12

INFORMATION SPREADING AND THE ROLE OF AUTOMATED ACCOUNTS ON TWITTER

Two case studies

Guido Caldarelli, Rocco De Nicola, Marinella Petrocchi, and Fabio Saracco

Introduction

Disruptive Innovation: a term that perfectly sums up the impact of social media on people's everyday life. Crucial information can be produced and disseminated among millions of people in a flash, even information concerned with real-time updates on important events. Unfortunately, new technologies have not only revolutionised traditional sectors such as retail and advertising, but they have also been fertile and ground-breaking even on a much more slippery ground: misinformation, hoaxes, and propaganda (National Endowment for Democracy 2017). Disinformation, widely defined as "the purposeful dissemination of false information intended to mislead or harm" (National Endowment for Democracy 2017), is probably as old as human relationships – in the 5th century B.C., the Chinese military theoretician Sun Tzu wrote that "all warfare is based on deception."[1] However, the advent and easiness of use of social media has served to enhance the scale (capability to reach billions of people), scope (capability to achieve a focussed objective, e.g. in terms of a particular audience group), and effectiveness of disinformation (Bradshaw and Howard 2018).

Moreover, the dis/information diffusion on social media is often supported by automated accounts, controlled totally or in part by computer algorithms, called bots. Designed to mimic human behaviour online, a dominant and worrisome use of automated accounts is far from being benign: they have been often used to amplify narratives or drown out political dissent (Yang et al. 2019). Recent studies demonstrate that bots are particularly active in spreading false information. Among other examples, Shao et al. (2018) report of a highly viral fabricated news story titled 'Spirit cooking', which claimed Clinton's campaign chairman practiced bizarre occult rituals, and was published four days before the 2016 US election, subsequently shared in over 30,000 tweets. Even more worryingly, the Global Inventory of Organised Social Media Manipulation reports that bot accounts are being used in 50 out of 70 investigated countries which make use of organised social media manipulation campaigns (Bradshaw and Howard 2019).

According to the 2019 report 'Weapons of mass distraction' (Gangware and Nemr 2019), strategists of false news can exploit – at least – three significant targets of the online information ecosystem: (1) the medium: the platforms on which fake news creeps in and expands; (2) the message: what is the information that one wants to convey?; (3) the audience: who consumes (and contributes to diffuse) this information. The work presented in this chapter focussed on the three mentioned aspects. Relying on two huge Twitter corpora about (1) migration in the Mediterranean Sea from North Africa to Italy, and (2) Covid-19-related discussions, the authors analyse the relevant (i.e. those not compatible with users' random activity) communication and interaction patterns, spotting out the accounts that contribute to the effective dissemination of messages.

Our main results are the following: first, after cleaning the system from the random activity of users, we detect the main hubs of the two networks, i.e. the most effective accounts in significantly propagating their messages, and we observe that those accounts have a higher number of bots among their followers than average. Second, for the migration topic, the strongest hubs in the network share a relatively high number of bots as followers, which most probably aim at further increasing the visibility of the hubs' messages via following and retweeting. As far as the Covid-19 topic is concerned, at least at the time of our investigation, the presence of bots is more limited, but we expect it will grow when the issue stops being only medical and starts becoming political.

To the best of our knowledge, the existence of formations of bots shared by a group of human-operated accounts has never been reported in the literature before.

Datasets

Our study is based on two large corpora of Twitter data, generated by collecting tweets in Italian concerned with migrations and Covid-19-related discussions. For data collection, we developed specific programmes which, by exploiting Twitter public filter API,[2] provided real-time tweet delivery and allowed the collection of sets of data filtered according to specified keywords. For both datasets, we selected a set of keywords compatible with recent chronicles.

The keywords for the dataset about migration have been selected because they are commonly used in Italy when talking and writing about immigration flows from Northern Africa to the Italian coasts, including the dispute about the holder of jurisdiction for handling emergencies involving European countries and NGOs.[3] We collected 1,082,029 tweets, posted by 127,275 unique account IDs over a period of one month (from 23 January to 22 February 2019). By relying on the bot detector classifier developed by Cresci et al. (2015), all the accounts have been classified either as human-operated or as bots. This classification led to 117,879 genuine accounts and 9396 social bots, representing around 7% of all accounts.

It may be worth pointing out that the period over which the data was collected was characterised by a lively political debate in Italy about the landing of the *Diciotti* ship, which was operated by NGOs and rescued migrants fleeing from North Africa to Italy. Rescuing almost 200 migrants on 16 August 2018, it initially received a veto to land from the Italian government; it was allowed to do so only after ten days. Mr. Matteo Salvini, at that time Minister of Internal Affairs, was afterwards investigated for kidnapping and abuse of office; the case was stopped on 19 February 2019, when the Italian Senate did not grant judges the possibility to prosecute him. Right before and after the Senate's decision, there was an intense debate on social networks about migrants and NGOs, and about the role of the Italian government and of the European Union (EU).

The collected tweets concerned with Covid-19 had hashtags related to the coronavirus contagion in the text of the tweet.[4] We collected almost 2.5 million tweets in Italian, from 21 February 2020 to 10 March 2020.[5] By relying on Botometer, the bot detector/classifier developed at the Indiana University (Varol et al. 2017), all the accounts have been classified either as human-operated or as bots. This classification led to 265,910 genuine accounts and 16,973 bots, representing 6% of all accounts. Also in this case, it is important to notice that the timing of the data collection is significant, as far as the topic and the Italian scenario are concerned. In fact, on 21 February the case of the so-called patient one in Lombardy broke out, giving rise to the escalation of the epidemic in Italy.

Users' affiliation

Following previous studies (Becatti et al. 2019; Caldarelli et al. 2020), we used the official certification of one account's authenticity, provided by Twitter, in order to get the polarisation of users. Indeed, upon request from its owner, an account can be certified by the platform and tagged as verified once its authenticity is confirmed. On the official portal, the verified accounts display a blue circle with a white tick at the centre, close to their name.

The intuition behind recent researches on Italian Twitter users' polarisation (Becatti et al. 2019; Caldarelli et al. 2020) is that two verified users are perceived as similar if their messages are retweeted by the same (unverified) users. In order to translate this intuition into a measure, we consider the bipartite network formed by verified users (on one layer) and unverified users (on the other layer). A link is present in the network if one of the two users retweeted the other one at least one time, no matter if the unverified user retweeted the verified one or vice-versa. We chose to focus on retweets since they represent the preferred way through which users spread messages they agree with (Conover et al. 2011).

Figures 12.1 and 12.2 show, respectively, the communities of verified users found for the migration flows scenario and the Covid-19 one. In particular, the network in Figure 12.1 presents a strong community structure. The accounts tied

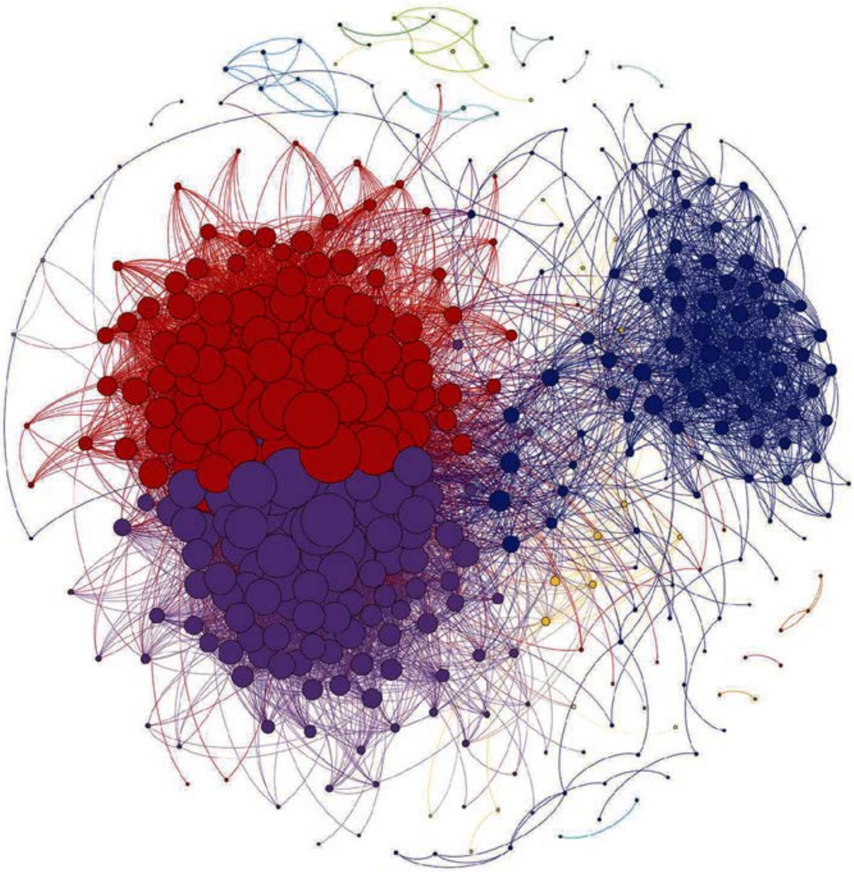

FIGURE 12.1 The communities of verified users in the migration flows dataset

to the Italian government – in office at the time of data collection (Lega and Movimento 5 Stelle) – and other right-wing parties are in blue. The accounts of the centre-left-wing parties (e.g. the Italian Democratic Party, PD) are in red. The violet group includes official media accounts, several NGOs, and left-wing politicians. Some official accounts related to the Catholic Church are in orange. In turquoise, we represent some smaller groups involved in the debate, such as the Maltese Prime Minister Joseph Muscat and some of his ministers, and in green we represent a soccer commentators' community.

As can be seen, the communities are mostly based on their political inclinations. In fact, it is well known that Twitter users tend to be strongly clustered in communities sharing similar ideas (Bessi et al. 2016; Del Vicario et al. 2016; Schmidt et al. 2018). Figure 12.2 shows a pretty clear correlation with respect to political ideas. Orange represents the accounts of the Movimento 5 Stelle. Light blue (on the left) are accounts of Forza Italia. The red vertices are those of the Democratic Party, the institutional users (embassies, police, *carabinieri*, ministers, local governments) are

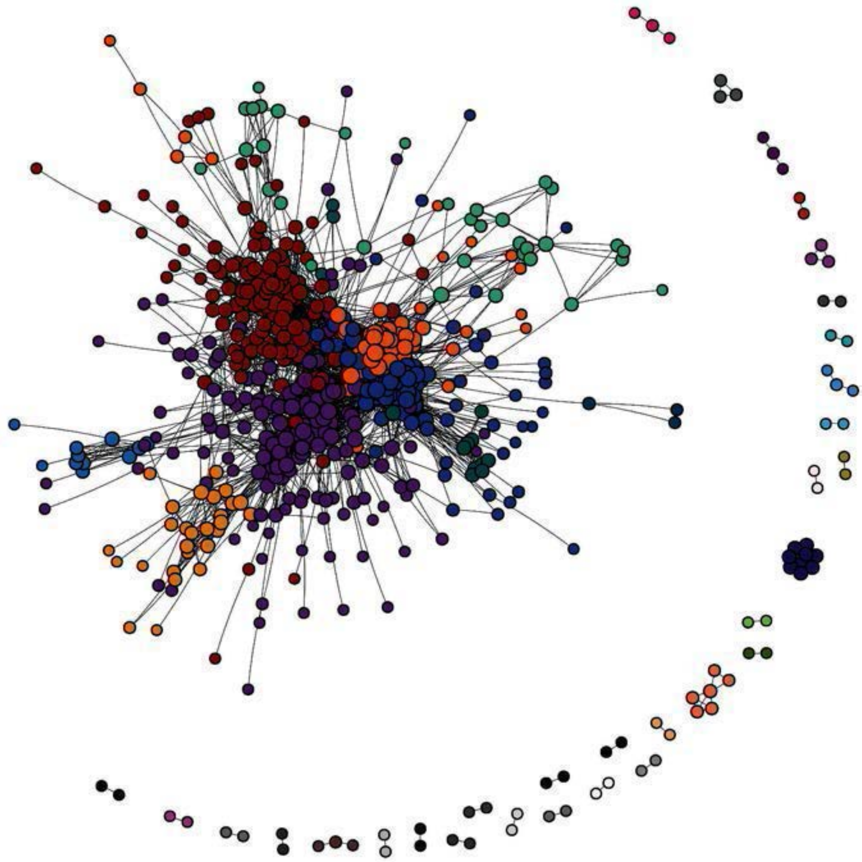

FIGURE 12.2 The communities of verified users in the Covid-19 dataset

represented by blue vertices. Purple vertices are related to Fratelli di Italia, Lega, and newspapers. Finally, the light green vertices on the right are related to TV pundits, journalists, actors, or theatres accounts, while the yellow community on the left is composed mainly by sport journals and journalists.

Verified accounts of politicians can be easily associated with a political party; thus, ideological inclination of unverified users can be guessed by considering their interactions with the communities of verified ones. Figures 12.3 and 12.4 show the matrix of online interactions between verified and unverified users. In Figure 12.3, we report the matrix describing the interactions between verified and unverified users for the migration flows dataset. Nodes are coloured according to their communities, i.e. violet for NGOs, media accounts, left-wing politicians, and for the Democratic Party community, orange for Catholic Church-related accounts, and blue for the pro-government users. In grey, there are users with lower values of polarisation. Figure 12.4 describes the interactions between verified and unverified users for the Covid-19 dataset. Also in this case,

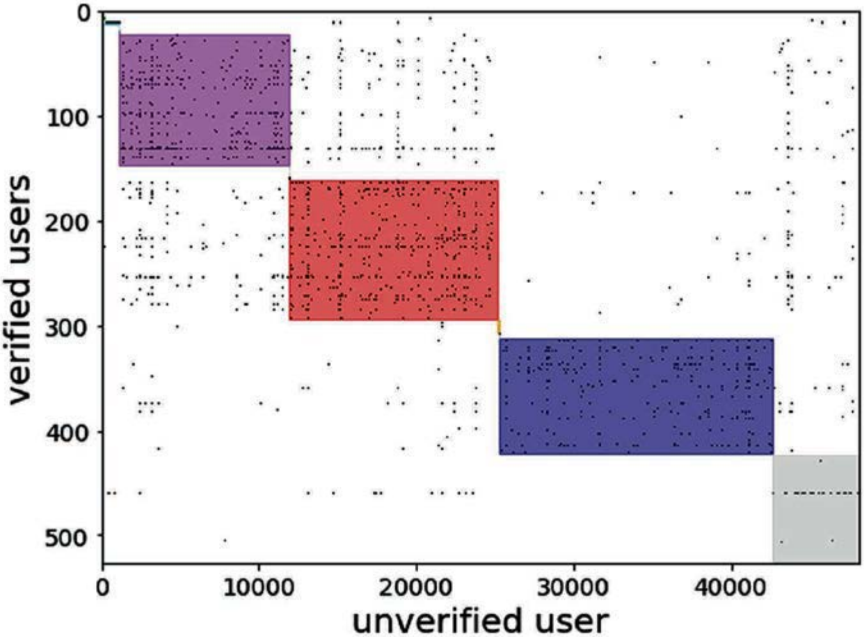

FIGURE 12.3 Interactions between verified and unverified users for the migration flows dataset

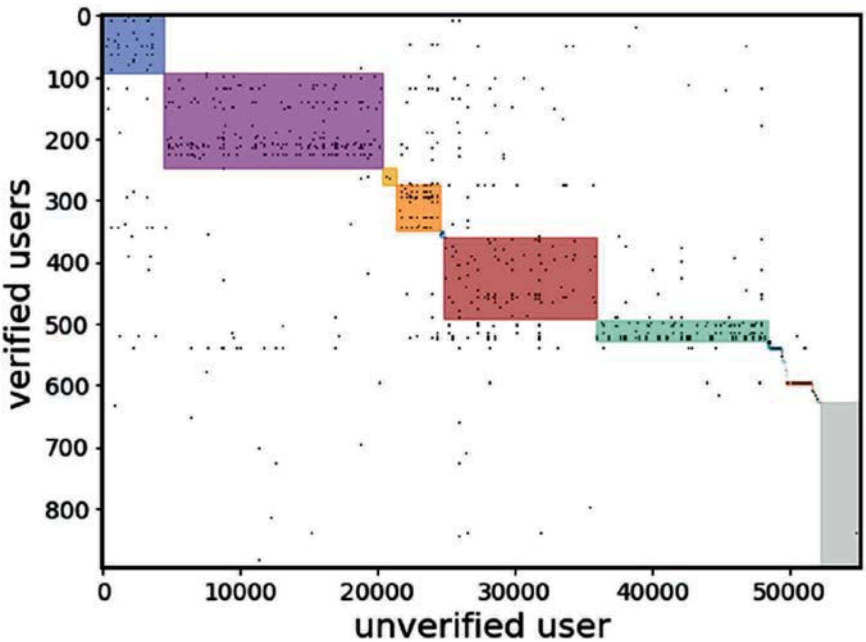

FIGURE 12.4 Interactions between verified and unverified users for the Covid-19 dataset

the community structure clearly reproduces the cluster of verified users with similar political background. Even in a situation of crisis, people tend to follow advice from experts only from a particular side of the political parties (Bessi et al. 2016; Del Vicario et al. 2016). A striking case is that of medical doctor Roberto Burioni, a physician with expertise in infectious diseases (Starr 2020), whose messages are retweeted only by the red community.

Like in a previous study (Becatti at al. 2019), even for the two case studies the community structure is strong.

Significant content exchange

As anticipated in the Introduction, in the analysis of a complex information system, one of the main issues is to skim relevant information from 'noise' (Cimini et al. 2019). Of course, the definition of noise itself depends on the system. In the previous section, we obtained the political affiliation of verified users by projecting the information in the bipartite network describing the interactions between verified and unverified users.

By applying the procedure proposed by van Lidth de Jeude et al. (2019), we filter the total exchange of content in our datasets after discounting the information regarding the activity of users and the virality of messages. Following the approach of Becatti et al. (2019) and Caldarelli et al. (2020), we build the network of users and messages. A link from a user to a message is present if the user authored the message, while there is a link from the message to the user if the latter retweeted the message. This network of users and messages is then used to determine the connections between the users: for every (ordered) couple of users *u* and *w*, we consider how many times *w* retweeted a message authored by *u*, compared to the activity of *u* as an author, the retweeting activity of *w*, and the virality of the messages.

At the end of the procedure, we obtain a 'validated' network: users in such a network contribute to spread the messages in a statistically significant way. The filtering procedure returns a directed network in which the arrows go from the authors to the retweeters. For the migration dataset, the number of nodes reduces to 14,883 users and the number of links reduces to 34,302. For the Covid-19 dataset, the final network contains 10,412 different users and 14,105 links.

Figures 12.5 and 12.6 show the structure of the validated networks in terms of communities for the two scenarios, respectively. The former figure describes the directed validated projection of the retweet activity network. Nodes are violet for NGOs, media accounts, and left-wing politicians, red for the Democratic Party community, and blue for the pro-government users; other colours identify smaller communities. An arrow between a source node and a target node is present if the target is a significant retweeter of the source. The dimension of each node is proportional to its hub score: the biggest node (in blue) is the account of Matteo Salvini, i.e. the leader of a major right-wing party and the Minister of Internal Affairs at the time of the data collection. The latter figure describes the directed

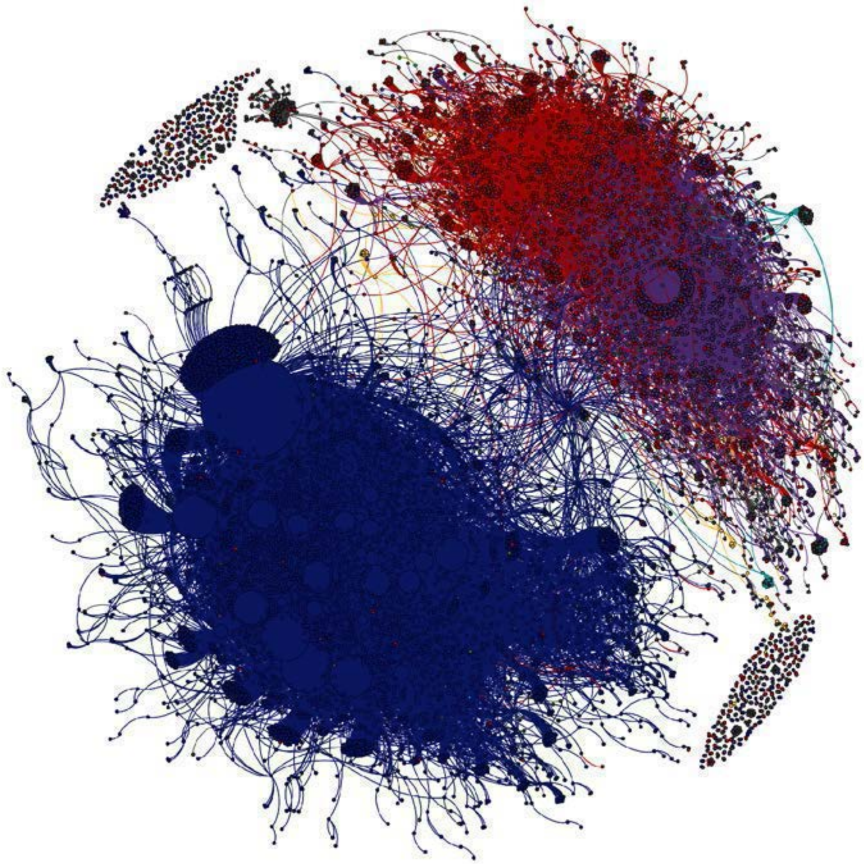

FIGURE 12.5 Mediterranean flows

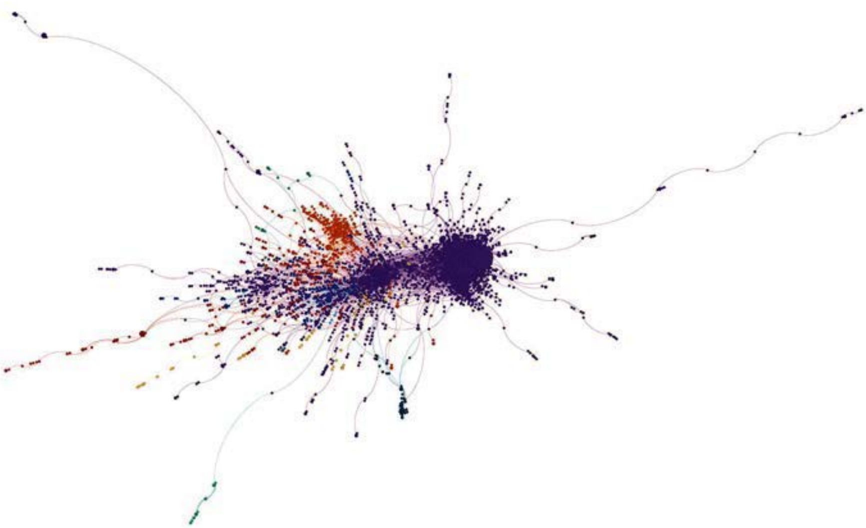

FIGURE 12.6 Covid-19

validated projection of the retweet activity network. In red, the Italian Democratic Party, in orange, accounts of politicians and journalists close to Movimento 5 Stelle. The violet cluster is divided into two poles: the one of media (in the centre) and the one of the extreme right parties (on the right). Interestingly, all communities are extremely linked to the core of the media.

Results

The effectiveness of a hub can be derived by its ability to reach a high number of relevant nodes: this principle is finely implemented in the Hubs-Authorities algorithm, originally introduced in Kleinberg (1999) to rate webpages. In the original version, the paradigm assigns two scores for each webpage: its authority, which estimates the value of the content of the page from the pages linking to it, and its hub value, which estimates the ability to redirect to the most relevant pages. In the scenario currently under investigation, hubs and authorities are Twitter accounts. In the following we will focus on hubs, because they represent the driving force of the discussion and are relatively popular users, and even if they are not verified by Twitter, we often have reliable information about their accounts.

Migration flows scenario

Among the top 20 nodes, in terms of hub scores, the first account is owned by Matteo Salvini, leader of the right-wing party Lega. The second and the third ones belong to two journalists of a news website supported by Casa Pound, a neo-fascist Italian party. The fourth is owned by Giorgia Meloni, leader of the right-wing party Fratelli d'Italia and former ally, during the 2018 Italian electoral campaign, of Lega. Salvini and Meloni have similar opinions on how to deal with migration in the Mediterranean. The fifth and sixth accounts belong, respectively, to a journalist of *Il Fatto Quotidiano* (a newspaper close to M5S) and an unverified user with opinions in line with the ones of the two above mentioned politicians. Notably, the top nodes belong to the blue community. The first account with a different membership (*TgLa*7, a popular newscast by a private TV channel, whose account belongs to the purple community) ranks 176th in the hub score ranking.

Remarkably, we observe a non-zero overlap among the bots in the lists of the validated followers of human-operated accounts. To the best of our knowledge, this is the first time that such a phenomenon is detected. In our opinion, the use of bot squads, retweeting the messages of two or more strong hubs, aims at increasing the visibility of their tweets. We have detected two main groups of such accounts, the other being composed by a maximum of two common bots. The first one includes 22 genuine accounts (nine of which are in the top 10 hubs), sharing 22 bots. In this set, some users share a relatively high fraction of bots; there is one right-wing account that shares all its automated followers with both Meloni and Salvini (see Figure 12.7).

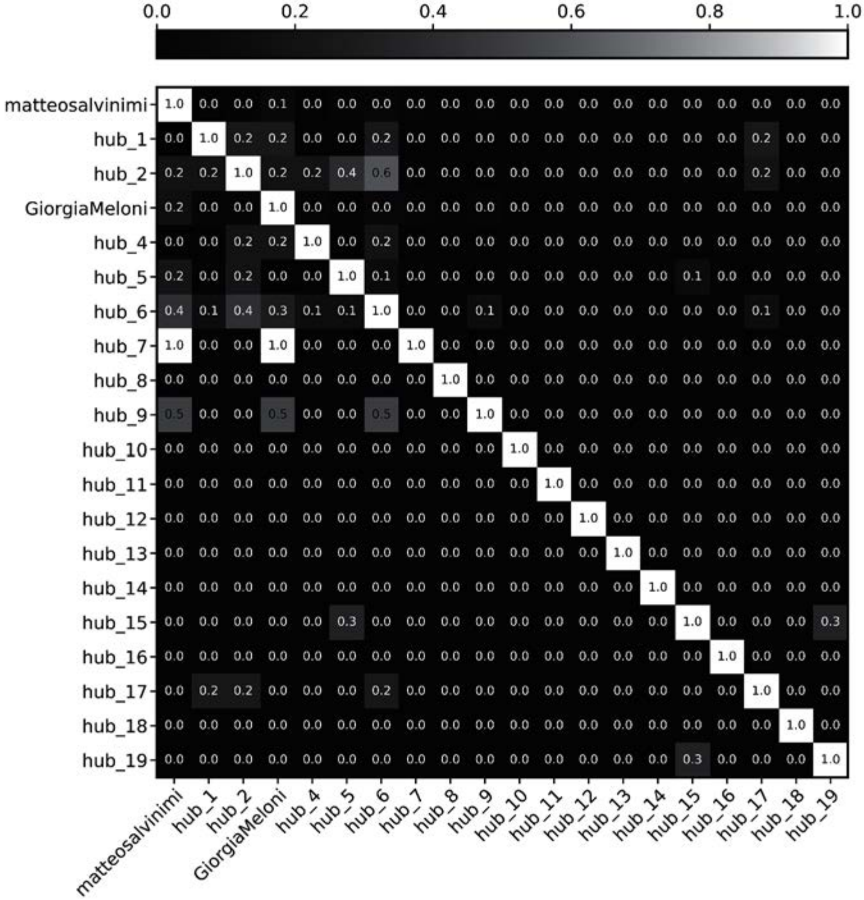

FIGURE 12.7 The relative overlap matrix among the list of bots following the top 20 hubs. A matrix entry represents the percentage of shared bots between users i and j over the number of bots following node i. There are 12 accounts sharing a relatively high number of bots

Figure 12.8 shows the first group of genuine accounts sharing bots and all their bot followers. The subgraph includes genuine accounts (in dark blue) and all the bots following them (in magenta). The dimension of the nodes is proportional to their hub score but normalised on the subgraph. The biggest node represents Salvini's account. In the picture, there are 22 bots shared by 22 humans. Among the latter, nine accounts are among the top 10 hubs. The subgraph contains 172 nodes. Notably, the accounts belong almost exclusively to the blue, i.e., pro-government community. The genuine accounts sharing bots and all their bot followers belong almost exclusively to the blue community, thus we can notice that in this community there is a strong cooperation between bots and humans. The hub scores, represented by the dimensions of the nodes, are nearly homogeneous among the hubs.

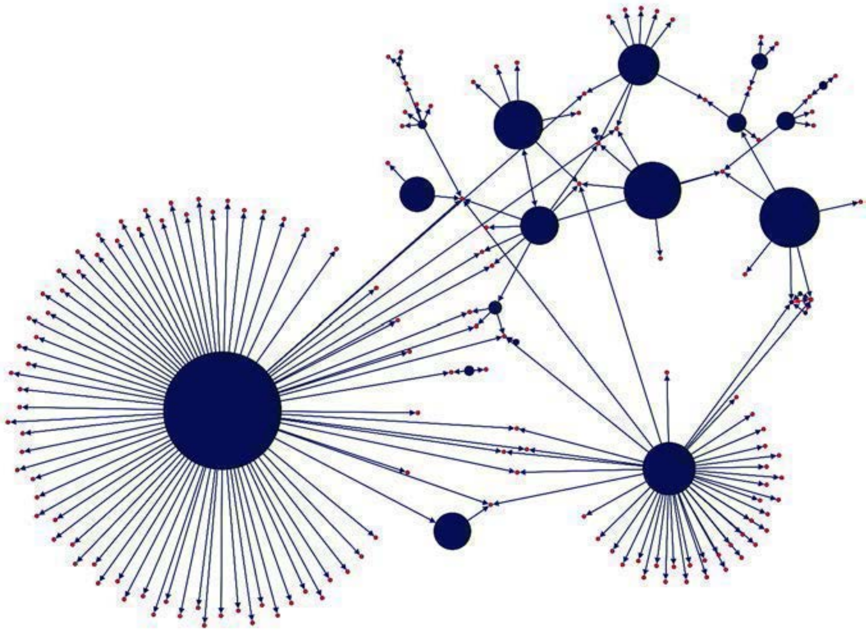

FIGURE 12.8 Subgraph of the largest group of users sharing bots

The incidence of bots in the subgraph of Figure 12.8 is 87%. The number of shared bots over the total number of genuine users is exactly 1. Interestingly, the hubs rarely retweet between each other in a significant way (in fact, only three links can be found among them). They leave it to the bots to spread the content of their partners.

The topmost panel of Figure 12.9 shows that the main activity of the bots in the largest bot squad is retweeting. As expected, they mostly retweet human-operated accounts connected to them (see the central panel of Figure 12.9). The same cannot be said for mentions that may be used either to provoke or to involve the target in a discussion. Accounts from different political sides are mentioned by bot squads; in fact, the bot accounts with more than 30 mentions point to members of the blue community as well as to the official account of the Italian Democratic Party (pdnetwork), a centre-left party. It is worth noticing that other 'non-partisan' verified accounts, e.g. the one of the President of the Republic (Quirinale) and the one of the President of the Chamber of the Deputies, are mentioned there and that, in most cases, the messages containing those mentions are sort of invites for the institutional figures to intervene in the management of migration flows (bottom panel, Figure 12.9). The most striking outcome of the analysis, however, concerns the sources cited by the bots in the blue squads: 89% of their original tweets (i.e. not replies, nor retweets or quoted tweets), contain a URL, and 97% of those URLs refers to www.voxnews.info, a website blacklisted as a source of political disinformation by two popular fact-checking websites, namely www.butac.it and www.bufale.net.

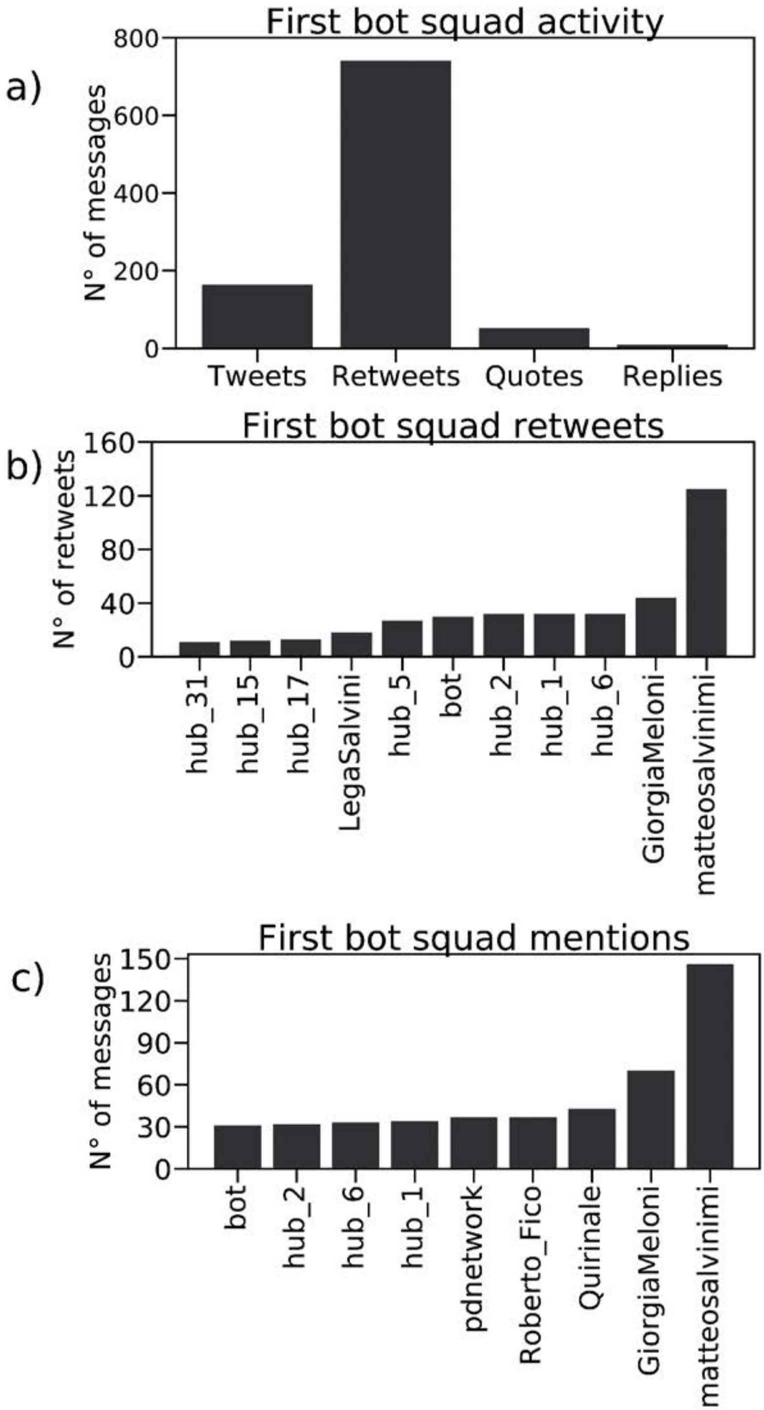

FIGURE 12.9 Statistics of the largest bot squad

Covid-19 scenario

The top 20 hubs are mostly unverified accounts, all from the right wing and the extreme right wing of the political spectrum (exactly as in the above scenario). Among the verified accounts, we have Matteo Salvini, Giorgia Meloni, two right-wing journalists, Salvini's political party, and a politician from the same party; all of them belong to the purple community. In the first position we have an unverified user that we did not spot in the previous analysis, with hub score = 1: this is an unverified account that can be put in relation with extreme right-wing parties. Interestingly enough, in the top 10 hubs, we found four accounts already present

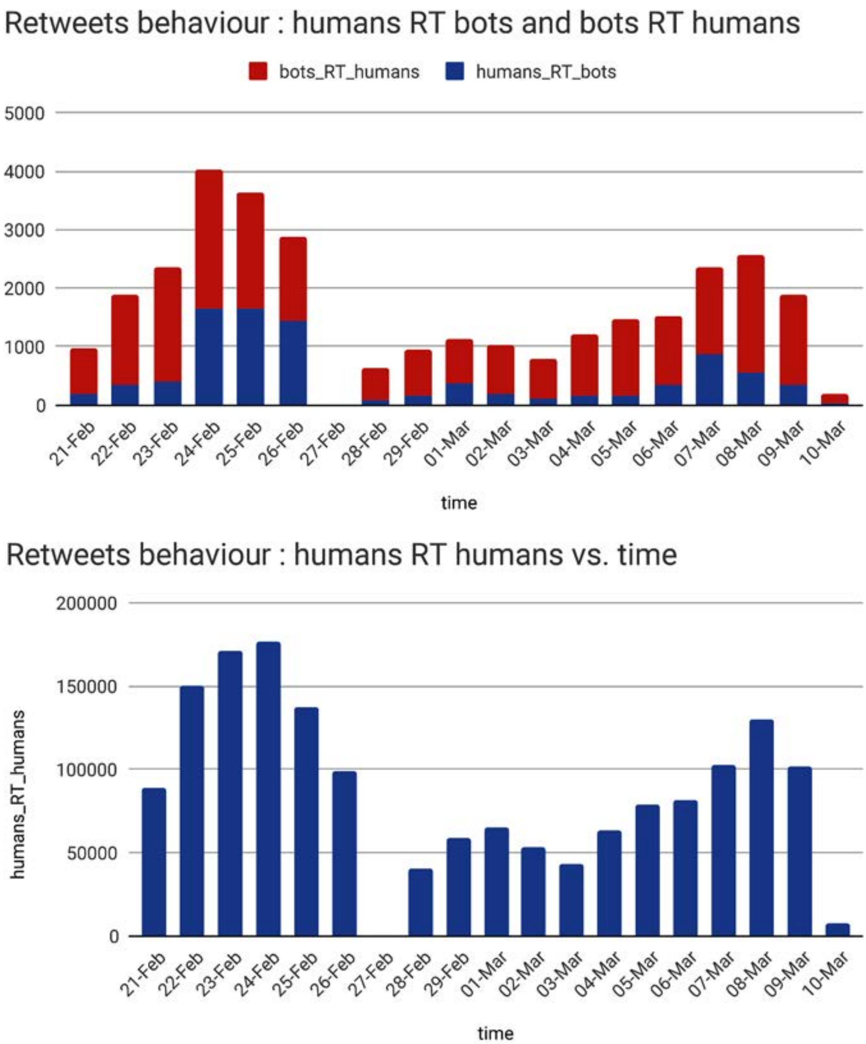

FIGURE 12.10 The retweeting activity of genuine and automated accounts

in the previous study on migration flows. Salvini and Meloni are respectively in the third and fifth places. It should be noticed that the distribution of the hub scores is very peaked (the second place in the ranking accounts for a hub score of 0.45, and it drops down to less than 0.10 for the 10th place).

The activity of bots consists mostly of retweeting human users, although there is a non-negligible activity of genuine accounts in retweeting automated ones (see Figure 12.10). Figure 12.10 shows the retweeting activity of genuine users in the top panel, while the bottom panel reports the the retweeting activity for bots. As it can be seen, the trend of the retweeting activity of automated accounts closely follows the one of genuine users. Looking at the scales, the interaction between human users and bots is limited. Furthermore, the number of genuine users retweeting bots is even smaller.

Amongst these bots, 300 also appeared in the analysis of the political discussion about migration flows. Only 1.7% of all detected bots passed the filtering procedure, thereby representing about 2.8% of validated accounts. Differently from the previous analysis, no bot is shared by the first 500 hubs in the network.

Conclusion

In this chapter, using Twitter as a benchmark, we have analysed the flow of information within and between members of different communities, studying the dynamics of their interactions and the role of automated accounts in such exchanges. For our study, we have relied on techniques developed by physicists and computer scientists. In particular, we have filtered the network of connections and focussed on the most effective accounts in tweets propagation by taking advantage of statistical physics techniques, and we have used machine learning techniques to single out the automatic accounts operating on the network. We have demonstrated the impact of our approach by considering the propagation of Italian tweets concerned with two topics: migration flows in the Mediterranean and Covid-19 related discussion. The analysis has shown that bots play a central role in the exchange of significant content, and that the so-called hub nodes (the most effective accounts in propagating messages) have, among their followers, a high number of bots. This is particularly evident for the migration flows discussion. In the Covid-19 discussion, at least at the time of our investigation, the presence of bots is more limited, but we expect it will grow when the issue stops being only medical and starts becoming political.

The conclusions we can now draw from our analysis do indeed depend on the specific topic of the discussion. As expected, the debate on migrants is dominated by the interplay of several communities strongly related to political parties and alliances operating in the Italian political arena. In this scenario, a rather clear structure of bot squads is present and its effect on the debate is evident. The total presence of bots represents 7% of all users, while their contribution to the backbone of content exchange reduces to the 2.5% of all validated users.

For the Covid-19 discussion, a preliminary analysis showed the presence of echo chambers too. Interestingly enough, those chambers are related to political opinions, even if the subject is mostly scientific. This is not completely unexpected. Indeed, the political countermeasures in order to face the Covid-19 contagion strongly influenced the public debate; by converse, the already present political echo chambers still shape the activity of users on Twitter. The total presence of bots represents 6% of all users, while their contribution to the backbone of content exchange reduces them to 2.8% of all validated users. In contrast to the migration case, we find no signs of a coherent squad of automated accounts acting in the discussion.

Notes

1 Geoff Nunberg (2019) '"Disinformation" is the word of the year – and a sign of what's to come', *NPR*, 30 December 2019.
2 https://developer.twitter.com/en/docs/tweets/filter-realtime/api-reference/post-statuses-filter.html.
3 We searched for the Italian translation of the following keywords: immigrants, migrants, ngo, boat drivers as human smugglers, seawatch, barges, illegal immigrants, Libyan coast guard, shipwreck, disembarkation.
4 #coronavirus, #WuhanCoronavirus, #CoronavirusOutbreak, #coronaviruschina, #coronaviruswuhan, #ChinaCoronaVirus, #nCoV, #coronaviruses, #ChinaWuHan, #nCoV2020, #nCov2019, coronavirus, coronaviruses, ncov, ncov2020, ncov2019, covid2019, covid-19, SARS-CoV2, #SARS_CoV2, #SARSCoV2, #COVID19.
5 We had an interruption of one day and four hours on 27 February due to connection breakdown.

References

Becatti, C., G. Caldarelli, R. Lambiotte, F. Saracco (2019) 'Extracting significant signal of news consumption from social networks: The case of Twitter in Italian political elections', *Palgrave Communications* 5, https://doi.org/10.1057/s41599-019-0300-3.

Bessi, A., F. Zollo, M. Del Vicario, M. Puliga, A. Scala, G. Caldarelli, B. Uzzi, W. Quattrociocchi (2016) 'Users polarization on Facebook and Youtube', *PLoS One* 11(8), doi:10.1371/journal.pone.0159641.

Bradshaw, S., P.N. Howard (2018) 'How does junk news spread so quickly across social media? Algorithms, advertising and exposure in public life', Knight Foundation.

Bradshaw, S., P.N. Howard (2019) 'The global disinformation order: 2019 Global Inventory of Organised Social Media Manipulation', Oxford Internet Institute.

Caldarelli, G., R. De Nicola, F. Del Vigna, M. Petrocchi, F. Saracco (2020) 'The role of bot squads in the political propaganda on Twitter', *Communication Physics* 3(81).

Cimini, G., T. Squartini, F. Saracco, D. Garlaschelli, Andrea Gabrielli & Guido Caldarelli (2019) 'The statistical physics of real-world networks', *Nature Physics Reviews* 1: 52–70.

Conover, M.D., J. Ratkiewicz, M. Francisco, B. Gonçalves, A. Flammini, F. Menczer (2011) 'Political polarization on Twitter', *ICSWM*, doi:10.1021/ja202932e.

Cresci, S., R. Di Pietro, M. Petrocchi, A. Spognardi, M. Tesconi (2015) 'Fame for sale: Efficient detection of fake Twitter followers', *Decision Support Systems* 80: 56–71, doi: 10.1016/j.dss.2015.09.003.

Del Vicario, M., G. Vivaldo, A. Bessi, F. Zollo, A. Scala, G. Caldarelli, W. Quattrociocchi (2016) 'Echo chambers: Emotional contagion and group polarization on Facebook', *Scientific Reports* 6, doi:10.1038/srep37825.

Gangware, C., W. Nemr (2019) 'Weapons of mass distraction: Foreign state-sponsored disinformation in the digital age', www.state.gov/weapons-of-mass-distraction-foreign-state-sponsored-disinformation-in-the-digital-age/.

Kleinberg, J.M. (1999) 'Authoritative sources in a hyperlinked environment', *Journal of the ACM* 46(5), doi:10.1145/324133.324140.

National Endowment for Democracy (2017) 'Issue brief: Distinguishing disinformation from propaganda, misinformation, and "fake news"', https://bit.ly/39E04xb.

Schmidt, A.L., F. Zollo, A. Scala, C. Betsch, W. Quattrociocchi (2018) 'Polarization of the vaccination debate on Facebook', *Vaccine* 36(25): 3606–12.

Shao, C., G.L. Ciampaglia, O. Varol, K.-C. Yang, A. Flammini, F. Menczer (2018) 'The spread of low-credibility content by social bots', *Nat Commun.* 9(1): 4787.

Starr, D. (2020) 'Fighting words', *Science*, 367: 16–19.

van Lidth de Jeude, J., R. Di Clemente, G. Caldarelli, F. Saracco, T. Squartini (2019) 'Reconstructing mesoscale network structures', *Complexity* 2019: 5120581:1–5120581:13.

Varol, O., E. Ferrara, C.A. Davis, F. Menczer, A. Flammini (2017) 'Online human-bot interactions: Detection, estimation, and characterization', *11th International Conference on Web and Social Media*, ICWSM 2017, Montréal, Québec, Canada, May 15–18, pp. 280–89.

Yang, K.-C., O. Varol, C.A. Davis, E. Ferrara, A. Flammini, F. Menczer (2019) 'Arming the public with AI to counter social bots' *Human Behaviour and Emerging Technologies* 1: 48–61.

Acknowledgements

This work has been supported by the Project TOFFEe (TOol for Fighting FakEs) funded by IMT School for Advanced Studies Lucca. The authors would like to thank Serena Giusti and Elisa Piras for the invitation to contribute to the volume and most of all for the feedback that helped us in polishing the paper.