Discovery and classification of Twitter bots

MARIA OIKONOMIDOU, ALEXANDER SHEVTSOV, DESPOINA ANTONAKAKI, POLYVIOS PRATIKAKIS, SOTIRIS IOANNIDIS, and PARASKEVI FRAGOPOULOU, Institute of Computer Science, FORTH, Greece

Twitter's popularity makes it a frequent target of malicious activity. There are plenty of studies analyzing the user activity in Twitter, including the detection of automated behavior. It is the 11th largest social network, with 330 million monthly active users [Twitter Inc. 2018] as of December 2017, and has attracted public figures, organizations, news media, and social authorities.

This openness of user communities and information dissemination capacity unfortunately attracts malicious entities as well, aiming to influence public opinion, or at least convince people of their ability to do so, for reasons such as personal popularity, political influence, or even influencing of international relations. The activity of such agents is often organized in the form of botnets: groups of *sybil* accounts that collectively seek to influence ordinary users. The percentage of Twitter users that are bots has been estimated to be between 9% and 15% [Varol et al. 2017]. Although not all automated accounts are malicious, the potential magnitude of damage has driven a lot of research into the detection of bot accounts [Ferrara et al. 2016; Subrahmanian et al. 2016; Varol et al. 2017], focusing mostly on English-speaking and Arabic-speaking parts of the network, often by machine-translating the latter to the former.

This work presents an analysis of Twitter content, crawled between August 2016 and January 2018, comprising of about 720 million tweets from mostly Greek-speaking users. Our analysis discovered several thousand accounts that exhibit automated content-injection behavior, detected to tweet the same content as other accounts almost concurrently, for multiple times. We mark the accounts that repeatedly tweet in such a synchronized fashion as bots, and study the content, usage patterns, position in the follow graph, as well as community infiltration of such automated bots.

Our method first analyzes individual tweets in the crawled corpus to detect synchronous or near-synchronous activity. Specifically, we assume the model where botnets of multiple accounts controlled by the same agent aim to promote and diffuse content, and affect legitimate users in some way.

The implementation is based on the twAwler Twitter crawler [Pratikakis 2018] and extends it with the detection of concurrent and similar tweets, the extraction of a concurrent content injection graph, and a graph filtering mechanism that aims to reduce noise. We have used these extensions to discover a large number of bot accounts, and propose a method for analysis of the resulting data, that classifies bots according to the community that they most engage, based on Twitter list memberships and their use of hashtags.

CCS Concepts: • **Networks** → **Online social networks**; • **Security and privacy** → *Network security*.

Additional Key Words and Phrases: Twitter, bot detection, botnets, click farms

Authors' address: Maria Oikonomidou, mareco@ics.forth.gr; Alexander Shevtsov, shevtsov@ics.forth.gr; Despoina Antonakaki, despoina@ics.forth.gr; Polyvios Pratikakis; Sotiris Ioannidis; Paraskevi Fragopoulou, Institute of Computer Science, FORTH, N. Plastira 100 Vassilika Vouton, Heraklion, Crete, Greece.

XXXX-XXXX/2018/6-ART $15.00
https://doi.org/10.1145/1122445.1122456

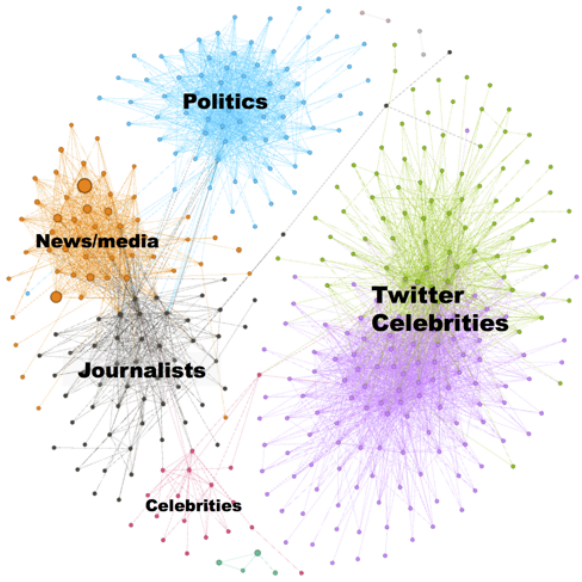

Fig. 1. Filtered list-similarity graph reveals clusters targeting specific interests.

ACM Reference Format:
Maria Oikonomidou, Alexander Shevtsov, Despoina Antonakaki, Polyvios Pratikakis, Sotiris Ioannidis, and Paraskevi Fragopoulou. 2018. Discovery and classification of Twitter bots. 1, 1 (June 2018), 2 pages. https://doi.org/10.1145/1122445.1122456

ACKNOWLEDGMENTS

This work was supported by the project EUNITY, with grant number 740507 (EUROPEAN COMMISSION , Horizon 2020).

REFERENCES

Qiang Cao, Xiaowei Yang, Jieqi Yu, and Christopher Palow. 2014. Uncovering large groups of active malicious accounts in online social networks. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 477–488.

Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. The rise of social bots. *Commun. ACM* 59, 7 (2016), 96–104.

Polyvios Pratikakis. 2018. twAwler: A lightweight twitter crawler. *CoRR* abs/1804.07748 (2018). arXiv:1804.07748 http://arxiv.org/abs/1804.07748

VS Subrahmanian, Amos Azaria, Skylar Durst, Vadim Kagan, Aram Galstyan, Kristina Lerman, Linhong Zhu, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. The DARPA Twitter bot challenge. *Computer* 49, 6 (2016), 38–46.

Twitter Inc. 2018. FORM 10-K annual report. https://investor.twitterinc.com/financial-information/annual-reports/default.aspx.

Onur Varol, Emilio Ferrara, Clayton A Davis, Filippo Menczer, and Alessandro Flammini. 2017. Online human-bot interactions: Detection, estimation, and characterization. *arXiv preprint arXiv:1703.03107* (2017).