less prepared against influence. Given the effect of accidental exposure to political content on social media on online participation (Valeriani & Vaccari, 2016), malicious social media entities aimed at influencing political opinion may have strong incentive to do so without revealing their authentic motivation and

identity. Furthermore, exploiting the Internet opportunity to gather together people based on personal interests (Ito et al., 2010), it is much easier to build a large follower base by presenting, in order to appeal to a wider audience, the entity as dedicated to entertainment or popular culture than politics. Once the follower base is established, the pages and groups can be used to convey political content to a largely unguarded audience.

Research Questions

Despite the approach focused on coordination and authenticity suits existing literature and is currently employed by major social media companies as a policy to fight information operations that seek to manipulate public debate and, in turn, remove network of actors (accounts, pages and groups) behind such operations, there is a shortage of scholarly evidence on the effectiveness of this approach in terms of surfacing malicious actors and problematic information. To address this gap, we put the idea to test it by analyzing Facebook shares of political news stories published in the run up of two Italian elections. Using an original method described in the next section, we detected several networks of coordinated and inauthentic actors that cooperated to boost certain political news stories in the lead up of both 2018 and 2019 elections. We thus formulated the following research questions:

RQ1: Did these coordinated networks share problematic content in the months preceding 2018 and 2019 Italian elections?

The evidence available in the literature clearly describe a range of motivations pushing actors to coordinate their activities to artificially boost the popularity of certain online content. Given this different range of motivations, we expect that networks entirely composed by

openly political entities (pages, groups and verified profiles belonging to political actors and/or presenting themselves as a venue to get information and discuss politics) and networks also composed, instead, by inauthentic entities that shared political contents under a misleading non political identity, would differ in terms of typology of content shared and structure of the network. We thus formulated the following research questions:

RQ2a: Did political and non-political coordinated networks employ different link sharing strategies?

Considering that existing research on online coordinated information spreading (Del Vicario et al., 2016) have suggested that specific network configurations might be more effective, and thus preferable to achieve a broader dissemination of content, we finally settled on analysing the structure of the coordinated networks. Based on these studies it was asked:

RQ2b: Are there significant structural differences between political and non-political coordinated networks?

Data and Methods

The analyses presented in this paper are based on two datasets of online Italian political news stories shared on Facebook during the six months preceding the 2018 Italian general election (N = 84,815) and the 2019 European election (N = 164,760). For both the elections, news items were collected in real-time using a technological infrastructure based on the open-source software Huginn[2] from three sources: Google News, the Global Database of Society (GDELT) and Twitter (filtering for tweets including a link and mention of a candidate or a political party).

CrowdTangle API link endpoint (CrowdTangle Team, 2019) was used to collect public Facebook/Instagram shares of the news stories URLs in our datasets performed in a period of seven days after the publication of each piece of news. CrowdTangle is a social media analytics tool owned by Facebook that tracks most of the public posts on Facebook, Instagram and Reddit. The numbers shown by this tool reflect public interactions (likes, reactions, comments, shares, upvotes and three second views), with the exception of reach, referral traffic and data around posts originally created as paid ads or made visible only to specific groups of followers[3]. The resulting datasets consisted of 107,842 shares performed by 6,217 unique entities (2018 election dataset) and 222,877 shares performed by 8,148 unique entities (2019 election dataset).

The detection of the networks of coordinated entities was designed as a two steps process (Fig 1). First, the algorithm[4] estimates a time threshold for identifying all the news items shared near simultaneously by different entities in a short period of time. Subsequently,

[2] Huginn is "a system for building agents that perform automated tasks": https://github.com/huginn/huginn.
[3] Please see https://help.crowdtangle.com/en/articles/1140930-what-is-crowdtangle-tracking for an overview of what CrowdTangle is tracking. For this study only Facebook and Instagram platforms have been used.
[4] The algorithm is developed in R and the code is available at https://github.com/fabiogiglietto/coord_link_share_ct.

the coordinated networks are identified by grouping just the entities that repeatedly shared the same news story near simultaneously.

INSERT FIG 1 HERE

While it is ordinary that several entities share the same URLs, the rationale of the method is that it is unlikely that this occurs within a very short time span and repeatedly. Such rapidity and regularity in sharing news items can be a signal of coordinated activity.

The idea is thus to operationalize, as a first step, the concept of "near-simultaneous sharing" by finding an appropriate time threshold. Given a CrowdTangle dataset of URLs shares, this threshold is estimated by analyzing the time differences between each share of the same URL ranked by date (i.e. the date-time when the links were shared) to identify a subset consisting of 10% URLs with the shortest time span between the first and second share, based on the assumption that quickness is necessary for online actors to occupy the social media space. We then identified the desired threshold by calculating the median time in seconds used by 10% of the quickest URLs to reach 50% of their total number of shares, assuming that networks aimed at spreading news items are likely to be closely associated to the news sources they spread. We used this threshold to identify a list of entities that performed "near-simultaneous link sharing".

Since a regular pattern of activity is a significant signal of the existence of an organized structure aimed at spreading news articles on social media, as a second step, we derived from the list of entities resulting from the previous step the networks of the entities that frequently (above the 90th percentile or more than 4 times for 2018 and more than 3 times for 2019) shared news links in a coordinated way.

By using this method, a total of 24 and 92 strongly coordinated networks which spread political news before the 2018 and 2019 elections, respectively, were identified. The 2018

networks were composed by 82 entities, while the 2019 networks by 606 entities. Given the conservative approach used in estimating the "near-simultaneous shares" threshold, the entities listed should be considered as the core of potentially larger networks.

The analyses focused on the news stories shared by these highly coordinated networks within a very short time from each other, that is 2,213 news items shared in the 2018 election dataset, and 5,863 in the 2019 election dataset, also comparing them with the news items shared by the non coordinated entities in both years, which were 38,233 in the 2018 dataset and 66,810 in the 2019 dataset. There is a small overlap in the news items shared by coordinated and non coordinated entities: in the 2018 election dataset about 3% of all the news items shared by non coordinated pages and groups were shared also by coordinated ones, while in the 2019 election dataset the overlap amounts to about 6%.

INSERT TAB 1 HERE

To answer the first research question about whether the coordinated networks identified through the method described above actually spread problematic content online, we checked the domains they shared against blacklists of already identified sources of "fake" and hyperpartisan news. The list of Italian problematic websites was retrieved from established debunking websites which were already used for the same purpose before. Merging the blacklists published by these websites we ended up with a list of 332 problematic news domains. Moreover, we checked the coordinated entities against a list of 87 Facebook pages already pointed out as sources of problematic information by the nonprofit organization Avaaz (Di Benedetto Montaccini, 2019; Mastinu, 2019).

To answer the second research question concerning the differences in terms of link sharing strategies, the analysis focused on the degree of politicalness of the coordinated networks and the variety of sources they shared. First, it was performed a qualitative analysis

of the profiles and cover photos of the coordinated entities, so as to understand their self-presentation strategy and classify them as "political", when their politicalness was explicit, or "non-political", when their self-presentation did not include any reference to politics. Then, a measure of politicalness ranging from 0 to 1 was computed for each network based on the proportion of openly political entities over the total entities of a network. Afterward, it was measured how much large or narrow was the set of domains shared by each coordinated network. To this end, it was computed the Gini coefficient on the proportions of unique domains they shared. Since the Gini coefficient is not computable on only one value, the networks that shared only a single domain were assigned value 1.

Considered the skewness of the data, the statistical analyses were based on non-parametric techniques, such as Spearman's rank correlation coefficient to calculate correlations. The chi-square test was used to evaluate statistical relations between categorical variables and the odds ratio to measure the strength of these relations. The analyses were performed with R (R Core Team, 2013).

To answer the third research question on the structure of the coordinated networks, the analysis of the pages performing coordinated sharing activity was focused on examining strongly coordinated networks to investigate if the political nature of the networks (politicalness) or their editorial strategy, measured through the Gini coefficient, are more frequently associated with a specific type of structure. We looked at network structures through two specific metrics: degree centralization and clustering coefficient. Degree centralization is a widely used metric of degree distribution concentration (Butts, 2006; Wasserman & Faust, 1994) and it has been observed as a measure for authoritarian structures where the opinion of a central node is imposed to, and shared with, external satellites (Sicilia, Korfiatis, Poulos, & Bokos, 2006). We used the value of degree centralization to measure

how much the observed co-sharing network was structured like a star-like network with a clear center of origin. Clustering coefficient (Watts & Strogatz, 1998) measures the degree to which nodes in a network tend to cluster together forming triangles and it has been often associated with the presence of strong community structures (Girvan & Newman, 2002; Rossi & Giglietto, 2016). We used clustering coefficient to measure how the observed networks were densely connected communities.

Given the nature of the chosen metrics, the analysis could only be performed on networks counting more than two entities. This meant that we removed 73 networks that were composed of only two nodes. Figure 2 shows the density functions of the size of the networks showing how dyads of only two nodes were, by far, the most common size.

INSERT FIG 2 HERE

Findings

Statistical significant relations emerged between coordinated activity and the problematicity of both the domains and the Facebook entities (pages or groups) that shared their stories.

In the 2018 election dataset, news items shared in a coordinated way were published by problematic domains significantly more frequently (39%) than those shared without coordinated activity (5%), χ^2 (1, N = 107,842) = 12,529, $p < 0.001$, odds ratio 12.07 (95% CI [11.46, 12.72]). The same, although weaker, relation emerged before the 2019 elections, when the news articles shared in a coordinated way were published by problematic news sources slightly more frequently (4.87%) than those shared without coordination (4.53%), χ^2 (1, N = 222,877) = 10.879, $p < 0.001$, odds ratio 1.08 (95% CI [1.03, 1.13]).

We also observed a strong relation between coordinated activities and well-known problematic Facebook pages and groups. Checking the list of coordinated and

non-coordinated entities against the Avaaz list of Italian problematic Facebook pages, it emerged that both in 2018 and 2019 the entities we detected as coordinated were significantly more present than the non coordinated entities. The 2018 coordinated pages and groups occurred in the list of signaled Facebook pages much more frequently (11%) than the non coordinated ones (1%), χ^2 (1, N = 6,217) = 110.3, $p < 0.001$. Computing the odds ratio it emerged that coordinated entities were 21.49 times (95% CI [9.97, 46.32]) more likely to be signaled than the non coordinated entities. The same relation emerged in the 2019 election dataset, where the number of coordinated entities included in the list of problematic pages was larger (6.8%) than that of the non coordinated ones (0.3%), χ^2 (1, N = 8,148) = 298.05, $p < 0.001$. In the 2019 case the coordinated entities were 24.8 times (95% CI [14.67, 41.93]) more likely to be mentioned in the list of problematic Facebook pages and groups than the non coordinated ones.

Based on the abovementioned evidence it was concluded that coordinated entities shared problematic information before the 2018 and 2019 elections in Italy. Also the answer to the second research question about whether political and non political coordinated networks employ different link sharing strategies was affirmative.

As expected, the qualitative inspection of the entities facade pinpointed a certain degree of deception. Although all of the pages in the datasets shared political news stories, some of them did not disclose their political nature but, on the contrary, conceal it under the appearance of venues exclusively devoted to entertainment, soft news stories or gossip. Considering the coordinated networks active before the 2018 election, 27% of the coordinated networks presented themselves as non political, 29% were composed by openly political and non political entities, and 44% were explicitly political in nature. Examining the coordinated networks active before the 2019 elections, it analogously emerged that 19%

presented themselves as non political, 64% were composed by explicitly political and non political entities, and 17% were composed by plainly political entities.

Besides the issue of deception, the percentage of political entities in a network was also associated with different strategies in terms of the variety of news domains a coordinated network shared. Considering the coordinated networks that spread news items before both the 2018 and 2019 elections in Italy, a strong relation emerged between politicalness of a network and the domains sharing strategies. Indeed, a Spearman correlation found that the more explicital the politicalness of a network, the lower the shares concentration around a few domains, both in the 2018 election dataset, $r_s = -.76$, $N = 24$, $p < 0.001$, and in the 2019 election dataset, $r_s = -.63$, $N = 92$, $p < 0.001$.

The analysis of the network structures associated with the coordinated networks revealed a tendency of the networks to assume either one or the other of the two ideal configurations we have identified: highly clustered or highly centralized networks. Figure 3 represents the density functions of the two metrics measured on the networks and shows how the networks seem to be either organized in one way or another, thus clustering into two groups: one dominated by a centralized structure and one dominated by a clustered structure. Obviously a network can not have both high clustering coefficient and high degree centralization, but a majority of values either on a single side or in the middle ground was entirely possible and it has not been observed.

INSERT FIG. 3 HERE

Figure 4 shows strongly coordinated networks detected in 2018 and 2019 plotted according to their values of centralization and clustering coefficient. It can be observed that networks present various types of structures with the tendency for the networks to assume either one of the two “ideal” structures (highly clustered or highly centralized).

INSERT FIG 4 HERE

We have then explored if there were a correlation between the two structures we identified and specific level of politicalness or editorial strategy (measured through the Gini coefficient), without finding any significant relation. Figure 5 shows the level of politicalness and the Gini coefficient for each strongly coordinated network plotting according to their clustering coefficient and level of centralization. Inspecting Figure 5, emerged that there was no relation between the structure of the strongly coordinated networks and their politicalness or their Gini index. We observed networks dominated by political pages both with a highly centralized structure and highly clustered structure. Similarly, we found highly centralized pages with extremely high Gini index and as well as highly clustered pages. Thus, the observed dichotomy of adopted structure is a finding that requires future works to be fully explained.

INSERT FIG. 5 HERE

Limitations

The algorithm used to detect the "coordinated link sharing behaviour" proved useful to surface subsets with highest concentration of problematic content and actors across the two different datasets of CrowdTangle shares. However, additional tests are needed on a wider range of different datasets of fine tune the algorithm. At the same time, while we tried to carefully avoid arbitrary choices when setting time and edges filters by linking these thresholds to the distributions, a certain amount of arbitrariness proved to be unavoidable.

Entities removed by Facebook as the results of a violation of their policies, disappear from CrowdTangle as well. Given the focus on the analysis of potentially malicious actors, we can not exclude the presence of additional entities or entire coordinated networks at work during both 2018 and 2019 elections. Under this perspective, a public database of the removed entities and the URLs they shared, similar to the one maintained by Twitter (2019), would be helpful for future studies.

Discussion and Conclusion

Given the widely recognized risks posed to democracy by information operations aimed at manipulating the public debate through social media, a wide range of studies attempted to define the phenomenon, estimate its prevalence and effects. However, the complexity of the challenge and the wide variety of motivations and strategies adopted by different malicious actors undermined the attempts to establish a sufficiently shared terminology required to build reliable measures of prevalence and impact. Given this lack of common ground, both content-based and actor-based approaches seem unable to provide compelling answers to the challenges at stake.

In this paper, we introduce an additional approach that focuses on the collective behaviour of the actors. The contribution is twofold. On the one hand, we frame the concept of

"coordinated inauthentic behaviour" in the existing literature on coordination and authenticity. On the other, we assess the reliability of the approach by detecting and analyzing networks of coordinated Facebook entities that boosted political news stories in the lead up of 2018 and 2019 Italian elections.

By analysing over three hundred thousand Facebook shares of thousands political news stories, we identified hundreds of networks of coordinated entities that cooperated to boost a wide variety of sources under a wide variety of political and non political identities.

Entities feeding political content to their subscribers while hiding their identity and intention are particularly distressing. While the scholarly debate about the role played by social media in fostering more selective or cross-cutting exposure flourished, our understanding of the prevalence and effect of this form of casual exposure requires more work (especially concerning cases where this *casual* exposure is in fact orchestrated by malicious actors).

Both the news outlets shared and Facebook coordinated entities detected tend to appear with a frequency well above other news outlets and entities in black-lists compiled by Italian fact-checkers. Furthermore, we show that networks predominantly composed by political entities tend to share a wider variety of news outlets than networks including entities with deceptive non political identities. Certain networks only share one specific domain and this domain is often problematic. In other terms, while political networks tend to share news stories from different sources as long as they support their worldview and even sometimes to shame the alternative worldview, the entire existence of certain non political networks is devoted to boost specific news outlets. While the first group are ideologically motivated, the second are mainly commercially motivated.

Thanks to the comparative perspective offered by two subsequent elections, we also observed several differences that may depend on changes in the strategies adopted by these networks or being the effect of the new policies enforced by Facebook before the EU Parliamentarian 2019 election (Woodford, 2019). The ever changing policies of social media platforms combined with the as much changing adversarial strategies conceived by malicious actors and their networks, pose serious challenges to those who intend to study this phenomenon. At the same time, while we used black-lists to assess the presence of previously known problematic outlets and entities, the comparison between 2018 and 2019 clearly points out that new outlets and entities keep substituting old one making static black-lists partially ineffective.

The analysis of the structural properties of the strongly coordinated networks produced mixed results. On the one side, the structural properties that have been identified, centralization and level of clustering, appears to be relevant since the networks seems to assume one of the two configurations associated with those properties. Nevertheless, our attempt to explain the structures using as explanatory variable the level of politicalness or the Gini index of the networks resulted inconclusive leaving the explanation of the observed duality in structures as a goal for further research.

References

Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, *6*(2), 2053168019848554. https://doi.org/10.1177/2053168019848554

Bastos, M., & Farkas, J. (2019). 'Donald Trump Is My President!': The Internet Research Agency Propaganda Machine. *Social Media + Society*, *5*(3), 2056305119865466. https://doi.org/10.1177/2056305119865466

Bastos, M., & Mercea, D. (2017). The Brexit Botnet and User-Generated Hyperpartisan News. *Social Science Computer Review*. https://doi.org/10.1177/0894439317734157

Bastos, M., & Mercea, D. (2018). The public accountability of social platforms: lessons from a study on bots and trolls in the Brexit campaign. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *376*(2128). https://doi.org/10.1098/rsta.2018.0003

Bennett, W. L., & Segerberg, A. (2012). THE LOGIC OF CONNECTIVE ACTION. *Information, Communication and Society*, *15*(5), 739–768. https://doi.org/10.1080/1369118X.2012.670661

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*, *21*(11). https://doi.org/10.5210/fm.v21i11.7090

boyd, D. (2017). Hacking the attention economy. *Data and Society: Points. Available at: Https://points. Datasociety. Net/hacking-the-Attention-Economy-9fa1daca7a37*. Retrieved from https://points.datasociety.net/hacking-the-attention-economy-9fa1daca7a37

Bruns, A., Highfield, T., & Burgess, J. (2013). The Arab Spring and Social Media Audiences: English and Arabic Twitter Users and Their Networks. *The American Behavioral*

Scientist, *57*(7), 871–898. https://doi.org/10.1177/0002764213479374

Butts, C. T. (2006). Exact bounds for degree centralization. *Social Networks*, *28*(4), 283–296. https://doi.org/10.1016/j.socnet.2005.07.003

Caplan, R., Hanson, L., & Donovan, J. (2018). *Dead Reckoning Navigating Content Moderation After 'Fake News'*. Data&Society.

Coleman, E. G. (2015). *Hacker, hoaxer, whistleblower, spy : the many faces of Anonymous*. Retrieved from https://www.worldcat.org/title/hacker-hoaxer-whistleblower-spy-the-many-faces-of-anonymous/oclc/982185564&referer=brief_results

CrowdTangle Team. (2019). CrowdTangle API. Retrieved 22 October 2019, from CrowdTangle Help website: https://help.crowdtangle.com/en/articles/1189612-crowdtangle-api

Daniels, J. (2009). Cloaked websites: propaganda, cyber-racism and epistemology in the digital era. *New Media & Society*, *11*(5), 659–683. https://doi.org/10.1177/1461444809105345

Daniels, J. (2014). From Crisis Pregnancy Centers to TeenBreaks.com: Anti-abortion Activism's Use of Cloaked Websites. In *Cyberactivism on the Participatory Web* (pp. 152–166). https://doi.org/10.4324/9781315885797-12

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., … Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(3), 554–559. https://doi.org/10.1073/pnas.1517441113

Di Benedetto Montaccini, V. (2019, May 14). Facebook fake news | Elenco di pagine e siti che diffondono notizie false. Retrieved 22 October 2019, from TPI website:

https://www.tpi.it/tecnologia/facebook-fake-news-pagine-20190514313552/

Donovan, J., & Friedberg, B. (2019). *Source Hacking: Media Manipulation in Practice*. Retrieved from Data&Society website: https://datasociety.net/output/source-hacking-media-manipulation-in-practice/

Farkas, J., Schou, J., & Neumayer, C. (2018). Cloaked Facebook pages: Exploring fake Islamist propaganda in social media. *New Media & Society*, *20*(5), 1850–1867. https://doi.org/10.1177/1461444817707759

Fletcher, R., & Nielsen, R. K. (2017). Are News Audiences Increasingly Fragmented? A Cross-National Comparative Analysis of Cross-Platform News Audience Fragmentation and Duplication. *The Journal of Communication*, *67*(4), 476–498. https://doi.org/10.1111/jcom.12315

Freelon, D., McIlwain, C., & Clark, M. (2018). Quantifying the power and consequences of social media protest. *New Media & Society*, *20*(3), 990–1011. https://doi.org/10.1177/1461444816676646

Giglietto, F., Iannelli, L., Valeriani, A., & Rossi, L. (2019). 'Fake news' is the invention of a liar: How false information circulates within the hybrid news system. *Current Sociology. La Sociologie Contemporaine*, *67*(4), 625–642. https://doi.org/10.1177/0011392119837536

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(12), 7821–7826. https://doi.org/10.1073/pnas.122653799

Gleicher, N. (2018, December 6). Coordinated Inauthentic Behavior Explained. Retrieved 19 August 2019, from Facebook Newsroom website: https://newsroom.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/

Gleicher, N. (2019, October 21). How We Respond to Inauthentic Behavior on Our Platforms: Policy Update | Facebook Newsroom. Retrieved 22 October 2019, from https://newsroom.fb.com/news/2019/10/inauthentic-behavior-policy-update/

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, *5*(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

HLEG EU Commission. (2018). *A multi-dimensional approach to disinformation. Report of the independent High level Group on fake news and online disinformation*. Retrieved from EU commission website: https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation

Howard, P. N., Woolley, S., & Calo, R. (2018). Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration. *Journal of Information Technology & Politics*, *15*(2), 81–93. https://doi.org/10.1080/19331681.2018.1448735

Ito, M., Baumer, S., Bittanti, M., boyd, D., Cody, R., Herr-Stephenson, B., … Tripp, L. (2010). *Hanging Out, Messing Around, and Geeking Out: Kids Living and Learning with New Media* (p. 432). Retrieved from http://www.amazon.com/Hanging-Out-Messing-Around-Geeking/dp/0262013363

Jack, C. (2017). *Lexicon of Lies: Terms for Problematic Information*. Retrieved from Data & Society website: https://datasociety.net/output/lexicon-of-lies/

Jenkins, H. (2006). *Fans, Bloggers, and Gamers: Exploring Participatory Culture*. Retrieved from https://play.google.com/store/books/details?id=jj2eKl3NcBEC

Jenkins, H. (2008). *Convergence Culture: Where Old and New Media Collide* (Revised ed).

Retrieved from http://www.amazon.co.uk/dp/0814742955

Jenkins, H., Ito, M., & boyd, D. (2015). *Participatory Culture in a Networked Era: A Conversation on Youth, Learning, Commerce, and Politics*. Retrieved from http://books.google.it/books/about/Participatory_Culture_in_a_Networked_Era.html?hl=&id=3V1XCQAAQBAJ

Khan, J. Y., Khondaker, M. T. I., Iqbal, A., & Afroz, S. (2019). A Benchmark Study on Machine Learning Methods for Fake News Detection. Retrieved from http://arxiv.org/abs/1905.04749

King, G., Pan, J., & Roberts, M. E. (2017). How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument. *The American Political Science Review*, *111*(3), 484–501. https://doi.org/10.1017/S0003055417000144

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The people's choice*. *178*. Retrieved from https://psycnet.apa.org/fulltext/1945-02291-000.pdf

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., … Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Loader, B. D., & Mercea, D. (2011). NETWORKING DEMOCRACY? *Information, Communication and Society*, *14*(6), 757–769. https://doi.org/10.1080/1369118X.2011.592648

Luceri, L., Deb, A., Giordano, S., & Ferrara, E. (2019). Evolution of bot and human behavior during elections. *First Monday*, *24*(9). https://doi.org/10.5210/fm.v24i9.10213

Marwick, A. (2018). WHY DO PEOPLE SHARE FAKE NEWS? A Sociotechnical MODEL OF MEDIA EFFECTS. *Georgetownlawtechreview.org*, *2*(2). Retrieved from https://georgetownlawtechreview.org/wp-content/uploads/2018/07/2.2-Marwick-pp-474-

512.pdf

Marwick, A., & Lewis, R. (2017). *Media Manipulation and Disinformation Online*. Retrieved from Data & Society website: https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

Mastinu, L. (2019, May 14). TPI propone un nuovo elenco di pagine Facebook dispensatrici di odio e falsità. Retrieved 22 October 2019, from Bufale website: https://www.bufale.net/tpi-propone-un-nuovo-elenco-di-pagine-facebook-dispensatrici-di-odio-e-falsita/

Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2019). 'Fake News' Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content. *The American Behavioral Scientist*, *20*, 000276421987822. https://doi.org/10.1177/0002764219878224

Morstatter, F., Wu, L., Nazer, T. H., Carley, K. M., & Liu, H. (2016). A new approach to bot detection: Striking the balance between precision and recall. *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 533–540. https://doi.org/10.1109/ASONAM.2016.7752287

Phillips, W. (2018). *The Oxygen of Amplification. Better Practices for Reporting on Far Right Extremists, Antagonists, and Manipulators*. Data & Society Research Institute.

Ratkiewicz, J., Conover, M. D., Meiss, M., Flammini, A., & Menczer, F. (2011). Detecting and tracking political abuse in social media. *In Proceedings of the 5th AAAI International Conference on Weblogs and Social Media (ICWSM'11*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.646.5073

R Core Team. (2013). *R: A language and environment for statistical computing*. Retrieved from ftp://ftp.uvigo.es/CRAN/web/packages/dplR/vignettes/intro-dplR.pdf

Reis, J. C. S., Correia, A., Murai, F., Veloso, A., & Benevenuto, F. (2019). Supervised Learning for Fake News Detection. *IEEE Intelligent Systems*, *34*(2), 76–81. https://doi.org/10.1109/MIS.2019.2899143

Rossi, L., & Giglietto, F. (2016). Twitter Use During TV: A Full-Season Analysis of #serviziopubblico Hashtag. *Journal of Broadcasting & Electronic Media*, *60*(2), 331–346. https://doi.org/10.1080/08838151.2016.1164162

Rotman, D., Vieweg, S., Yardi, S., Chi, E., Preece, J., Shneiderman, B., … Glaisyer, T. (2011). From Slacktivism to Activism: Participatory Culture in the Age of Social Media. *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, 819–822. https://doi.org/10.1145/1979742.1979543

Salisbury, M., & Pooley, J. D. (2017). The #nofilter Self: The Contest for Authenticity among Social Networking Sites, 2002–2016. *Social Sciences*, *6*(1), 1–24. Retrieved from https://ideas.repec.org/a/gam/jscscx/v6y2017i1p10-d88346.html

Shirky, C. (2008). *Here Comes Everybody: The Power of Organizing Without Organizations*. Retrieved from http://www.librarything.com/work/book/31983348

Shu, K., Zhou, X., Wang, S., Zafarani, R., & Liu, H. (2019). The Role of User Profile for Fake News Detection. Retrieved from http://arxiv.org/abs/1904.13355

Sicilia, M.-A., Korfiatis, N. T., Poulos, M., & Bokos, G. (2006). Evaluating authoritative sources using social networks: an insight from Wikipedia. *Online Information Review*.

Silverman, C. (2017, December 31). I Helped Popularize The Term 'Fake News' And Now I Cringe Whenever I Hear It. Retrieved 28 August 2019, from BuzzFeed News website: https://www.buzzfeednews.com/article/craigsilverman/i-helped-popularize-the-term-fake-news-and-now-i-cringe

Twitter. (2019). Elections integrity. Retrieved 25 October 2019, from Twitter About website:

https://about.twitter.com/en_us/values/elections-integrity.html

Valeriani, A., & Vaccari, C. (2016). Accidental exposure to politics on social media as online participation equalizer in Germany, Italy, and the United Kingdom. *New Media & Society*, *18*(9), 1857–1874. https://doi.org/10.1177/1461444815616223

Wardle, C., & Derakhshan, H. (2017). Information Disorder: Toward an interdisciplinary framework for research and policy making. *Council of Europe Report*, *27*. Retrieved from http://www.theewc.org/content/download/2105/18430/file/INFORMATION%20DISORDER.pdf

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Retrieved from https://play.google.com/store/books/details?id=CAm2DpIqRUIC

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Weeks, B. E., & Gil de Zúñiga, H. (2019). What's Next? Six Observations for the Future of Political Misinformation Research. *The American Behavioral Scientist*, 0002764219878236. https://doi.org/10.1177/0002764219878236

Woodford, A. (2019, April 25). Protecting the EU Elections From Misinformation and Expanding Our Fact-Checking Program to New Languages. Retrieved 24 October 2019, from Facebook Newsroom website: https://newsroom.fb.com/news/2019/04/protecting-eu-elections-from-misinformation/

Woolley, S. C., & Howard, P. N. (2016). Automation, Algorithms, and Politics| Political Communication, Computational Propaganda, and Autonomous Agents — Introduction. *International Journal of Communication Systems*, *10*(0), 9. Retrieved from https://ijoc.org/index.php/ijoc/article/view/6298

Yang, K., Varol, O., Davis, C. A., Ferrara, E., Flammini, A., & Menczer, F. (2019). Arming the public with artificial intelligence to counter social bots. *Human Behavior and Emerging Technologies*, *1*(1), 48–61. https://doi.org/10.1002/hbe2.115

Zhang, J., Carpenter, D., & Ko, M. (2013). *Online Astroturfing: A Theoretical Perspective*. Retrieved from https://aisel.aisnet.org/amcis2013/HumanComputerInteraction/GeneralPresentations/5/

Zhang, Y., Wells, C., Wang, S., & Rohe, K. (2017). Attention and amplification in the hybrid media system: The composition and activity of Donald Trump's Twitter following during the 2016 presidential election. *New Media & Society*, 1461444817744390. https://doi.org/10.1177/1461444817744390