Report

What is the proportion of robots in Tweets? - BOTS

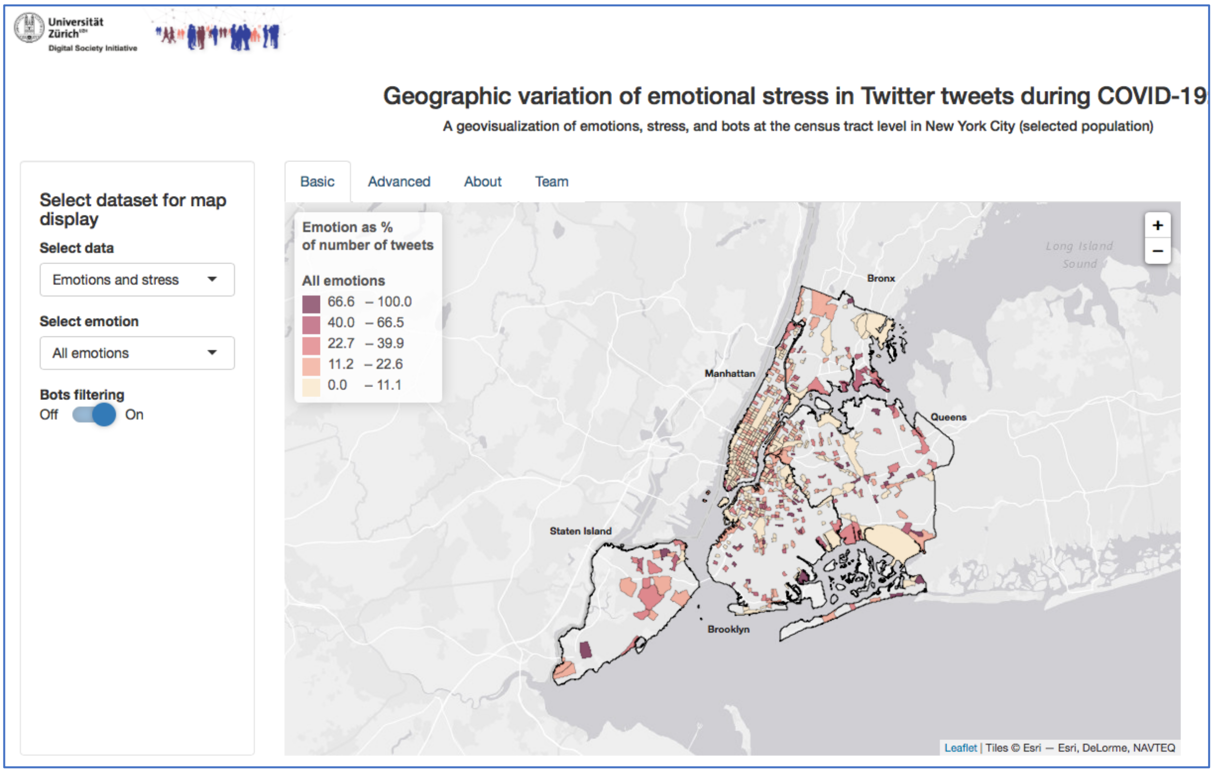

Figure taken from the publication Edry T, Maani N, Sykora M, Elayan S, Hswen Y, Wolf M, Rinaldi F, Galea S, Gruebner O. Real-time geospatial surveillance of localized emotional stress responses to COVID-19: A proof of concept analysis. Health Place. 2021 Jul;70:102598. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1353829221000940 and https://givauzh.shinyapps.io/NYC_tweets/

UZH team

Dr. Oliver Gruebner (PI, Department of Geography and EBPI, oliver.gruebner@uzh.ch)
MSc. Tamar Edry (Department of Geography, assistant, employed for this project)
Dr. Markus Wolf (Department of Psychology)
Dr. Fabio Rinaldi (Dalle Molle Institute for Artificial Intelligence, IDSIA USI-SUPSI and Department of Quantitative Biomedicine, University of Zurich)

External project collaborators

Dr. Suzanne Elayan (Loughborough University, UK)
Dr. Martin Sykora (Loughborough University, UK)
Dr. Yulin Hswen (Harvard Medical School)
Dr. Nason Maani (Boston University School of Public Health)

We thank the Digital Society Initiative for funding this project and report the project outcomes below.

Project outcomes

We applied Botometer[1], the current gold standard for automated twitter account (bot) detection, as well as EMOTIVE and Stresscapes, that is, ontology based natural language processing tools that can monitor expressions associated with general emotion and stress as seen in social media in real-time, respectively. Furthermore, we analysed the geographic trajectories of tweets over time with the help of geographic information systems to assess whether spatial information could help in advancing current bot detection measures. The two studies were approved by the UZH Faculty of Arts and Social Sciences Ethics Committee (Number # 20.12.9).

In our first study, we applied bot detection in geo-referenced tweet data in New York City related to the COVID-19 pandemic[2]. For this proof-of-concept analysis, we developed a tool that potentially can monitor real-time general emotions and stress as seen in social media, can detect stress hotspots at the local level over time, and includes a tool to filter out artificial tweets produced by bots. We argued that if applied to large anonymized datasets, this tool could help inform local responses. The application can be found under the following link[3].

In the second study, we employed a geographic trajectory analysis to evaluate potential improvements in bot detection in these data since current approaches such as Botometer[1] are location agnostic[4]. Specifically, we first investigated the distribution of tweets in single users in a dataset from Europe spanning across the years 2015 to 2018 (Table 1).

Table 1: Descriptive statistics for the data used in this study. The table shows minimum, mean (standard deviations), median, and maximum numbers of tweets per user, as well as total tweets and total users in the dataset across four years.

Year	Min	Mean (SD)	Median	Max	Total tweets	Total users
2015	1	99.50 (± 734.18)	29	193,191	9,356,329	94,033
2016	1	80.21 (± 845.41)	19	209,359	6,864,585	85,584
2017	1	63.49 (± 620.91)	14	126,362	4,453,799	70,148
2018	1	39.40 (± 672.46)	8	137,473	1,931,568	49,026

From these data, we identified 106,798 (89%) valid users and within them, 61,698 (58%) as bots and 45,178 (42%) as non-bots according to Botometer[5]. We also explored the geographic trajectories in these users over time and generated seven new variables that are listed in Table 2.

Table 2: Descriptive statistics for tweets of users across the entire time range 2015-2018

Variable	Description	Min	Mean (SD)	Median	Max	Total
# tweets	Total tweets per user for the entire sampling duration	2	186.36 (±1,574.21)	47	402,550	19,902,700
Tweets per duration	Total tweets per user normalized by the duration of the user's activity (time between first and last tweet in the sample)	0	2.19 (±140.07)	0.10	24,685.71	233,993.97

[1] https://botometer.iuni.iu.edu/#!/
[2] Edry T, Maani N, Sykora M, Elayan S, Hswen Y, Wolf M, Rinaldi F, Galea S, Gruebner O. Real-time geospatial surveillance of localized emotional stress responses to COVID-19: A proof of concept analysis. Health Place. 2021 Jul;70:102598. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1353829221000940
[3] https://givauzh.shinyapps.io/NYC_tweets/
[4] In preparation: Edry T, Elayan S, Sykora M, Maani N, Rinaldi F, Wolf M, Gruebner O. Exploring the geography of Twitter users' behavior for bot detection

# countries	Total number of countries that were visited by a user	1	15.23 (±719.13)	4	215,338	1,626,444
Total length (km)	Total distance travelled (km)	0	28,196.58 (±1,000,160.61)	8,659.16	287,222,629.28	3,014,468,694.22
Countries per duration	Number of countries normalized by the duration of activity (time between first and last tweet in the sample)	0	0.90 (±69.86)	0.01	12,342.86	96,436.94
Mean duration in country (days)	Mean time (days) spend in the visited countries	0	143.44 (±180.33)	81.62	1,343.71	15,318,747.46
Length per duration	Total length (km) normalized by the duration of activity (time between first and last tweet in the sample)	0	1,210.96 (±96,750.82)	18.60	20,483,972.56	129,462,160.53
Median speed (km/h)	Median value of speed (km/h) between each pair of tweets (per user)	0	208.20 (±42,227.55)	0.25	1,364,7900.90	22,234,838.34

In addition, we recoded four of the newly created variables at the user level to reflect extreme behavior in geographic trajectories of users, that is, unrealistic movements across geographic space.
For this, the variable ‘Countries per duration’ was recoded 1 for the upper 5% of the number of countries visited per duration of activity and 0 otherwise and ‘Mean duration in a country’ was coded as 1 if the user spent less than a day in a country on average and 0 otherwise. The variable ‘Length per duration’ was recoded as 1 for the upper 5% of the length travelled in km per duration of activity and 0 otherwise. The variable ‘Median speed’ was coded as 1 for a median speed of 1000km/h and above and 0 otherwise.

Finally, we studied the associations of the recoded variables reflecting extrem geographic trajectories of users and the likelihood of the users being classified as bots in a set of multivariable regression models. We found that the 'Mean duration in a country’, ‘Length per duration’, ‘Countries per duration’, and 'Median speed’ were all significantly and positively associated with the likelihood of a twitter account being classified as a bot. As ‘Length per duration’ and ‘Countries per duration’ were found to be correlated, we compared the two regression models, each using only one of the correlated variables, using the AIC values that estimate the prediction error. The model that included ‘Length per duration’ showed a slightly lower AIC value (145,234 versus 145,245) and was therefore chosen to represent the best fit to the data (Table 3).

Table 3: Results of logistic regression model predicting bots, using and odds ratio (OR), 95% Confidence Interval (CI). Significance code: *** p< .001, ** p< .01, * p< .05, .<.1

Variable	Category	OR (95% CI)
	(Intercept)	1.33 (1.32 - 1.35)***
Mean duration in country (day)	Reference	/
	Less than a day in a country on average (mean)	1.32 (1.12 - 1.56)**
Length per duration	Reference	/
	Upper 5% of the length travelled (km) per duration of activity	1.54 (1.45 - 1.64)***
Median speed	Reference	/
	Median speed >1000 km/h	1.42 (1.14-1.77)**

With this study, we could show that introducing a geographic based approach to identify extreme spatial movement can improve existing bot detection abilities. Furthermore, if applied to large datasets in real time and combined with natural language processing tools to monitor expressions associated with general emotion and stress, our approach could help inform local interventions.