The spread of misinformation by social bots

Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Alessandro Flammini, and Filippo Menczer

Indiana University, Bloomington

Abstract

The massive spread of digital misinformation has been identified as a major global risk and has been alleged to influence elections and threaten democracies. Communication, cognitive, social, and computer scientists are engaged in efforts to study the complex causes for the viral diffusion of misinformation online and to develop solutions, while search and social media platforms are beginning to deploy countermeasures. However, to date, these efforts have been mainly informed by anecdotal evidence rather than systematic data. Here we analyze 14 million messages spreading 400 thousand claims on Twitter during and following the 2016 U.S. presidential campaign and election. We find evidence that social bots play a disproportionate role in spreading and repeating misinformation. Automated accounts are particularly active in amplifying misinformation in the very early spreading moments, before a claim goes viral. Bots target users with many followers through replies and mentions, and may disguise their geographic locations. Humans are vulnerable to this manipulation, retweeting bots who post misinformation. Successful sources of false and misleading claims are heavily supported by social bots. These results suggest that curbing social bots may be an effective strategy for mitigating the spread of online misinformation.

1 Introduction

If you get your news from social media, as most Americans do [9], you are exposed to a daily dose of false or misleading content — hoaxes, conspiracy theories, fabricated reports, click-bait headlines, and even satire. We refer to such claims collectively as "misinformation." The incentives are well understood: traffic to fake news sites is easily monetized through ads [23], but political motives can be equally or more powerful [26, 32]. The massive spread of digital misinformation has been identified as a major global risk [14]. Claims that fake news can influence elections and threaten democracies [10] are hard to prove. Yet we have witnessed abundant demonstrations of real harm caused by misinformation and disinformation spreading on social media, from dangerous health decisions [13] to manipulations of the stock market [8].

A complex mix of cognitive, social, and algorithmic biases contribute to our vulnerability to manipulation by online misinformation [18]. Even in an ideal world where individuals tend to recognize and avoid sharing low-quality information, information overload and finite attention limit the capacity of social media to discriminate information on the basis of quality. As a result, online misinformation is just as likely to go viral as reliable information [30]. Of course, we do not live in such an ideal world. Our online social networks are strongly polarized and segregated along political lines [4, 3]. The resulting "echo chambers" [39, 29] provide selective exposure to news sources, biasing our view of the world [28]. Furthermore, social media platforms are designed to prioritize engaging rather than trustworthy posts. Such algorithmic popularity bias may well hinder the selection of quality content [35, 12, 27]. All of these factors play into confirmation bias and motivated reasoning [37, 17, 19], making the truth hard to discern.

While fake news are not a new phenomenon [21], the online information ecosystem is particularly fertile ground for sowing misinformation. Social media can be easily exploited to manipulate public opinion thanks to the low cost of producing fraudulent websites and high volumes of software-controlled profiles or pages, known as *social bots* [32, 8, 38, 40]. These fake accounts can post content and interact with each other and with legitimate users via social connections, just like real people. People tend to trust social contacts [15] and can be manipulated into believing and spreading content produced in this way [1]. To make matters worse, echo chambers make it easy to tailor misinformation and target those who are most likely to believe it. Moreover, amplification of content through social bots overloads our fact-checking capacity due to our finite attention, as well as our tendencies to attend to what appears popular and to trust information in a social setting [16].

The fight against misinformation requires a grounded assessment of the mechanism by which it spreads online. If the problem is mainly driven by cognitive limitations, we need to invest in news literacy education; if social media platforms are fostering the creation of echo chambers, algorithms can be tweaked to broaden exposure to diverse views; and if malicious bots are responsible for many of the falsehoods, we can focus attention on detecting this kind of abuse. Here we focus on gauging the latter effect. There is plenty of anecdotal evidence that social bots play a role in the spread of misinformation. The earliest manifestations were uncovered in 2010 [26, 32]. Since then, we have seen influential bots affect online debates about vaccination policies [8] and participate actively in political campaigns, both in the U.S. [1] and other countries [45, 7]. However, a quantitative analysis of the effectiveness of misinformation-spreading attacks based on social bots is still missing.

A large-scale, systematic analysis of the spread of misinformation by social bots is now feasible thanks to two tools developed in our lab: the *Hoaxy* platform to track the online spread of claims [36] and the *Botometer* machine learning algorithm to detect social bots [5, 40]. Let us examine how social bots promoted hundreds of thousands of false and misleading articles spreading through millions of Twitter posts during and following the 2016 U.S. presidential campaign.

2 Results

We crawled the articles published by seven independent fact-checking organizations and 121 websites that, according to established media, routinely publish false and/or misleading news (see Methods). The present analysis focuses on the period from mid-May 2016 to the end of March 2017. During this time, we collected 15,053 fact-checking articles and 389,569 unsubstantiated or debunked *claims*. Using the Twitter API, Hoaxy collected 1,133,674 public posts that included links to fact checks and 13,617,425 public posts linking to claims. See Methods for details.

Misinformation sources each produced approximately 100 articles per week, on average. By the end of the study period, the mean popularity of these claims was approximately 30 tweets per article per week (see Supplementary Materials). However, as shown in Fig. 1, success is extremely heterogeneous across articles. Whether we measure success by number of accounts sharing an article or number of posts containing a link, we find a very broad distribution of popularity spanning several orders of magnitude: while the majority of articles go unnoticed, a significant fraction go viral. Unfortunately, and consistent with prior analysis using Facebook data [30], we observe that the popularity profiles of false news are indistinguishable from those of fact-checking articles. Most claims are spread through original tweets and retweets, while few are shared in replies; this is different from fact-checking articles, that are shared mainly via retweets but also replies (Fig. 2).

Fig. 3(a) plots the distribution of the number of tweets used by individual users sharing the same claim. While it is normal behavior for a person to share an article once, the long tail of the distribution highlights inorganic support. A single account posting the same article over and over — hundreds or thousands of times in some cases — is likely controlled by software. We expect the average number of same-article shares per user to decrease for more viral claims, indicating organic spreading. But Fig. 3(b) demonstrates that for the most viral claims, much of the spreading activity originates from a small portion of accounts.

We suspect that these super-spreaders of misinformation are social bots that automatically post links to articles, retweet other accounts, or perform more sophisticated autonomous tasks, like following and replying to other users. To test this hypothesis, we used the Botometer service to evaluate the Twitter accounts that posted links to claims. For each user we computed a bot score, which can be interpreted as the likelihood that the account is controlled by software. Details of the Botometer system can be found in Methods. We considered a random sample of 915 accounts that shared at least one link to a claim. We classified each of these accounts as likely bot or human by comparing its bot score to a threshold of 0.5, which yields high accuracy. Only 8% of accounts in the sample are labeled as likely bots using this method, but they are responsible for spreading 33% of all tweets with links to claims, and 36% of all claims (see details in Supplementary Materials).

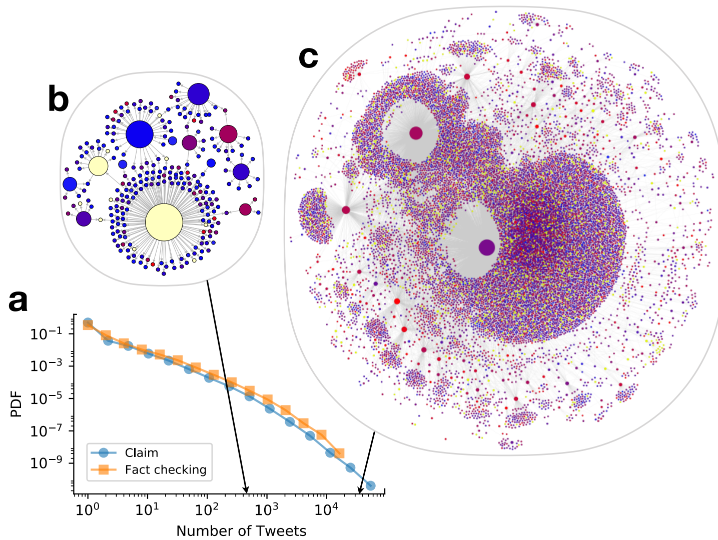

Figure 1: Online virality of content. (a) Probability distribution (density function) of the number of tweets per link, for both claim and fact-checking articles. The distribution of the number of accounts sharing an article is very similar (see Supplementary Materials). As illustrations, the diffusion networks of two claims are shown: (b) a medium-virality article titled *FBI just released the Anthony Weiner warrant, and it proves they stole election,* published a month after the 2016 U.S. election and shared in over 400 tweets; and (c) a highly viral article titled *"Spirit cooking": Clinton campaign chairman practices bizarre occult ritual,* published four days before the 2016 U.S. election and shared in over 30 thousand tweets. In both cases, only the largest connected component of the network is shown. Nodes and links represent Twitter accounts and retweets of the claim, respectively. Node size indicates account influence, measured by the number of times an account is retweeted. Node color represents bot score, from blue (likely human) to red (likely bot); yellow nodes cannot be evaluated because they have either been suspended or deleted all their tweets. An interactive version of the larger network is available online (`iunetsci.github.io/HoaxyBots/`).

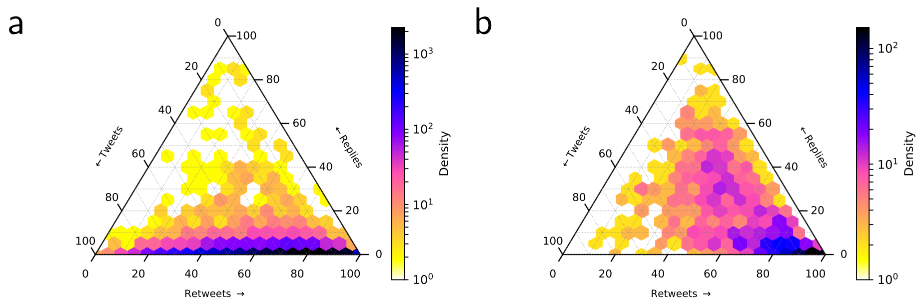

Figure 2: Distribution of types of tweet spreading (a) claims and (b) fact checks. Each article is mapped along three axes representing the percentages of different types of messages that share it: original tweets, retweets, and replies. When user Alice retweets a tweet by user Bob, the tweet is rebroadcast to all of Alice's followers, whereas when she replies to Bob's tweet, the reply is only seen by Bob. Color represents the number of articles in each bin, on a log-scale.

Fig. 4 presents further analysis of the super-spreaders, confirming that they are significantly more likely to be bots compared to the population of users who share claims. We hypothesize that these bots play a critical role in driving the viral spread of misinformation. To test this conjecture, we examined the different spreading phases of viral claims. In each of these phases we examined the accounts posting these claims. As shown in Fig. 5, bots actively share links in the first few seconds after they are first posted. This early intervention exposes many users to false or misleading articles, effectively boosting their viral diffusion.

Another strategy used by bots is illustrated in Fig. 6(a): influential users are often mentioned in tweets that link to misinformation claims. Bots seem to employ this targeting strategy repetitively; for example, a single account produced 18 tweets linking to the claim shown in the figure and mentioning `@realDonaldTrump`. The number of followers of a Twitter user is often used as a proxy for their influence. For a systematic investigation, let us consider all tweets in our corpus that mention or reply to a user and include a link to a viral misinformation story. Tweets tend to mention popular people, of course. However, Figs. 6(b,c) show that when accounts with the highest bot scores share these links, they tend to target users with a higher number of followers (median and average). In this way bots expose influential people, such as journalists and politicians, to a claim, creating the appearance that it is widely shared and the chance that the targeted users will spread it.

We examined whether bots (or rather their programmers) tended to target voters in certain states by creating the appearance of users posting claims from those locations. To this end, we considered accounts with high bot scores that shared claims in the three months before the election, and focused on those with a state location in their profile. The location is self-reported and thus trivial to fake. We compared the distribution of bot account locations across states

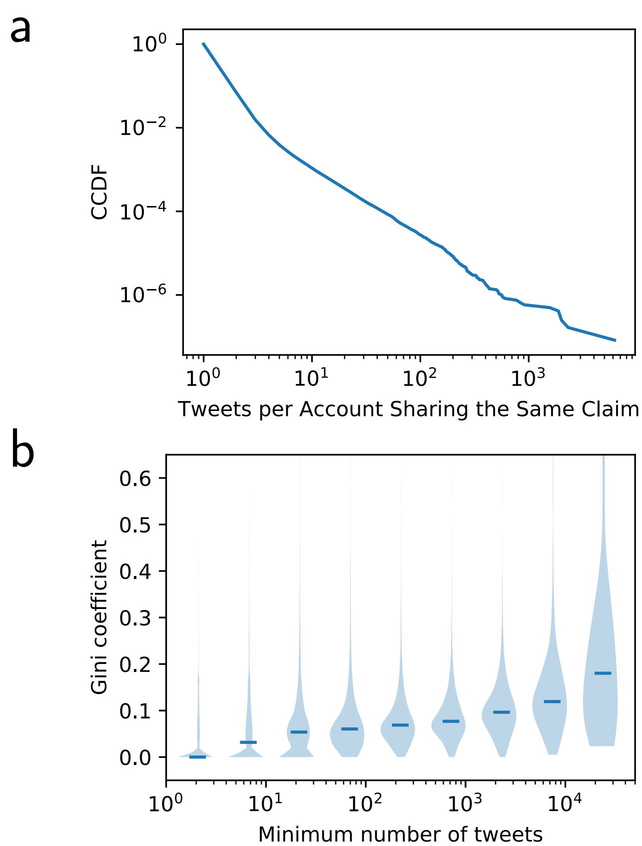

Figure 3: Concentration of claim-posting activity. (a) Cumulative distribution of the number of tweets used by one account to post one and the same claim. (b) Source concentration for claims with different popularity. We consider a collection of articles shared by a minimum number of tweets as a popularity group. We use Gini coefficients to compute source concentration for claims in each of these groups. For each claim, the Lorenz curve plots the cumulative share of tweets versus the cumulative share of accounts generating these tweets. The Gini coefficient is the ratio of the area that lies between the line of equality (diagonal) and the Lorenz curve, over the total area under the line of equality. For claims in each popularity group, a violin plot shows the distribution of Gini coefficients. A high coefficient indicates that a small subset of accounts was responsible for a large portion of the posts. In this and the following violin plots, the width of a contour represents the probability of the corresponding value, and the median is marked by a colored line.

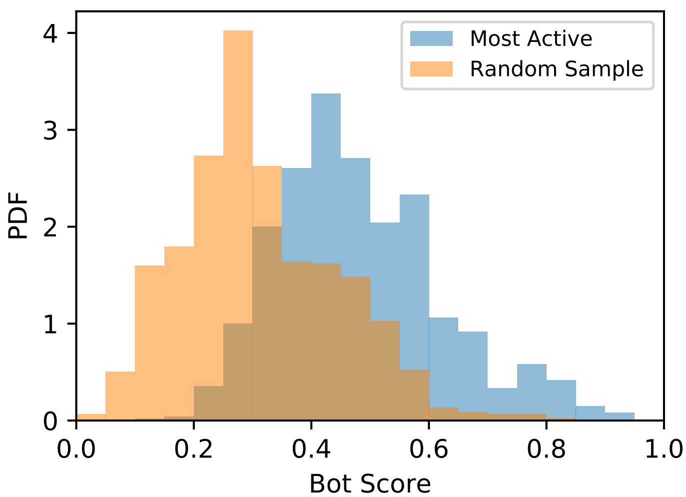

Figure 4: Bot score distributions for a random sample of 915 users who posted at least one link to a claim, and for the 961 accounts that most actively share misinformation (super-spreaders). The two groups have significantly different scores ($p < 10^{-4}$ according to a Mann-Whitney U test). 8% of accounts in the random sample and 38% of accounts in the most active group have bot score above 0.5. Details in Supplementary Materials.

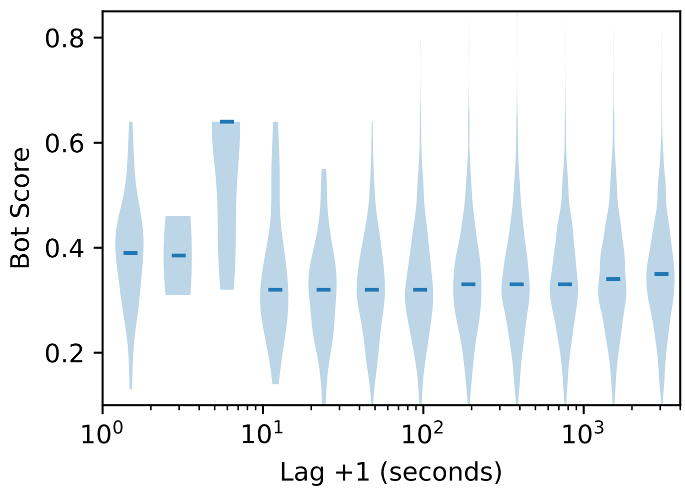

Figure 5: Temporal evolution of bot support after a viral claim is first shared. We consider a sample of 60,000 accounts that participate in the spread of the 1,000 most viral claims. We align the times when each claim first appears. We focus on a one-hour early spreading phase following each of these events, and divide it into logarithmic lag intervals. The plot shows the bot score distribution for accounts sharing the claims during each of these lag intervals.

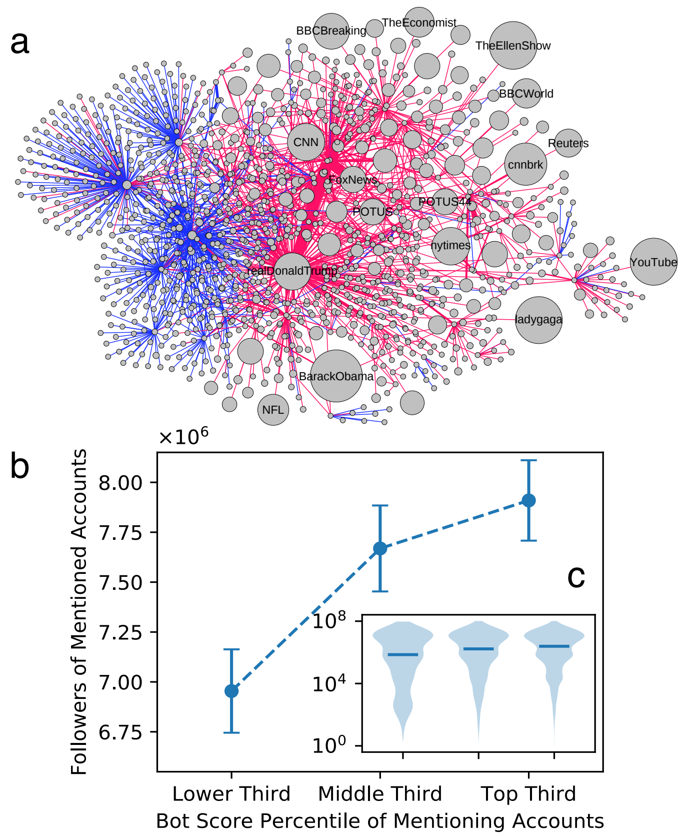

Figure 6: (a) Example of targeting for the claim *Report: three million votes in presidential election cast by illegal aliens*, published by `Infowars.com` on November 14, 2016 and shared over 18 thousand times on Twitter. Only a portion of the diffusion network is shown. Nodes stand for Twitter accounts, with size representing number of followers. Links illustrate how the claim spreads: by retweets and quoted tweets (blue), or by replies and mentions (red). (b) Average number of followers for Twitter users who are mentioned (or replied to) by accounts that link to the most viral 1000 claims. The mentioning accounts are aggregated into three groups by bot score percentile. Error bars indicate standard errors. (c) Distributions of follower counts for users mentioned by accounts in each percentile group.

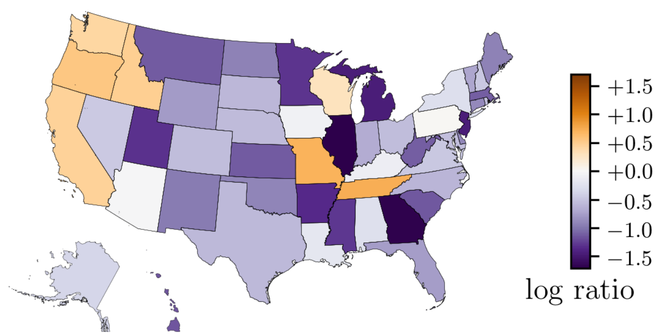

Figure 7: Map of location targeting by misinformation bots, relative to a baseline. To gauge sharing activity by likely bots, we considered tweets posting links to claims by accounts with bot score above 0.6 that reported a U.S. state location in their profile. We compared the tweet frequencies by states with those expected from a large sample of tweets about the elections in the same period. Positive log ratios indicate states with higher than expected bot activity. See Methods for details.

with a baseline obtained from a large sample of tweets about the elections in the same period (see details in Methods). A χ^2 test indicates that the location patterns produced by bots are inconsistent with the geographic distribution of political conversations on Twitter ($p < 10^{-4}$). This suggests that as part of their disguise, social bots are more likely to report certain locations than others. For example, Fig. 7 shows geographic anomalies in Tennessee and Missouri, where bot activity is over five times above baseline.

Having found that bots are employed to drive the viral spread of misinformation, let us explore how humans interact with the content shared by bots, which may provide insight into whether and how bots are able to affect public opinion. Fig. 8 shows that humans do most of the retweeting (upper panel), and they retweet claims posted by bots as much as by other humans (left panel). This suggests that collectively, people do not discriminate between misinformation shared by humans versus social bots.

Finally, we compared the extent to which social bots successfully manipulate the information ecosystem in support of different sources of online misinformation. We considered the most popular sources in terms of median and aggregate article posts, and measured the bot scores of the accounts that most actively spread their claims. As shown in Fig. 9, one site (`beforeitsnews.com`) stands out in terms of manipulation, but other well-known sources also have many bots among their promoters. Satire sites like *The Onion* and fact-checking websites do not display the same level of bot support.

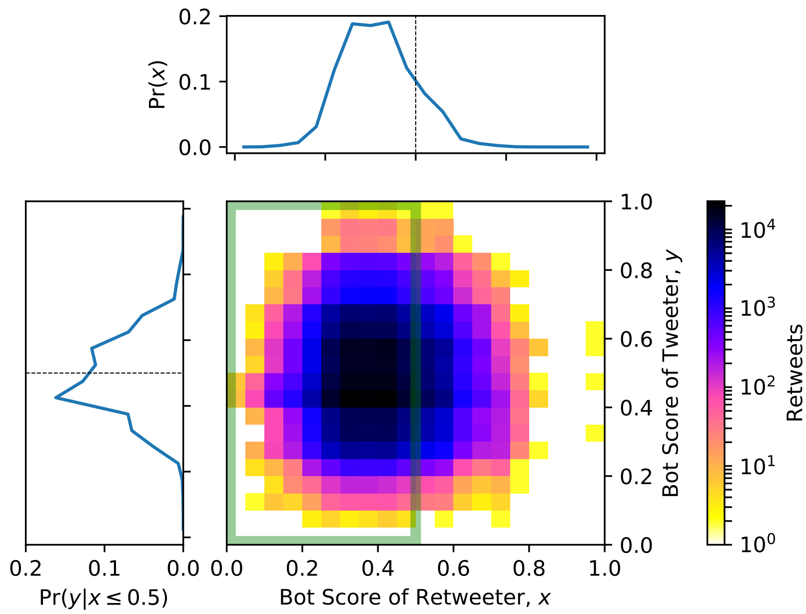

Figure 8: Joint distribution of the bot scores of accounts that retweeted links to claims and accounts that had originally posted the links. Color represents the number of retweeted messages in each bin, on a log scale. Projections show the distributions of bot scores for retweeters (top) and for accounts retweeted by likely humans (left).

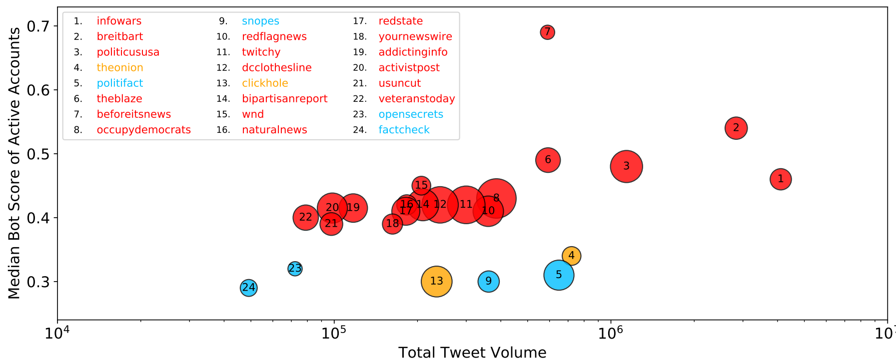

Figure 9: Popularity and bot support for the top claim and fact-checking sources. Top satire websites are shown in orange, top fact-checking sites in blue, and other claim sources in red. Popularity is measured by total tweet volume (horizontal axis) and median number of tweets per claim (circle area). Bot support is gauged by the median bot score of the 100 most active accounts posting links to articles from each source (vertical axis).

3 Discussion

Our analysis provides quantitative empirical evidence of the key role played by social bots in the viral spread of online misinformation. Relatively few accounts are responsible for a large share of the traffic that carries misinformation. These accounts are likely bots, and we uncovered several manipulation strategies they use. First, bots are particularly active in amplifying misinformation in the very early spreading moments, before a claim goes viral. Second, bots target influential users through replies and mentions. Finally, bots may disguise their geographic locations. People are vulnerable to these kinds of manipulation, retweeting bots who post misinformation just as much as they retweet other humans. Successful sources of misinformation in the U.S., including those on both ends of the political spectrum, are heavily supported by social bots. As a result, the virality profiles of false news are indistinguishable from those of fact-checking articles. Social media platforms are beginning to acknowledge these problems and deploy countermeasures, although their effectiveness is hard to evaluate [44, 25].

Our findings demonstrate that social bots are an effective tool to manipulate social media. While the present study focuses on the spread of misinformation, similar bot strategies may be used to spread other types of content, such as propaganda and malware. And although our spreading data is collected from Twitter, there is no reason to believe that the same kind of abuse is not taking place on other digital platforms as well. In fact, viral conspiracy theories spread on Facebook [6] among the followers of pages that, like social bots, can easily be managed automatically and anonymously. Furthermore, just like on Twitter, false claims on Facebook are as likely to go viral as reliable news [30]. While the difficulty to access spreading data on platforms like Facebook is a concern, the growing popularity of ephemeral social media like Snapchat may make future studies of this abuse all but impossible.

The results presented here suggest that curbing social bots may be an effective strategy for mitigating the spread of online misinformation. Progress in this direction may be accelerated through partnerships between social media platforms and academic research. For example, our lab and others are developing machine learning algorithms to detect social bots [8, 38, 40]. The deployment of such tools is fraught with peril, however. While platforms have the right to enforce their terms of service, which forbid impersonation and deception, algorithms do make mistakes. Even a single false-positive error leading to the suspension of a legitimate account may foster valid concerns about censorship. This justifies current human-in-the-loop solutions, which unfortunately do not scale with the volume of abuse that is enabled by software. It is therefore imperative to support research both on improved abuse detection algorithms and on countermeasures that take into account the complex interplay between cognitive and technological factors that favor the spread of misinformation [20].

An alternative strategy would be to employ CAPTCHAs [42], challenge-response tests to determine whether a user is human. CAPTCHAs have been deployed widely and successfully to combat email spam and other types of online abuse. Their use to limit automatic posting or resharing of news links could stem bot abuse, but also add undesirable friction to benign applications of automation by legitimate entities, such as news media and emergency response coordinators. These are hard trade-offs that must be studied carefully as we contemplate ways to address the fake news epidemics.

4 Methods

The online article-sharing data was collected through Hoaxy, an open platform developed at Indiana University to track the spread of misinformation and fact checking on Twitter [36]. A search engine, interactive visualizations, and open-source software are freely available (`hoaxy.iuni.iu.edu`). The data is accessible through a public API.

Our definition of "misinformation" follows the industry convention and includes the following classes: fabricated content, manipulated content, imposter content, false context, misleading content, false connection, and satire [43]. To these seven categories we also add claims that cannot be verified. Since fact-checking and categorizing millions of claims is unfeasible, the links to the stories considered here were crawled from websites that routinely publish these types of unsubstantiated and debunked claims, according to lists compiled by reputable third-party news and fact-checking organizations. We started the collection in mid-May 2016 with 71 sites and added 50 more in mid-December 2016. The collection period for the present analysis extends until the end of March 2017. During this time, we collected 389,569 claims. We also tracked 15,053 stories published by independent fact-checking organizations, such as `snopes.com`, `politifact.com`, and `factcheck.org`. The full list of sources is reported in Supplementary Materials. We did not exclude satire because many fake-news sources label their content as satirical, and viral satire is often mistaken for real news. *The Onion* is the satirical source with the highest total volume of shares. We repeated our analyses of most viral claims (e.g., Fig. 5) with articles from `theonion.com` excluded and the results were not affected. Our source-based analysis of misinformation diffusion does not require a complete list of sources, but does rely on the assumption that the vast majority of claims published by these sources falls under our definition of misinformation or unsubstantiated information. To validate this assumption, we analyzed the content of a random sample of articles, finding that fewer that 15% of claims could be verified. More details are available in Supplementary Materials.

Using Twitter's public streaming API, we collected 13,617,425 public posts that included links to claims and 1,133,674 public posts linking to fact checks. This is the complete set of tweets linking to these claims and fact checks in the study period, rather than a sample. We extracted metadata about the source of each link, the account that shared it, the original poster in case of retweet or quoted tweet, and any users mentioned or replied to in the tweet.

We transformed URLs to their canonical forms to merge different links referring to the same article. This happens mainly due to shortening services (44% links are redirected) and extra parameters (34% of URLs contain analytics tracking parameters), but we also found websites that use duplicate domains and snapshot services. Canonical URLs were obtained by resolving redirection and removing analytics parameters.

The bot score of Twitter accounts is computed using the Botometer service, which evaluates the extent to which an account exhibits similarity to the characteristics of social bots [5]. The system is based on a supervised machine learning algorithm leveraging more than a thousand features extracted from public data and meta-data about Twitter accounts. These features include various descriptors of information diffusion networks, user metadata, friend statistics, temporal patterns of activity, part-of-speech and sentiment analysis. The classifier is trained using publicly available datasets of tens of thousands of Twitter users that include both humans and bots of varying sophistication. The model has high accuracy in discriminating between human and bot accounts of different nature; five-fold cross-validation yields an area under the ROC curve of 94% [40]. The Botometer system is available through a public API (`botometer.iuni.iu.edu`) and is widely adopted, serving over 100 thousand requests daily. For the present analysis we use the Twitter Search API to collect up to 200 of an account's most recent tweets and up to 100 of the most recent tweets mentioning the account. From this data we extract the features used by the classifier to compute the bot score.

In the targeting analysis (Fig. 6), we exclude mentions of sources using the pattern "via @screen_name."

The location analysis is based on 3,971 tweets that meet four conditions: they were shared in the period between August and October 2016, included a link to a claim, originated from an account with high bot score (above 0.6), and included one of the 51 U.S. state names or abbreviations (including District of Columbia) in the location metadata. The baseline frequencies were obtained from a 10% random sample of public posts from the Twitter streaming API. This yielded 164,276 tweets in the same period that included hashtags with the prefix `#election` and a U.S. state location. Though the sample of tweets is unbiased, when one extracts the accounts posting these tweets, active accounts are more likely to be represented. We assume that being active is independent of reporting one's location truthfully, and therefore the baseline distribution of locations is representative of the Twitter population. Both baseline and bot tweet counts are normalized, then we consider the log-ratio in Fig. 7 so that positive (negative) values indicate higher (lower) bot activity than expected.

Acknowledgments. We are grateful to Ben Serrette and Valentin Pentchev of the Indiana University Network Science Institute (`iuni.iu.edu`), as well as Lei Wang for supporting the development of the Hoaxy platform. Clayton A. Davis developed the Botometer API. Nic Dias provided assistance with claim verification. We are also indebted to Twitter for providing data through their API. C.S. thanks the Center for Complex Networks and Systems Research (`cnets.indiana.edu`) for the hospitality during his visit at the Indiana University School of Informatics and Computing. He was supported by the China Scholarship Council. G.L.C. was supported by IUNI. The development of the Botometer platform was supported in part by DARPA (grant W911NF-12-1-0037). A.F. and F.M. were supported in part by the James S. McDonnell Foundation (grant 220020274) and the National Science Foundation (award CCF-1101743). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Appendix: Supplementary Materials

List of sources

Our list of 'misinformation' sources was obtained by merging several lists compiled by third-party news and fact-checking organizations. It should be noted that these lists were compiled independently of each other, and as a result they have uneven coverage. However, there is some overlap between them. The full list of sources is shown in Table 1. We also tracked the websites of seven independent fact-checking organizations: `badsatiretoday.com`, `factcheck.org`, `hoax-slayer.com`,[1] `opensecrets.org`, `politifact.com`, `snopes.com`, and `truthorfiction.com`. In April 2017 we added `climatefeedback.org`, which does not affect the present analysis.

[1] `hoax-slayer.com` includes its older version `hoax-slayer.net`.

Table 1: Misinformation sources. For each source, we indicate which lists include it: Fake News Watch (FNW), opensources.co (MZ), Daily Dot (DD), US News & World Report (USN), New Republic (NR), CBS, Urban Legends at About.com (A), NPR, and Snopes Field Guide. Headers link to the original lists.

Source	**FNW**	**MZ**	**DD**	**USN**	**NR**	**CBS**	**A**	**NPR**	**Snopes**	**Date Added**
21stcenturywire.com	Yes	Yes	Yes	No	No	No	No	No	No	2016-06-29
70news.wordpress.com	No	Yes	Yes	No	No	Yes	No	No	No	2016-12-20
abcnews.com.co	No	No	Yes	No	No	Yes	No	No	No	2016-12-20
activistpost.com	Yes	Yes	Yes	Yes	No	No	No	No	No	2016-06-29
addictinginfo.org	No	Yes	Yes	No	No	No	No	No	No	2016-12-20
americannews.com	Yes	Yes	Yes	Yes	No	No	No	No	No	2016-06-29
americannewsx.com	No	Yes	No	No	No	No	No	No	No	2016-12-20
amplifyingglass.com	Yes	No	No	No	No	No	No	No	No	2016-06-29
anonews.co	No	No	Yes	No	No	No	No	No	No	2016-12-20
beforeitsnews.com	Yes	Yes	No	Yes	No	No	No	No	No	2016-06-29
bigamericannews.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
bipartisanreport.com	No	Yes	Yes	No	No	No	No	No	No	2016-12-20
bluenationreview.com	No	Yes	Yes	No	No	No	No	No	No	2016-12-20
breitbart.com	No	Yes	Yes	No	No	No	No	No	No	2016-12-20
burrardstreetjournal.com	No	No	No	No	No	Yes	No	No	No	2016-12-20
callthecops.net	No	No	Yes	No	No	No	Yes	No	No	2016-12-20
christiantimes.com	No	No	No	No	No	Yes	No	No	No	2016-12-20
christwire.org	Yes	Yes	Yes	No	No	No	No	No	No	2016-06-29
chronicle.su	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
civictribune.com	Yes	Yes	Yes	No	No	Yes	No	No	No	2016-06-29
clickhole.com	Yes	Yes	Yes	Yes	No	No	No	No	No	2016-06-29
coasttocoastam.com	Yes	Yes	Yes	No	No	No	No	No	No	2016-06-29
collective-evolution.com	No	No	Yes	No	No	No	No	No	No	2016-12-20
consciouslifenews.com	Yes	Yes	Yes	No	No	No	No	No	No	2016-06-29
conservativeoutfitters.com	Yes	Yes	Yes	No	No	No	No	No	No	2016-12-20
countdowntozerotime.com	Yes	Yes	Yes	No	No	No	No	No	No	2016-06-29
counterpsyops.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
creambmp.com	Yes	Yes	Yes	No	No	No	No	No	No	2016-06-29
dailybuzzlive.com	Yes	Yes	No	Yes	No	No	No	No	No	2016-06-29
dailycurrant.com	Yes	Yes	No	No	No	No	Yes	No	No	2016-06-29
dailynewsbin.com	No	Yes	No	No	No	No	No	No	No	2016-12-20
dcclothesline.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
demyx.com	No	No	No	No	Yes	No	No	No	No	2016-12-20
denverguardian.com	No	No	No	No	No	No	No	Yes	No	2016-12-20
derfmagazine.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
disclose.tv	Yes	Yes	No	Yes	No	No	No	No	No	2016-06-29
duffelblog.com	Yes	Yes	Yes	Yes	No	No	No	No	No	2016-06-29
duhprogressive.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
empireherald.com	No	Yes	No	No	No	Yes	No	No	No	2016-12-20
empirenews.net	Yes	Yes	Yes	No	No	Yes	Yes	No	Yes	2016-06-29
empiresports.co	Yes	No	No	No	Yes	No	Yes	No	Yes	2016-06-29
en.mediamass.net	Yes	Yes	Yes	No	Yes	No	Yes	No	No	2016-06-29
endingthefed.com	No	Yes	No	No	No	No	No	No	No	2016-12-20
enduringvision.com	Yes	Yes	Yes	No	No	No	No	No	No	2016-06-29
flyheight.com	No	Yes	No	No	No	No	No	No	No	2016-12-20
fprnradio.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
freewoodpost.com	No	No	No	No	No	No	Yes	No	No	2016-12-20
geoengineeringwatch.org	Yes	Yes	No	No	No	No	No	No	No	2016-06-29

Continued on next page

Table 1 – continued from previous page

Source	FNW	MZ	DD	USN	NR	CBS	A	NPR	Snopes	Date Added
globalassociatednews.com	No	No	No	No	Yes	No	Yes	No	No	2016-12-20
globalresearch.ca	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
gomerblog.com	Yes	No	No	No	No	No	No	No	No	2016-06-29
govtslaves.info	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
gulagbound.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
hangthebankers.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
humansarefree.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
huzlers.com	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	2016-06-29
ifyouonlynews.com	No	Yes	No	No	No	Yes	No	No	No	2016-12-20
infowars.com	Yes	Yes	Yes	Yes	No	Yes	No	No	No	2016-06-29
intellihub.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
itaglive.com	Yes	No	No	No	No	No	No	No	No	2016-06-29
jonesreport.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
lewrockwell.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
liberalamerica.org	No	Yes	No	No	No	No	No	No	No	2016-12-20
libertymovementradio.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
libertytalk.fm	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
libertyvideos.org	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
lightlybraisedturnip.com	No	No	No	No	Yes	No	No	No	No	2016-12-20
nationalreport.net	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	2016-06-29
naturalnews.com	Yes	Yes	Yes	Yes	No	No	No	No	No	2016-06-29
ncscooper.com	No	Yes	No	No	No	No	No	No	Yes	2016-12-20
newsbiscuit.com	Yes	Yes	Yes	No	No	No	No	No	No	2016-06-29
newslo.com[a]	Yes	Yes	Yes	Yes	No	No	No	No	No	2016-06-29
newsmutiny.com	Yes	Yes	Yes	No	No	No	No	No	No	2016-06-29
newswire-24.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
nodisinfo.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
now8news.com	No	Yes	No	No	No	Yes	No	No	Yes	2016-12-20
nowtheendbegins.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
occupydemocrats.com	No	Yes	Yes	No	No	No	No	No	No	2016-12-20
other98.com	No	Yes	Yes	No	No	No	No	No	No	2016-12-20
pakalertpress.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
politicalblindspot.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
politicalears.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
politicops.com	Yes	Yes	No	No	No	Yes	No	No	No	2016-06-29
politicususa.com	No	Yes	No	No	No	No	No	No	No	2016-12-20
prisonplanet.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
react365.com	No	Yes	No	No	No	Yes	No	No	Yes	2016-12-20
realfarmacy.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
realnewsrightnow.com	Yes	Yes	Yes	No	No	Yes	No	No	No	2016-06-29
redflagnews.com	Yes	Yes	No	Yes	No	No	No	No	No	2016-06-29
redstate.com	No	Yes	Yes	No	No	No	No	No	No	2016-12-20
rilenews.com	Yes	Yes	Yes	No	No	Yes	No	No	No	2016-06-29
rockcitytimes.com	Yes	No	No	No	No	No	No	No	No	2016-06-29
satiratribune.com	No	Yes	No	No	No	No	No	No	Yes	2016-12-20
stuppid.com	No	No	No	No	No	No	No	No	Yes	2016-12-20
theblaze.com	No	Yes	No	No	No	No	No	No	No	2016-12-20
thebostontribune.com	No	No	No	No	No	Yes	No	No	No	2016-12-20
thedailysheeple.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
thedcgazette.com[b]	Yes	Yes	Yes	Yes	No	Yes	No	No	No	2016-06-29
thefreethoughtproject.com	No	Yes	Yes	No	No	No	No	No	No	2016-12-20
thelapine.ca	Yes	No	No	No	No	No	Yes	No	No	2016-06-29
thenewsnerd.com	Yes	Yes	No	No	Yes	No	No	No	No	2016-06-29
theonion.com	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	2016-06-29
theracketreport.com	No	No	No	No	No	No	Yes	No	No	2016-12-20

Continued on next page

Table 1 – continued from previous page

Source	FNW	MZ	DD	USN	NR	CBS	A	NPR	Snopes	Date Added
therundownlive.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
thespoof.com	Yes	No	No	No	No	No	Yes	No	No	2016-06-29
theuspatriot.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
truthfrequencyradio.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
twitchy.com	No	No	Yes	No	No	No	No	No	No	2016-12-20
unconfirmedsources.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
USAToday.com.co	No	No	No	No	No	No	No	Yes	Yes	2016-12-20
usuncut.com	No	Yes	Yes	No	No	No	No	No	No	2016-12-20
veteranstoday.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
wakingupwisconsin.com	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
weeklyworldnews.com	Yes	No	No	Yes	No	No	Yes	No	No	2016-06-29
wideawakeamerica.com	Yes	No	No	No	No	No	No	No	No	2016-06-29
winningdemocrats.com	No	Yes	No	No	No	No	No	No	No	2016-12-20
witscience.org	Yes	Yes	No	No	No	No	No	No	No	2016-06-29
wnd.com	No	Yes	No	No	No	No	No	No	No	2016-12-20
worldnewsdailyreport.com	Yes	Yes	Yes	No	No	No	Yes	No	Yes	2016-06-29
worldtruth.tv	Yes	Yes	No	Yes	No	No	No	No	No	2016-06-29
yournewswire.com	No	No	No	No	No	Yes	No	No	No	2016-12-20

[a] `politicops.com` is a mirror of `newslo.com`.
[b] `thedcgazette.com` is a mirror of `dcgazette.com`.

Hoaxy data

Our analysis focuses on the period from mid-May 2016 to the end of March 2017. During this time, we collected 15,053 fact-checking articles and 389,569 misinformation *claims*. Using the Twitter API, the Hoaxy system collected 1,133,674 public posts that included links to fact checks and 13,617,425 public posts linking to claims. We did not use a streaming sample, but rather the "POST statuses/filter" API endpoint, which provides all public tweets matching our query — namely, all tweets linking to the misinformation and fact-checking sites in our list. As shown in Fig. 10, misinformation websites each produced approximately 100 articles per week, on average. Toward the end of the study period, these claims were shared in approximately 30 tweets per article per week, on average. However, as discussed in the main text, success is extremely heterogeneous across articles. This is the case irrespective of whether we measure success through the number of tweets (Fig. 11(a)) or accounts (Fig. 11(b)) sharing a claim.

Content Analysis

Our analysis considers content published by a set of websites flagged as sources of misinformation by third-party journalistic and fact-checking organizations (Table 1). This source-based approach relies on the assumption that most of the claims published by our compilation of sources are some type of misinformation, as we cannot fact-check each individual claim. We validated this assumption by estimating the rate of false positives, i.e, verified claims, in the corpus. We manually evaluated a random sample of articles ($N = 50$) drawn from our corpus, stratified by source. We considered only those sources whose articles were tweeted at least once in the period of interest. To draw an article, we first select a source at random with replacement, and then choose one of the articles it published, again at random but without replacement. We repeated our analysis on an additional sample ($N = 50$) in which the chances of drawing an article are proportional to the number of times it was tweeted. This 'sample by tweet' is thus biased toward more popular sources.

It is important to note that articles with unverified claims are sometimes updated after being debunked. This happens usually late, after the claim has spread, and could lead to overestimating the rate of false positives. To mitigate this phenomenon, the earliest snapshot of each article was retrieved from the Wayback Machine at the Internet Archive (`archive.org`). If no snapshot was available, we retrieved the version of the page current at verification time. If the page was missing from the website or the website was down, we reviewed the title and body of the article crawled by Hoaxy. We gave priority to the current version over the possibly more accurate crawled version because, in deciding whether a piece of content is misinformation, we want to consider any form of visual evidence included with it, such as images or videos.

After retrieving all articles in the two samples, each article was evaluated independently by two reviewers (two of the authors), using a rubric summarized in Fig. 12. Each article was then labeled with the majority label, with ties broken by a third reviewer (another author). Fig 13 shows the results of the analysis. We report the fractions of articles that were verified and that could not be verified (inconclusive), out of the total number of articles that contain any factual claim. The rate of false positives is below 15% in both samples.

Bot classification

To show that a few social bots are disproportionately responsible for the spread of misinformation, we considered a random sample of accounts that shared at least one claim, and evaluated them using the bot classification system Botometer. Out of 1,000 sampled accounts, 85 could not be inspected because they had been either suspended, deleted, or turned private. For each of the remaining 915, Botometer returned a probabilistic score of likelihood that the account is automated, or *bot score*. To quantify how many account are likely bots, we binarize bot scores using a threshold of 0.5. This is a conservative choice to minimize false negatives and especially false positives, as shown in prior work [40]. Table 2 shows the fraction of accounts with scores above the threshold. To give a sense of their overall impact in the spreading of misinformation, Table 2 also shows the fraction of tweets with claims posted by accounts that are likely bots, and the number of unique claims included in those tweets overall. As a comparison, we also tally the fact-checks shared by these accounts, showing that accounts that are likely bots tended to focus on sharing misinformation.

In the main text we show the distributions of bot scores for this sample of accounts, as well as for a sample of accounts that have been most active in spreading claims (*super-spreaders*). To select the super-spreaders, we ranked

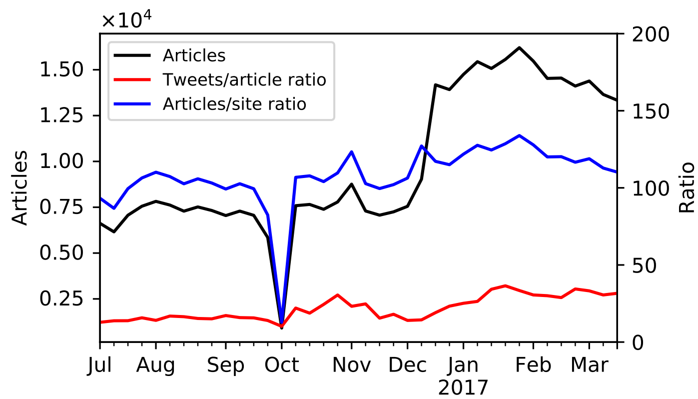

Figure 10: Weekly tweeted claim articles, tweets/article ratio and articles/site ratio. The collection was briefly interrupted in October 2016. In December 2016 we expanded the set of claim sources, from 71 to 121 websites.

Table 2: Analysis of likely bots and their misinformation spreading activity based on a random sample of Twitter accounts sharing at least one claim.

	Total	Likely bots	Percentage
Accounts	915	77	8%
Tweets with claims	11,656	3,857	33%
Unique claims	7,726	2,819	36%
Tweets with fact-checks	598	27	5%
Unique fact-checks	395	25	6%

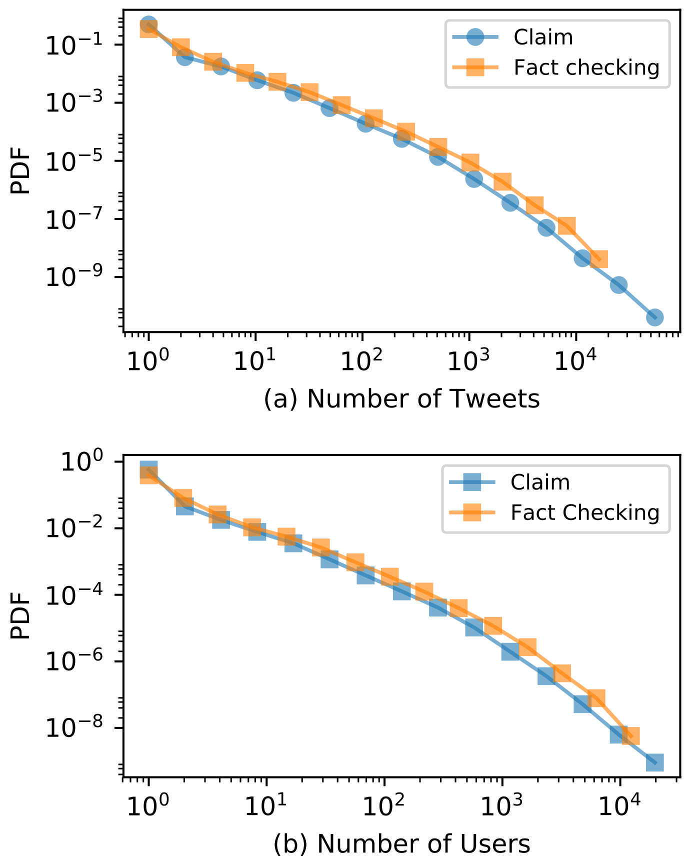

Figure 11: Probability distributions of popularity of claims and fact-checking articles, measured by (a) the number of tweets and (b) the number of accounts sharing links to an article.

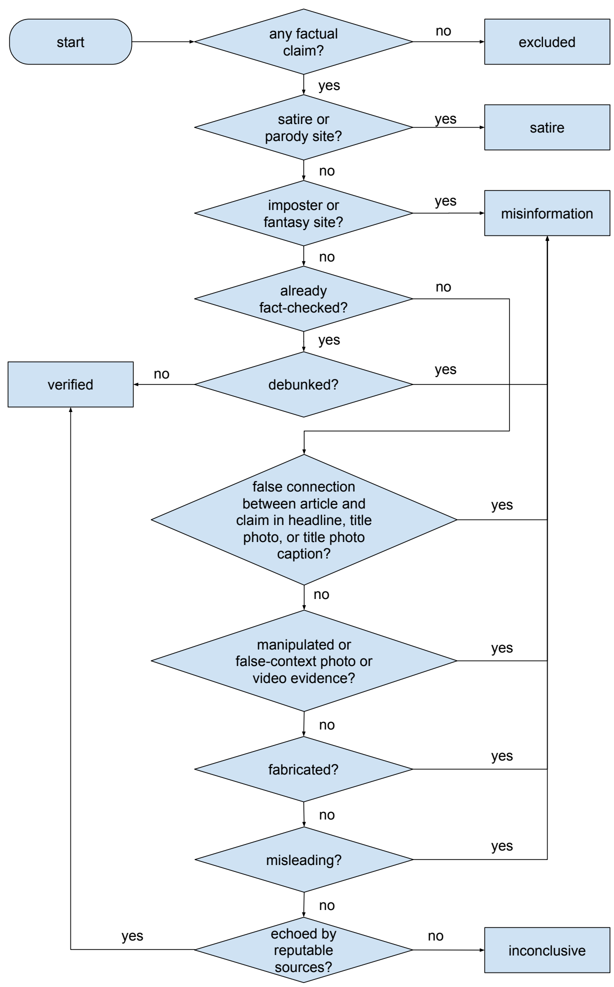

Figure 12: Flowchart summarizing the annotation rubric employed in the content analysis.

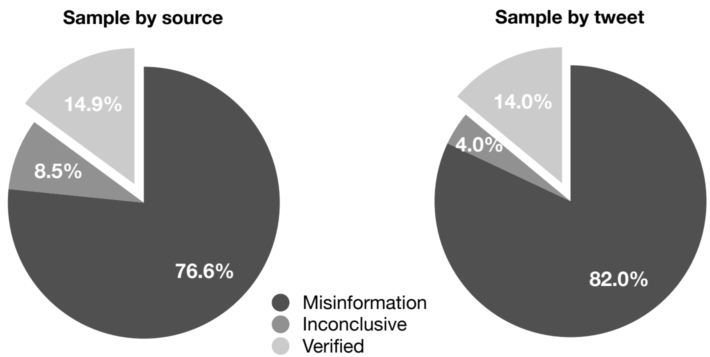

Figure 13: Content analysis based on two samples of claims. Sampling articles by source gives each source equal representation, while sampling by tweets biases the analysis toward more popular sources. We excluded from the sample by source three articles that did not contain any factual claims. Satire articles are grouped with misinformation, as explained in the main text.

all accounts by how many tweets with claims they posted, and considered the top 1,000 accounts. We then performed the same classification steps discussed above. For the same reasons mentioned above, we could not obtain scores for 39 of these accounts, leaving us with a sample of 961 scored accounts. Our notion of super-spreader is based upon ranking accounts by activity and taking those above a threshold. We experimented with different thresholds, and found that they do not change our conclusions that super-spreaders are more likely to be social bots.

Background

Tracking abuse of social media has been a topic of intense research in recent years. The analysis in the main text leverages Hoaxy, a system focused on tracking the spread of fake news. Here we give a brief overview of other systems designed to monitor the spread of misinformation on social media. This is related to the problem of detecting and resolving rumors, which is the subject of a recent survey by Zubiaga *et al.* [46].

Beginning with the detection of simple instances of political abuse like *astroturfing* [31], researchers noted the need for automated tools for monitoring social media streams. Several such systems have been proposed in recent years, each with a particular focus or a different approach. The Truthy system relied on network analysis techniques [33]. The TweetCred system [2] focuses on content-based features and other kind of metadata, and distills a measure of overall information credibility.

More recently, specific systems have been proposed to detect rumors. These include RumorLens [34], TwitterTrails [24], FactWatcher [11], and News Tracer [22].

The news verification capabilities of these systems range from completely automatic (TweetCred), to semi-automatic (TwitterTrails, RumorLens, News Tracer). In addition, some of them let the user explore the propagation of a rumor with an interactive dashboard (TwitterTrails, RumorLens). These systems vary in their capability to monitor the social media stream automatically, but in all cases the user is required to enter a seed rumor or keyword to operate them.

Finally, since misinformation can be propagated by coordinated online campaigns, it is important to detect whether a meme is being artificially promoted. Machine learning has been applied successfully to the task of early discriminating between trending memes that are either organic or promoted by means of advertisement [41].

References

[1] A. Bessi and E. Ferrara. Social bots distort the 2016 us presidential election online discussion. *First Monday*, 21(11), 2016.

[2] C. Castillo, M. Mendoza, and B. Poblete. Information credibility on Twitter. In *Proceedings of the 20th International Conference on World Wide Web*, page 675, 2011.

[3] M. Conover, J. Ratkiewicz, M. Francisco, B. Gonçalves, A. Flammini, and F. Menczer. Political polarization on twitter. In *Proc. 5th International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2011.

[4] M. D. Conover, B. Gonçalves, A. Flammini, and F. Menczer. Partisan asymmetries in online political activity. *EPJ Data Science*, 1:6, 2012.

[5] C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer. Botornot: A system to evaluate social bots. In *Proceedings of the 25th International Conference Companion on World Wide Web*, pages 273–274, 2016.

[6] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi. The spreading of misinformation online. *Proc. National Academy of Sciences*, 113(3):554–559, 2016.

[7] E. Ferrara. Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election. *First Monday*, 22(8), 2017.

[8] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini. The rise of social bots. *Comm. ACM*, 59(7):96–104, 2016.

[9] J. Gottfried and E. Shearer. News use across social media platforms 2016. White paper, Pew Research Center, May 2016.

[10] L. Gu, V. Kropotov, and F. Yarochkin. The fake news machine: How propagandists abuse the internet and manipulate the public. Trendlabs research paper, Trend Micro, 2017.

[11] N. Hassan, A. Sultana, Y. Wu, G. Zhang, C. Li, J. Yang, and C. Yu. Data in, fact out: Automated monitoring of facts by factwatcher. *Proc. VLDB Endow.*, 7(13):1557–1560, Aug. 2014.

[12] N. O. Hodas and K. Lerman. How limited visibility and divided attention constrain social contagion. In *Proc. ASE/IEEE International Conference on Social Computing*, 2012.

[13] P. J. Hotez. Texas and its measles epidemics. *PLOS Medicine*, 13(10):1–5, 10 2016.

[14] L. Howell et al. Digital wildfires in a hyperconnected world. In *Global Risks*. World Economic Forum, 2013.

[15] T. Jagatic, N. Johnson, M. Jakobsson, and F. Menczer. Social phishing. *Communications of the ACM*, 50(10):94–100, October 2007.

[16] Y. Jun, R. Meng, and G. V. Johar. Perceived social presence reduces fact-checking. *Proceedings of the National Academy of Sciences*, 114(23):5976–5981, 2017.

[17] D. M. Kahan. Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making*, 8(4):407–424, 2013.

[18] D. Lazer, M. Baum, Y. Benkler, A. Berinsky, K. Greenhill, F. Menczer, M. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. Sloman, C. Sunstein, E. Thorson, D. Watts, and J. Zittrain. The science of fake news. *Science*, in press.

[19] M. S. Levendusky. Why do partisan media polarize viewers? *American Journal of Political Science*, 57(3):611—-623, 2013.

[20] S. Lewandowsky, U. K. Ecker, and J. Cook. Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition*, 6(4):353–369, 2017.

[21] W. Lippmann. *Public Opinion*. Harcourt, Brace and Company, 1922.

[22] X. Liu, A. Nourbakhsh, Q. Li, R. Fang, and S. Shah. Real-time rumor debunking on twitter. In *Proceedings of the 24^{th} ACM International on Conference on Information and Knowledge Management*, CIKM '15, pages 1867–1870, New York, NY, USA, 2015. ACM.

[23] B. Markines, C. Cattuto, and F. Menczer. Social spam detection. In *Proc. 5th International Workshop on Adversarial Information Retrieval on the Web (AIRWeb)*, 2009.

[24] P. T. Metaxas, S. Finn, and E. Mustafaraj. Using twittertrails.com to investigate rumor propagation. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing*, CSCW'15 Companion, pages 69–72, 2015.

[25] A. Mosseri. News feed fyi: Showing more informative links in news feed. press release, Facebook, June 2017.

[26] E. Mustafaraj and P. T. Metaxas. From obscurity to prominence in minutes: Political speech and real-time search. In *Proc. Web Science Conference: Extending the Frontiers of Society On-Line*, 2010.

[27] A. Nematzadeh, G. L. Ciampaglia, F. Menczer, and A. Flammini. How algorithmic popularity bias hinders or promotes quality. Preprint 1707.00574, arXiv, 2017.

[28] D. Nikolov, D. F. M. Oliveira, A. Flammini, and F. Menczer. Measuring online social bubbles. *PeerJ Computer Science*, 1(e38), 2015.

[29] E. Pariser. *The filter bubble: How the new personalized Web is changing what we read and how we think*. Penguin, 2011.

[30] X. Qiu, D. F. M. Oliveira, A. S. Shirazi, A. Flammini, and F. Menczer. Limited individual attention and online virality of low-quality information. *Nature Human Behavior*, 1:0132, 2017.

[31] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer. Detecting and tracking political abuse in social media. In *Proc. 5th International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2011.

[32] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer. Detecting and tracking political abuse in social media. In *Proc. 5th International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2011.

[33] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, S. Patil, A. Flammini, and F. Menczer. Truthy: Mapping the spread of astroturf in microblog streams. In *Proceedings of the 20th International Conference Companion on World Wide Web*, WWW '11, pages 249–252, 2011.

[34] P. Resnick, S. Carton, S. Park, Y. Shen, and N. Zeffer. Rumorlens: A system for analyzing the impact of rumors and corrections in social media. In *Proc. Computational Journalism Conference*, 2014.

[35] M. J. Salganik, P. S. Dodds, and D. J. Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854–856, 2006.

[36] C. Shao, G. L. Ciampaglia, A. Flammini, and F. Menczer. Hoaxy: A platform for tracking online misinformation. In *Proceedings of the 25th International Conference Companion on World Wide Web*, pages 745–750, 2016.

[37] N. Stroud. *Niche News: The Politics of News Choice*. Oxford University Press, 2011.

[38] V. Subrahmanian, A. Azaria, S. Durst, V. Kagan, A. Galstyan, K. Lerman, L. Zhu, E. Ferrara, A. Flammini, F. Menczer, A. Stevens, A. Dekhtyar, S. Gao, T. Hogg, F. Kooti, Y. Liu, O. Varol, P. Shiralkar, V. Vydiswaran, Q. Mei, and T. Hwang. The darpa twitter bot challenge. *IEEE Computer*, 49(6):38–46, 2016.

[39] C. R. Sunstein. *Going to Extremes: How Like Minds Unite and Divide*. Oxford University Press, 2009.

[40] O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini. Online human-bot interactions: Detection, estimation, and characterization. In *Proc. Intl. AAAI Conf. on Web and Social Media (ICWSM)*, 2017.

[41] O. Varol, E. Ferrara, F. Menczer, and A. Flammini. Early detection of promoted campaigns on social media. *EPJ Data Science*, 6(1):13, 2017.

[42] L. von Ahn, M. Blum, N. J. Hopper, and J. Langford. Captcha: Using hard ai problems for security. In E. Biham, editor, *Advances in Cryptology — Proceedings of EUROCRYPT 2003: International Conference on the Theory and Applications of Cryptographic Techniques*, pages 294–311. Springer, 2003.

[43] C. Wardle. Fake news. It's complicated. White paper, First Draft News, February 2017.

[44] J. Weedon, W. Nuland, and A. Stamos. Information operations and facebook. white paper, Facebook, April 2017.

[45] S. C. Woolley and P. N. Howard. Computational propaganda worldwide: Executive summary. Working Paper 2017.11, Oxford Internet Institute, 2017.

[46] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter. Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys*, 50, 2018. Forthcoming.