arXiv:1707.07592v1 [cs.SI] 24 Jul 2017

The spread of fake news by social bots

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

Indiana University, Bloomington

Abstract

The massive spread of fake news 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 digital misinformation 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 key role in the spread of fake news. Accounts that actively spread misinformation are significantly more likely to be bots. Automated accounts are particularly active in the early spreading phases of viral claims, and tend to target influential users. Humans are vulnerable to this manipulation, retweeting bots who post false news. Successful sources of false and biased claims are heavily supported by social bots. These results suggests 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 [7], you are exposed to a daily dose of false or misleading content — hoaxes, rumors, conspiracy theories, fabricated reports, click-bait headlines, and even satire. We refer to this misinformation collectively as false or fake news. The incentives are well understood: traffic to fake news sites is easily monetized through ads [16], but political motives can be equally or more powerful [18, 23]. The massive spread of false news has been identified as a major global risk [11]. Claims that fake news can influence elections and threaten democracies [8] are hard to prove. Yet we have witnessed abundant demonstrations of real harm caused by misinformation spreading on social media, from dangerous health decisions [10] to manipulations of the stock market [6].

A complex mix of cognitive, social, and algorithmic biases contribute to our vulnerability to manipulation by online misinformation. 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 [22]. Of course, we do not live in such an ideal world. Our online social networks are strongly polarized and segregated along political lines [3, 2]. The resulting "echo chambers" [28, 21] provide selective exposure to news sources, biasing our view of the world [20]. 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 [24, 9, 19]. All of these factors play into confirmation bias and motivated reasoning [26, 14], making the truth hard to discern.

While fake news are not a new phenomenon [15], 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* [23, 6, 27, 29]. 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 [12] 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 be believe it. Moreover, the amplification of fake news 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 [13].

The fight against fake news requires a grounded assessment of the mechanism by which misinformation 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 fake news. The earliest manifestations were uncovered in 2010 [18, 23]. Since then, we have seen influential bots affect online debates about vaccination policies [6] and participate actively in political campaigns, both in the U.S. [1] and other countries [32]. 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 fake news and its manipulation by social bots is now feasible thanks to two tools developed in our lab: the *Hoaxy* platform to track the online spread of claims [25] and the *Botometer* machine learning algorithm to detect social bots [4, 29]. Let us examine how social bots promoted hundreds of thousands of fake news articles spreading through millions of Twitter posts during and following the 2016 U.S. presidential campaign.

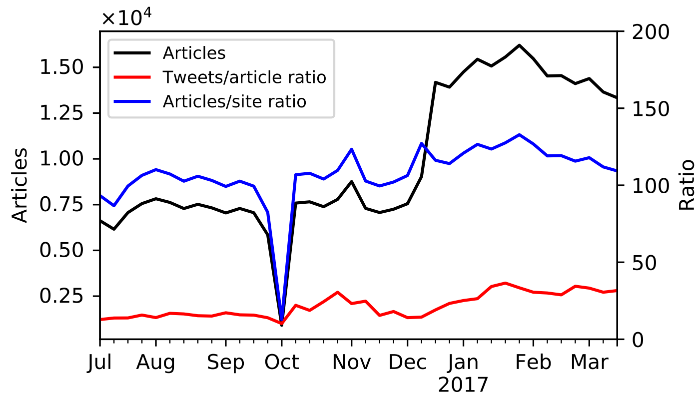

Figure 1: Weekly tweeted claim articles, tweets/article ratio and articles/site ratio. The collection was briefly interrupted in October 2016. In December 2016 we expanded the set of claim sources, from 71 to 122 websites.

2 Results

We crawled the articles published by seven independent fact-checking organizations and 122 websites that, according to established media, routinely publish false and/or misleading news. 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.

As shown in Fig. 1, fake news websites each produced approximately 100 articles per week, on average. The virality of these claims increased to approximately 30 tweets per article per week, on average. However, success is extremely heterogeneous across articles. Fig. 2(a) illustrates an example of viral claim. Whether we measure success by number of people 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 (Fig. 2(b,c)). Unfortunately, and consistently with prior analysis using Facebook data [22], we find that the popularity profiles profiles of false news are indistinguishable from those of fact-checking articles. Most claims are spread through original tweets and especially retweets, while few are shared in replies (Fig. 3).

The claim-posting patterns shown in Fig. 4(a) highlight inorganic support. The points aligned along the diagonal lines (on the left of the plot) indicate that for many articles, one or two accounts are responsible for the entirety of the activity. Furthermore, some accounts share the same claim up to 100

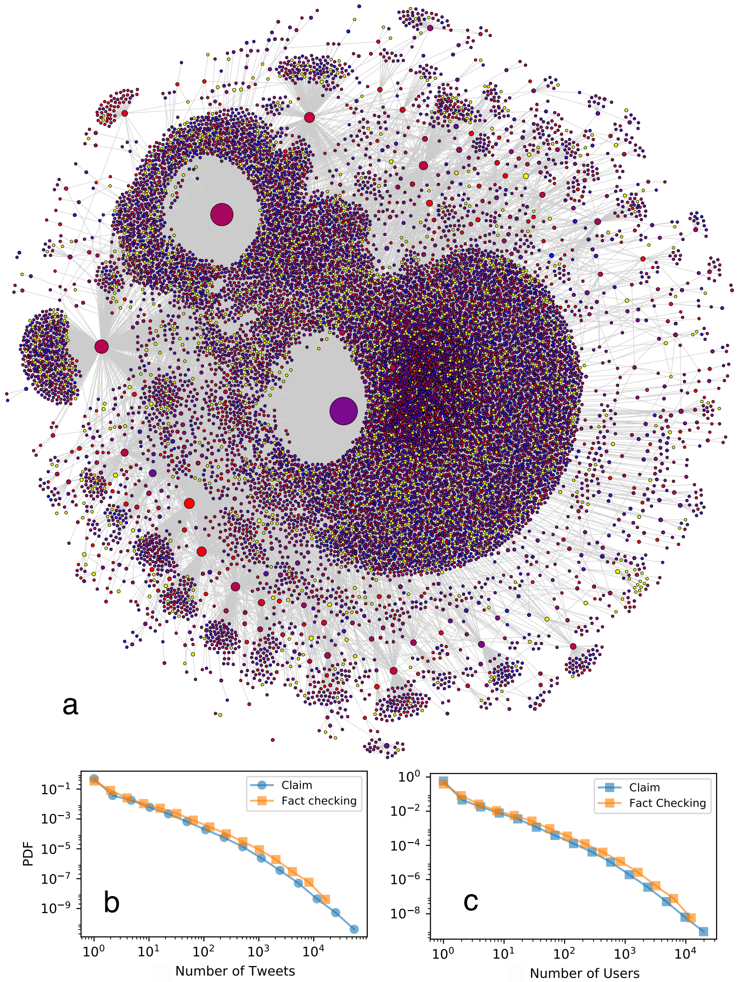

Figure 2: Virality of fake news. (a) Diffusion network for the article titled *"Spirit cooking": Clinton campaign chairman practices bizarre occult ritual,* published by the conspiracy site `Infowars.com` four days before the 2016 U.S. election. Over 30 thousand tweets shared this claim; only the largest 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 this network is available online (`iunetsci.github.io/HoaxyBots/`). The two charts plot the probability distributions (density functions) of (b) number of tweets per article and (c) number of users per article, for claims and fact-checking articles.

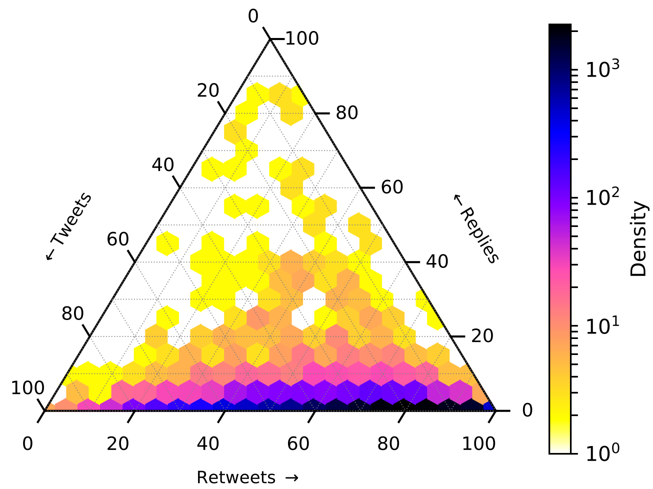

Figure 3: Distribution of types of tweet spreading claims. Each article is mapped along three axes representing the percentages of different types of messages that share it: original tweets, retweets, and replies. Color represents the number of articles in each bin, on a log-scale.

times or more. The ratio or tweets per user decreases for more viral claims, indicating more organic spreading. But Fig. 4(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 fake news 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 our detection systems can be found in Methods.

Fig. 5 confirms that the super-spreaders 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 fake news. To test this conjecture, we examined the accounts that post viral claims at different phases of their spreading cascades. As shown in Fig. 6, bots actively share links in the first few seconds after they are first posted. This early intervention exposes many users to the fake news article, effectively boosting its viral diffusion.

Another strategy used by bots is illustrated in Fig. 7(a): influential users are often mentioned in tweets that link to debunked claims. Bots seem to employ this targeting strategy repetitively; for example, a single account mentioned `@realDonaldTrump` in 18 tweets linking the claim shown in the figure. For a systematic investigation, let us use the number of followers of a Twitter user as a

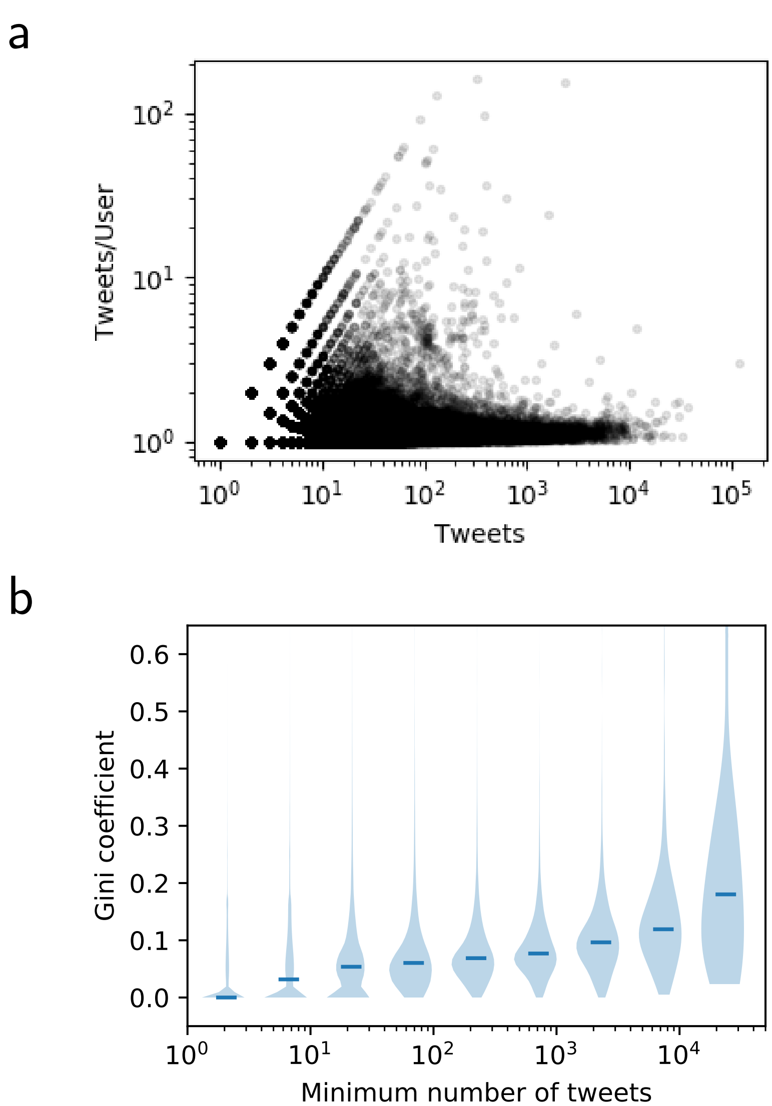

Figure 4: Concentration of claim-sharing activity. (a) Scatter plot of tweets/account ratio versus number of tweets sharing a claim. The darkness of a point represents the number of claims. (b) Source concentration for claims with different popularity. We consider a collection of articles shared by a minimum number of tweets as a popularity group. For claims in each of these groups, we show 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.

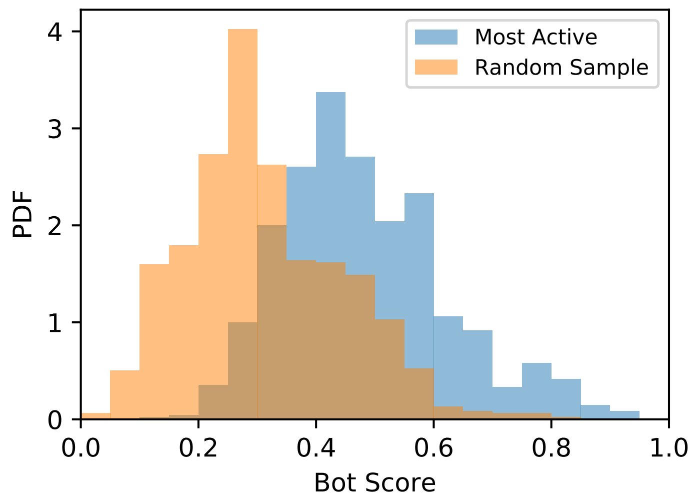

Figure 5: Bot score distributions for a random sample of 915 users who posted at least one link to a claim, and for the 961 accounts that most actively share fake news (super-spreaders). The two groups have significantly different scores ($p < 10^{-4}$ according to a Welch's unequal-variances t-test).

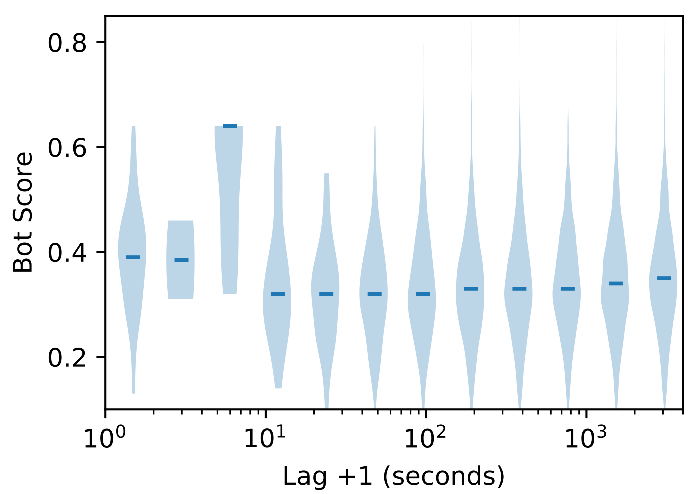

Figure 6: Temporal evolution of bot score distributions for a sample of 60,000 accounts that participate in the spread of the 1,000 most viral claims. We focus on the first hour since a fake news article appears, and divide this early spreading phase into logarithmic lag intervals.

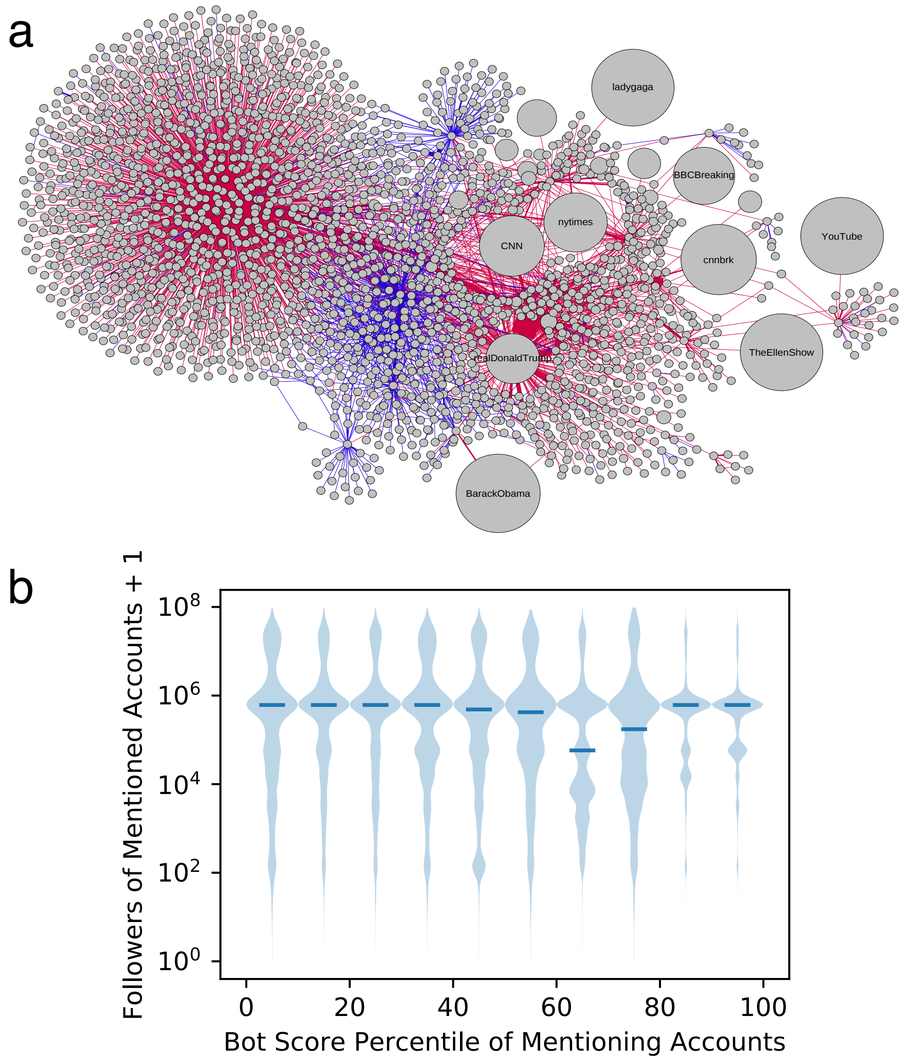

Figure 7: (a) Example of targeting for the claim *Report: three million votes in presidential election cast by illegal aliens*, published by `Infowars.com` on November 14, 2016 and shared over 18 thousand times on Twitter. Only a portion of the diffusion network is shown. Nodes stand for Twitter accounts, 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) Distributions of the number of followers for Twitter users who are mentioned or replied to in posts that link to the most viral 1000 claims. The distributions are grouped by bot score of the account that creates the mention or reply.

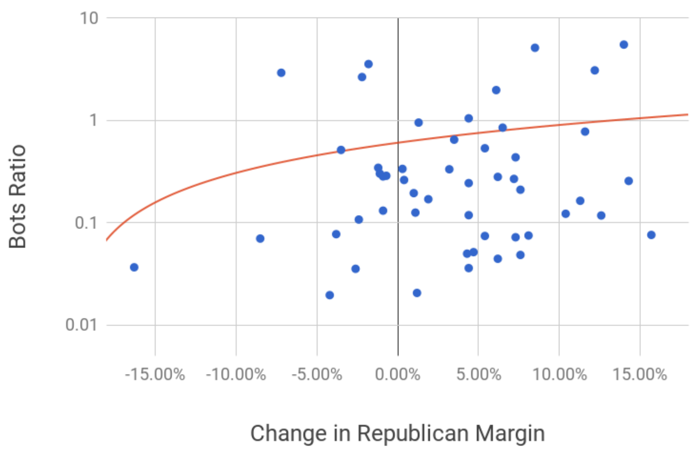

Figure 8: Scatter plot of bot activity vs. difference between actual and predicted vote margin by U.S. states. For each state, we compared the vote margin with forecasts based on the final polls on election day. A positive percentage indicates a larger Republican margin or smaller Democratic margin. To gauge fake news sharing activity by 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. Ratios above one indicate states with higher than expected bot activity. We also plot a linear regression (red line). Pearson's correlation is $\rho = 0.15$.

proxy for their influence. We consider tweets that mention or reply to a user and include a link to a viral fake news story. Tweets tend to mention popular people, of course. However, Fig. 7(b) shows that when accounts with the highest bot scores share these links, they tend to target users with a higher median number of followers and lower variance. 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 targets will spread it.

We examined whether bots 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. As a baseline, we extracted state locations 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 ($p < 10^{-4}$). Given the widespread but unproven allegations that fake news may have influenced the 2016 U.S. elections,

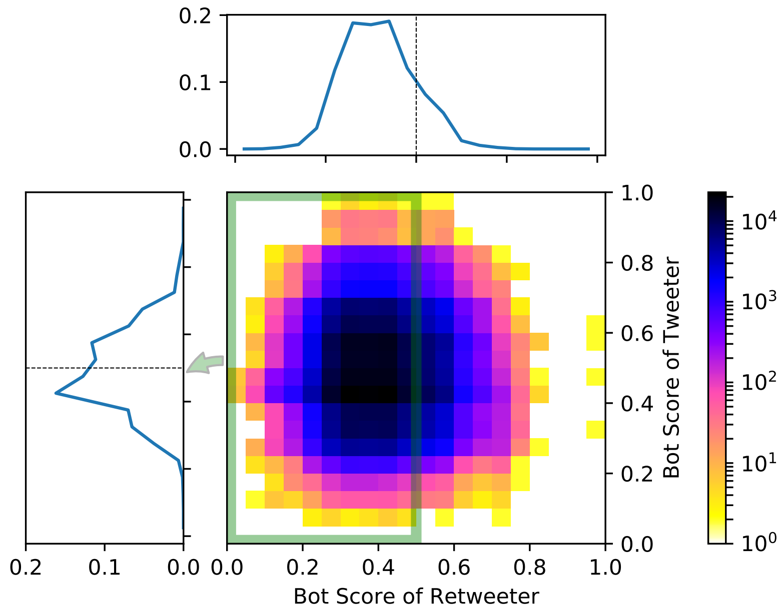

Figure 9: Joint distribution of the bot scores of accounts that retweeted links to claims and accounts that had originally posted the links. Color represents the number of retweeted messages in each bin, on a log scale. Projections show the distributions of bot scores for retweeters (top) and for accounts retweeted by humans (left).

we explored the relationship between bot activity and voting data. The ratio of bot frequencies with respect to state baselines provides an indication of claim-sharing activity by state. Fig. 8 shows a weak correlation between this ratio and the change in actual vote margin with respect to state forecasts (see Methods). Naturally this correlation does not imply that voters were affected by bots sharing fake news; many other factors can explain the election outcome. However it is remarkable that states most actively targeted by misinformation-spreading bots tended to have more surprising election results.

Having found that bots are employed to drive the viral spread of fake news, 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. 9 shows that human do most of the retweeting, and they retweet claims posted by bots as much as by other humans. This suggests that humans can be successfully manipulated through social bots.

Finally, we compared the extent to which social bots successfully manipulate the information ecosystem in support of different sources of online misinforma-

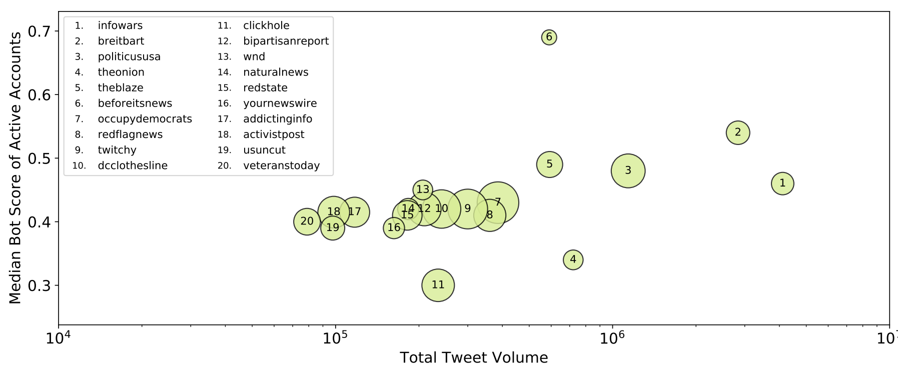

Figure 10: Popularity and bot support for the top 20 fake news websites. Popularity is measured by total tweet volume (horizontal axis) and median number of tweets per claim (circle area). Bot support is gauged by the median bot score of the 100 most active accounts posting links to articles from each source (vertical axis).

tion. 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. 10, one site (`beforeitsnews.com`) stands out in terms of manipulation, but other well-known sources also have many bots among their promoters. At the bottom we find satire sites like *The Onion*.

3 Discussion

Our analysis provides quantitative empirical evidence of the key role played by social bots in the viral spread of fake news online. 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 fake news 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 false news just as much as other humans. Successful sources of fake news 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 [31, 17].

Our findings demonstrate that social bots are an effective tool to manipulate social media and deceive their users. 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 [5] 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 [22]. 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 [6, 27, 29]. 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 on improved abuse detection technology.

An alternative strategy would be to employ CAPTCHAs [30], 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 fake news and fact checking on Twitter [25]. A search engine, interactive visualizations, and open-source software are freely available (`hoaxy.iuni.iu.edu`). The data is accessible through a public API.

The links to the stories considered here were crawled from websites that routinely publish unsubstantiated or 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 51 more in mid-December 2016. The full list of sources is available on the Hoaxy website. 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`.

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. 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.

We apply no editorial judgment about the truthfulness of individual claims; some may be accurate (false positives) and some fake news may be missed (false negatives). The great majority of claims are misleading, including fabricated news, hoaxes, rumors, conspiracy theories, click bait, and politically biased content. We did not exclude satire because many fake-news sources label their content as satirical, making the distinction problematic. Furthermore, 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. 6) with articles from `theonion.com` excluded and the results were not affected.

The bot score of Twitter accounts was computed using the Botometer service, developed at Indiana University and available through a public API (`botometer.iuni.iu.edu`). Botometer evaluates the extent to which an account exhibits similarity to the characteristics of social bots [4]. 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 features capturing various dimensions of information diffusion as well as user metadata, friend statistics, temporal patterns, part-of-speech and sentiment analysis. These features are fed to a machine learning algorithm trained on thousands of examples of human and bot accounts. The system has high accuracy [29] and is widely adopted, serving over 100 thousand requests daily.

The location analysis in Fig. 8 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% sample of public posts from the Twitter streaming API. We considered 164,276 tweets in the same period that included hashtags with the prefix `#election` and a U.S. state location. 2016 election forecast data was obtained from from FiveThirtyEight (`projects.fivethirtyeight.com/2016-election-forecast/`) and vote margins data from the Cook Political Report (`cookpolitical.com/story/10174`).

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. 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.

References

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

[2] 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.

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

[4] 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.

[5] 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.

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

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

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

[9] 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.

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

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

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

[13] 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.

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

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

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

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

[18] 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.

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

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

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

[22] 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.

[23] 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.

[24] 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.

[25] 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.

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

[27] 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.

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

[29] 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.

[30] 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.

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

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