Real-time geospatial surveillance of localized emotional stress responses to COVID-19: a proof of concept analysis

1. The COVID-19 pandemic has highlighted the need for improved disease surveillance efforts to guide local responses. These surveillance efforts stand to contribute to the detection and mitigation of viral spread, but also for the detection, and mitigation of the chronic health conditions that are also increasing in parallel with the spread of the virus. Among these, emerging evidence suggests that there will be substantial mental health consequences of the COVID-19 pandemic (Ettman et al., 2020a; Lima et al., 2020) . Far from being felt equally, the mental health burden has fallen most on those with the least means (Ettman et al., 2020b). Although effective approaches do exist to treat the emotional toll of the pandemic, there are limited tools to document burden in near real-time. In order to fill that void, we aimed to build a conceptual framework, or a road map, for the use of geo-located social media data, e.g. Twitter, in tracking emotional responses across time and space. The main innovations of the framework include the syntheses and integration of an emotion detection tool to geospatial and temporal analyses, filtering out automated bot tweets and allowing the full exploration of their spatial activity, contextualization with social and physical environments, and offering an on-the-fly hotspot analysis of emotional expressions. A system based on these elements could prove to be beneficial to govenments or NGOs, health practitioners, and other stakeholders to prepare for targeted health interventions where they are needed the most.

2. Social media analyses have emerged as a means to examine population level mental health. Twitter in particular has been shown to be an effective platform for the detection of mood and affective states (Hswen et al., 2018) and has been suggested as a way to identify populations in need following natural or human-made disasters through the use of automated emotive analysis, particularly when combined with aggregated geolocation (Gruebner et al., 2018, 2016). However, the use of such data is complicated by automated social media activity (in particular “bot” accounts), which can account for a disproportionate amount of social media posts, particularly around contested or politically sensitive topics, where political actors seek to sow confusion or panic (Shao et al., 2018). Other than influencing political support, numerous malicious applications have been associated with social bots, such as spreading rumours and conspiracy theories,

promoting terrorist propaganda and recruitment, and manipulation of the stock markets (Varol et al., 2017).

3. We investigated the possibility of an anonymized geospatial analytics and surveillance system that illustrates our main concepts and allowing for the assessment of 1) emotional stress content using the EMOTIVE tool of Twitter content, 2) of automated bot like behavior and 3) of spatial hotspots of emotional stress and present these in interactive and dynamic maps by census tract over time.

4. As a proof of concept, we used the EMOTIVE (Sykora et al., 2013) and Stresscapes (Elayan et al., 2020) tools to categorize emotions and overall stress of tweets in a sample dataset of Twitter users geo-located in NYC between January 1st and October 23, 2020. The data consists of 34,140 geo-tagged tweets from 2,204 users who were originally identified during Superstorm Sandy in NYC in 2012, with an average of 15.49 tweets per user in 2020. Emotions and stress found in tweets were calculated as proportions of all tweets (number of tweets with a specific emotion or stress indication divided by all tweets) at census tract level per month. Botometer v.3 (Yang et al., 2019) was used to identify potentially automated Twitter accounts. We used the Moran's I statistic to identify spatial hotspots of above average emotions or stress identified at the census tract level per month.

5. We incorporated this approach into a web-based geo-visualization tool that can depict geographic clusters of negative emotion or stress over time in accordance with the spread of the COVID-19 outbreak, allowing users to toggle to identify specific emotions identified, or negative emotions in general, and to identify the extent to which bot accounts are the source of that emotion by adjusting bot detection parameters (Figure 1).

6. General estimations on the proportion of automated accounts (bots) on twitter range between 9%-15% (Varol et al., 2017), in this specific dataset the number of bots is relatively small, as it uses Twitter-accounts that were originally extracted in 2012 for another project (Gruebner et al., 2018, 2017). Automated accounts at that time were likely less frequent, less sophisticated and more easily identifiable and removable by Twitter. Further, bots from 2012 may be more likely to be inactive, biasing our sample

towards fewer bots and more human users. In this dataset the number of bot accounts was 21 (332 tweets) out of 2,204 users (34,140 tweets), representing 0.95% of accounts and 0.97% of tweets. Despite the low percentage of bots, the proportion of emotions found in the bot tweets (9.9%) was slightly higher than the non bot users (9%). Our observation of bots expressing emotions is supported by previous studies (Kušen and Strembeck, 2019, 2018), and was our main motivation of integrating a bot filtering feature in our tool. Screening the bots is aimed at minimizing the bias that they introduce when portraying emotional expressions, particularly for natural disasters that can be highly politicized, and so may attract bot activity via a range of state actors (Wirth et al., 2019). The tool allows the user to manually explore the influence of bots and the different thresholds applied, a possibility that is rarely provided in other approaches, as the bot filtering stage usually takes place in the pre-processing phase of a study by scholars, or presented with a single threshold.

7. In this sample dataset of New York residents, we identified hotspots of stress in census tracts of Manhattan and Brooklyn appearing after the lockdown in April 2020 (See Supplementary Figure for additional screen capture) that persisted after filtering for bot activity using Botometer. These spatial patterns could have been produced by at least two different spatial processes. First, we have not accounted for potential influencing factors in a census tract (e.g., poor socio-economic status SES, proximity to hospitals) that may drive above average stress response in these areas during lockdown. Second, hotspots can also point to spatial interaction effects in which locally relevant influencing factors in one census tract may have an effect on neighboring census tracts. Future studies could spatially integrate other data from e.g. the American Community Survey or hospital locations to address locally relevant influencing factors in regression models.

8. As Twitter datasets are being made available online to support COVID-19 research, this approach offers the potential for surveillance of negative emotions and stress for large populations over time. The use of bot detection software such as Botometer allows us to assess the likelihood of a Twitter account to be automated. This is an important component of such surveillance. The extent to which the emotions and stress observed in tweets may be a result of bot-related activity as has been previously reported in the case of e-cigs (Allem et al., 2017), and as seen in our current dataset. Twitter data use for mental health studies has been criticized for its non-representativeness, but this

criticism has largely ignored the role of bots and other automated accounts, which may substantially skew non-representativeness further. Identifying bots is one way to limit this issue, which otherwise compounds existing selection bias.

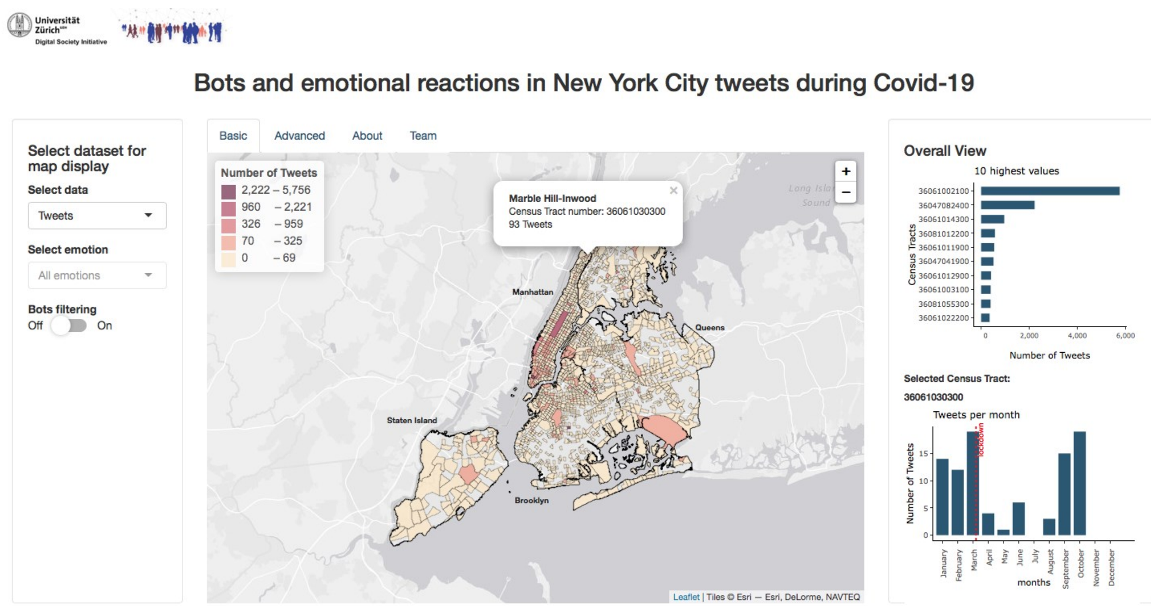
9. The current proof of concept tool as presented has several limitations, with the fixed, small dataset being the most prominent. The Twitter accounts within this study were originally identified in 2012, and the dataset size has since been reduced due to closures and suspensions of accounts, or individuals moving outside of New York City. As we are bound to a data confidentiality agreement and have a small dataset, a higher granularity of analyses is not possible for this particular data. Nevertheless, working with this smaller example dataset allowed us to more easily develop the tool itself and explore its possibilities, and a near real-time Twitter data flow could be further integrated into the existing platform. The knowledge that we pre-obtained on the population, including place of residence and demographic properties, showed great value in the rapid development of the conceptual dataframe and the proof of concept tool. In addition, we acknowledge that social media use is a natural and self-selected behaviour that is context-dependent. There are a number of important features that reflect this, of which increased/decreased activity is one prominent example. Other features are psychological or emotional processes, which provide – in combination with other features such as activity – important context-sensitive information. A further development of the current example tool would integrate additional data to provide well-needed contextual information on the socio-ecological environment (e.g. employment rate, air pollution) and more individual level information on the social media users themselves, such as sex, age, and socio-economic status, while acknowledging ethical and legal user rights. Other filtering options could also be added, such as by topic (such as covid-19) using topic modelling approaches. Such approaches could provide better insights on the spatio-temporal patterns found. Although, this should be done with caution as studies have shown that previous filtering based on hashtags introduce additional bias and loss of information (Tufekci, 2014). Furthermore, stressors may be unconscious and latent and users may not be aware that these may be related to covid-19 in the widest sense.

10. Notwithstanding the limitations, if applied to real-time, large anonymized datasets obtained using the Twitter API, the current approach would allow for the detection of “hotspots” of above average negative emotions or stress that may be signals or

precursors to changes in mental health in reaction to localized events (such as focused lock-downs, new outbreaks, or superspreader events), allowing the development of tailored supportive local measures accordingly.

Figures:

FIGURE 1: Screenshot of web-based geovisualization tool. Website available here: https://givauzh.shinyapps.io/NYC_tweets/



References

Allem, J.-P., Ferrara, E., Uppu, S.P., Cruz, T.B., Unger, J.B., 2017. E-Cigarette Surveillance With Social Media Data: Social Bots, Emerging Topics, and Trends. JMIR Public Heal. Surveill. 3. https://doi.org/10.2196/publichealth.8641

Elayan, S., Sykora, M., Shankardass, K., Robertson, C., Feick, R., Shaughnessy, K., Haydn, L., Jackson, T., 2020. The stresscapes ontology system: Detecting and measuring stress on social media, in: Proceedings of the 7th European Conference on Social Media, ECSM 2020. https://doi.org/10.34190/ESM.20.046

Ettman, C.K., Abdalla, S.M., Cohen, G.H., Sampson, L., Vivier, P.M., Galea, S., 2020a. Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. JAMA Netw. Open 3. https://doi.org/10.1001/jamanetworkopen.2020.19686

Ettman, C.K., Abdalla, S.M., Cohen, G.H., Sampson, L., Vivier, P.M., Galea, S., 2020b. Low assets and financial stressors associated with higher depression during COVID-19 in a nationally representative sample of US adults. J. Epidemiol. Community Health. https://doi.org/10.1136/jech-2020-215213

Gruebner, O., Lowe, S.R., Sykora, M., Shankardass, K., Subramanian, S. V., Galea, S., 2018. Spatio-temporal distribution of negative emotions in New York city after a natural disaster as seen in social media. Int. J. Environ. Res. Public Health 15. https://doi.org/10.3390/ijerph15102275

Gruebner, O., Lowe, S.R., Sykora, M., Shankardass, K., Subramanian, S. V., Galea, S., 2017. A novel surveillance approach for disaster mental health. PLoS One 12. https://doi.org/10.1371/journal.pone.0181233

Gruebner, O., Sykora, M., Lowe, S.R., Shankardass, K., Trinquart, L., Jackson, T., Subramanian, S. V., Galea, S., 2016. Mental health surveillance after the terrorist attacks in Paris. Lancet. https://doi.org/10.1016/S0140-6736(16)30602-X

Hswen, Y., Naslund, J.A., Brownstein, J.S., Hawkins, J.B., 2018. Monitoring Online Discussions About Suicide Among Twitter Users With Schizophrenia: Exploratory Study. JMIR Ment. Heal. 5. https://doi.org/10.2196/11483

Kušen, E., Strembeck, M., 2019. Something draws near, I can feel it: An analysis of human and bot emotion-exchange motifs on Twitter. Online Soc. Networks Media 10–11. https://doi.org/10.1016/j.osnem.2019.04.001

Kušen, E., Strembeck, M., 2018. Why so emotional? An analysis of emotional bot-generated content on twitter, in: COMPLEXIS 2018 - Proceedings of the 3rd International Conference on Complexity, Future Information Systems and Risk. https://doi.org/10.5220/0006699500130022

Lima, C.K.T., Carvalho, P.M. de M., Lima, I. de A.A.S., Nunes, J.V.A. de O., Saraiva, J.S., de Souza, R.I., da Silva, C.G.L., Neto, M.L.R., 2020. The emotional impact of Coronavirus 2019-nCoV (new Coronavirus disease). Psychiatry Res. https://doi.org/10.1016/j.psychres.2020.112915

Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F., 2018. The spread of low-credibility content by social bots. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-06930-7

Sykora, M.D., Jackson, T.W., Obrien, A., Elayan, S., 2013. National security and social media monitoring: A presentation of the emotive and related systems, in: Proceedings - 2013 European Intelligence and Security Informatics Conference, EISIC 2013. https://doi.org/10.1109/EISIC.2013.38

Tufekci, Z., 2014. Big Questions for social media big data: Representativeness, validity and other methodological pitfalls, in: Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014.

Varol, O., Ferrara, E., Davis, C.A., Menczer, F., Flammini, A., 2017. Online human-bot interactions: Detection, estimation, and characterization, in: Proceedings of the 11th International Conference on Web and Social Media, ICWSM 2017.

Wirth, K., Menchen-Trevino, E., Moore, R.T., 2019. Bots by topic: Exploring differences in bot activity by conversation topic, in: ACM International Conference Proceeding Series. https://doi.org/10.1145/3328529.3328547

Yang, K.C., Varol, O., Davis, C.A., Ferrara, E., Flammini, A., Menczer, F., 2019. Arming the public with artificial intelligence to counter social bots. Hum. Behav. Emerg. Technol. 1. https://doi.org/10.1002/hbe2.115

SUPPLEMENTARY FILE Figure

Spatial hotspots of above average stress detected by local Moran's I statistic in NYC Tweets for April 2020.