Future Geospatial Disinformation Campaigns

Lieutenant Colonel David M. Beskow
Kathleen M. Carley, Ph.D.

ABSTRACT

Social media is increasingly used as a source of data to provide situational awareness and decision support tools for world events including sporting events, democratic elections, and natural disasters. As this data is increasingly used in these scenarios, it also becomes vulnerable to manipulation. This manipulation can take several forms which have been variously explored. This paper will highlight the vulnerability of future manipulation of social media geospatial data.

I. INTRODUCTION

As social media has become the medium through which many people consume news and information[7], it has also become the marketplace for beliefs and ideas. This marketplace has recently been manipulated by highly organized disinformation campaigns by both state and non-state actors.[3] These actions generally involve manipulation of the actual network or of the information in order to gain an unfair advantage in the information marketplace.

While these disinformation campaigns are emerging to manipulate social media platforms, the same platforms and associated data are used by a variety of organizations to help understand world events. Social media data is used by financial companies[4], national security organizations[1], emergency response organizations[6], news outlets[5], and political organizations[2] to provide situational awareness and decision support tools. The known use of this data to support decision making in these events will likely increase the incentives to launch disinformation campaigns to manipulate decision making or simply to sow discord. To date, there has not been a widespread, publicized attempt to manipulate the geographic dimension of this data. This paper will highlight the future possibility of this by expanding on the innocent manipulation of this data by a Twitter bot hobbyist.

Lt. Col. David Beskow, U.S. Army, is a PhD candidate in the School of Computer Science at Carnegie Mellon University. During his career, Beskow served as an infantry leader in the 82nd Airborne Division and the 4th Infantry Division. As an FA49 operations research and systems analyst (ORSA), Beskow served as an assistant professor at West Point and as an ORSA analyst at the U.S. Army Intelligence and Security Command. Beskow's current research develops machine learning algorithms to detect and characterize online bots and the disinformation campaigns they inhabit.

II. BACKGROUND

Our team has been monitoring Twitter and other social media outlets for NATO-related conversations for several months leading up to NATO Trident Juncture 2018 Exercise, held in October and November 2018 in Scandinavia. As the largest military exercise ever held in Norway, we expected NATO-related disinformation campaigns from Russia and Russian proxies to target this event.

While monitoring the NATO-related conversation, we periodically visualized the geospatial nature of the Twitter conversation in Scandinavia. During one such investigation, we found that a bot hobbyist in Finland created a bot that tweeted the Finnish numbers while geo-locating these tweets in a uniform distribution across the longitude and latitude of the bounding box of Finland. This bot was discovered in the geo-spatial visualization provided in Figure 1. In this figure, the two-dimensional, uniform distribution across the bounding box of Finland is clearly evident.

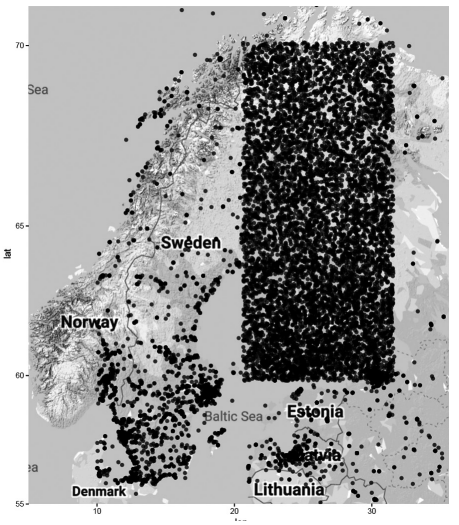


Figure 1. Map showing Finnish bot designed to locate tweets uniformly in the bounding box of Finland

Kathleen M. Carley, PhD, is a professor of societal computing in the School of Computer Science at Carnegie Mellon University, an IEEE Fellow, the director of the Center for Computational Analysis of Social and Organizational Systems (CASOS), and the CEO of Netanomics. She is the 2011 winner of the Simmel Award from the International Network for Social Network Analysis and the 2018 winner of the USGA Academic Award from GEOINT.

This bot produces a tweet once every minute that slowly goes through positive integers in the Finnish language. It has been tweeting since October 2014 and has produced 1.69 million tweets (we captured 6,734 of these in Figure 1). The very simple Python function used by this account was available on Github and is provided in Listing 1.1.

```
def random_point_in(bbox):
    '''Given a bounding box of (swlat, swlon, nelat, nelon),
     return random (lat, long)'''
    lat = random.uniform(bbox[0], bbox[2])
    lon = random.uniform(bbox[1], bbox[3])
    return (lat, lon)
```

Listing 1.1. Python Code from Finnish Bot

III. POTENTIAL FOR MALICIOUS EFFECTS

Although this account has good-natured intentions, the power of this geospatial data manipulation is clearly evident in Figure 1. In this case it was easy to clean the data since the tweets were generated by a single account in an easily recognizable rectangular pattern. A determined actor, however, could activate a dormant *bot army* to generate Tweets, and geo-locate them in a manner to either

1. Carpet the area in spatial tweets to create enough noise to mask the underlying signal of interest (i.e. true calls for help in natural disaster), rendering the underlying data useless for situational awareness or decision support (most likely).
2. Create a fake social event or fake social signal to sow discord or enable an elaborate deception operation (most dangerous).

If we consider this manipulation a type of *offensive information operation*, then we need to consider the resulting *defensive information operation*. In this case, data scientists at social media companies and government agencies would attempt to identify all accounts and content associated with the offensive disinformation operation. These efforts would seek to find any pattern in the attack, and then leverage machine learning

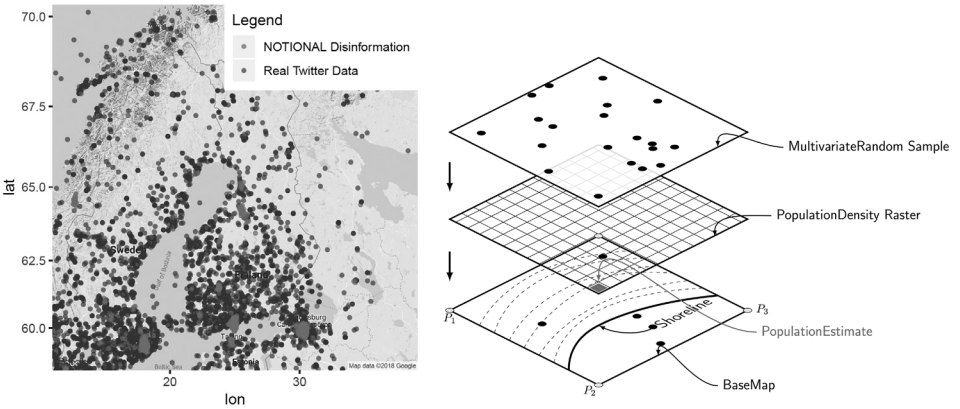
algorithms to identify malicious accounts at scale, similar to current methods that identify traditional bots[3]. They would look for any patterns in the account features (similar names, descriptions, language settings) or account activity (similar language, content, temporal correlation, etc) or geospatial distribution (easily identifiable distributions, such as the Finnish bot).

Therefore, given the known techniques used to *clean* the data, for an offensive information operation to achieve full success with geospatial disinformation, it would need to have the following characteristics:

- Leverage a large number of accounts (i.e. a *bot army*) that appear to be local to the target area (i.e. have reasonable language, time zone, and life patterns).
- Leverage data sampling and varied random distributions to create an elaborate and realistic geospatial pattern.
- Create content that blends easily with local conversation.
- Duration must be just long enough to achieve success, and, after, with accounts would blend back into the conversation.

The desired end state for an offensive geospatial disinformation operation is to cause confusion and operational paralysis, while the offensive actor maintains access to most of their *bot army.*

Note that a sophisticated actor is not confined to geometric shapes or even to simple random distributions. They could use a variety of methods to generate synthetic geo-spatial coordinates that approximate the true social media geospatial distribution. They can create these patterns by producing Gaussian *jitter* around a sample of real data (see Figure 2a) or by sampling a multivariate uniform distribution through a population density raster as illustrated in Figure 2b. Either of these would accomplish the same effect, namely creating seemingly genuine social media geo-coordinates with which they can execute their geospatial disinformation campaign.



(a) Gaussian *jitter* (b) Sampling by population density

Figure 2. Two methods that could be used to create synthetic geospatial coordinates that approximate genuine social media geospatial distributions

Given these requirements, these inevitable geospatial information operations will require a degree of sophistication to achieve operational success. With the requisite degree of sophistication and planning, these operations would be difficult to defend and would create operational confusion and possibly, paralysis. The eventual result could be that specific social media platforms and data may be rendered useless for situational awareness and decision making for key leaders in finance, emergency response, and national security.

IV. NOTIONAL GEOSPATIAL DISINFORMATION OPERATION

In this section we will illustrate the potential danger of geospatial disinformation campaigns. We will do this with a **NOTIONAL** disinformation campaign inserted into the real-world data associated with Hurricane Michael, a Category 4 hurricane which struck the Gulf Coast in October 2018. All synthetic data was inserted after the fact; our team did not create or manipulate Twitter to create this notional scenario.

In this scenario, we create a notional actor who wants to create confusion and chaos around Tyndall Air Force Base (AFB) during the hurricane and use this chaos to gain unauthorized access to the base for malicious or espionage purposes. Given that the base was in the eye of the storm, all but essential personnel were evacuated from the base, and therefore very few social media posts were emanating from Tyndall AFB during and immediately after the storm. To create the chaos, our notional actor posts numerous fake cries for help on Twitter, all geo-spatially located inside Tyndall AFB cantonment area. The notional actor would then attempt to infiltrate the base when numerous off-base first responders attempt to gain access to the base to respond to these false alarms.

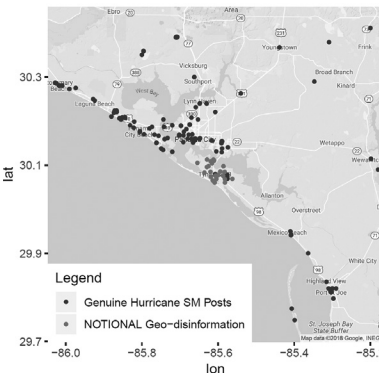


Figure 3. Notional Example of Geospatial Disinformation

In Figure 3, our team has inserted *notional* data (red) on top of the genuine hurricane related Twitter posts (blue). To create the simulated geospatial disinformation below, our team created a multivariate random normal distribution centered on Tyndall AFB and used a point-in-polygon algorithm to remove all points that were in the water. The notional actor would then attach fake calls for help to these geographic coordinates and post them in the form of a tweet or other social media post. To do this, the malicious actor could create their own content, or more likely mimic or copy the real calls for help already associated with the natural disaster.

This seemingly real surge in calls for help from the Tyndall AFB would undoubtedly cause multiple off-base first responders attempt to get base access to rescue the supposed victims. The malicious actor could then use the chaos that ensues to insert their own agents onto Tyndall AFB to sabotage or conduct espionage operations. In the notional example above, we have illustrated a targeted, geospatial, disinformation operation associated with a natural disaster. These type of information operations could be deployed in conjunction with a terrorist attack, a humanitarian crisis, or combat operations. In all cases the geospatial disinformation would create confusion and chaos, alter decision making, and in the end, render the underlying data source unreliable and unusable.

V. CONCLUSIONS

In this paper, we have highlighted how the manipulation of geospatial information in social bot disinformation campaigns can deceive and disrupt organizations who use that data for situational awareness and decision support. These geospatial disinformation campaigns may be simply trying to hide the signal in noise, or they may be trying to support an elaborate deception operation. Regardless, the initial effect will be confusion and operational paralysis. Long term strategic effects could include degraded value for large open source data (i.e. neutralization of *big data* advantage).

DISCLAIMER

The views expressed in this work are those of the author(s) and do not reflect the official policy or position of the Army Cyber Institute, the United States Military Academy, the Department of the Army, or the Department of Defense.

ACKNOWLEDGMENT

This work was supported in part by the Office of Naval Research (ONR) Multidisciplinary University Research Initiative Award N000141812108, Office of Naval Research Minerva Awards N00014-13-1-0835/N00014-16-1-2324, and the Center for Computational Analysis of Social and Organization Systems (CASOS). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the ONR or the U.S. government.

NOTES

1. M. Benigni and K. M. Carley (2016), From tweets to intelligence: Understanding the islamic jihad supporting community on twitter, In *Social, Cultural, and Behavioral Modeling: 9th International Conference, SBP-BRiMS 2016, Washington, DC, USA, June 28-July 1, 2016, Proceedings* 9, Springer, 346– 355.
2. W. L. Bennett and A. Segerberg (2012), The logic of connective action: Digital media and the personalization of contentious politics, *Information, Communication & Society* 15(5), 739–768.
3. D. Beskow and K. M. Carley (2018), Introducing bothunter: A tiered approach to detection and characterizing automated activity on twitter, In H. Bisgin, A. Hyder, C. Dancy, and R. Thomson (Eds.), *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*. Springer.
4. M. J. Culnan, P. J. McHugh, and J. I. Zubillaga (2010), How large us companies can use twitter and other social media to gain business value, *MIS Quarterly Executive* 9(4).
5. R. Goolsby, (2010), Social media as crisis platform: The future of community maps/crisis maps, *ACM Transactions on Intelligent Systems and Technology (TIST)* 1(1), 7.
6. M. Latonero and I. Shklovski (2013), Emergency management, twitter, and social media evangelism. In *Using Social and Information Technologies for Disaster and Crisis Management*, IGI Global, 196–212.