
Detecting Coordinated Behavior in the Twitter Campaign to Reopen America

Thomas Magelinski & Kathleen M. Carley
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
{tmagelin, kathleen.carley}@cs.cmu.edu.edu

Abstract

This study aims to uncover coordinated Twitter users urging people to protest to “Reopen America.” From April to September of 2020, a number of protests occurred across the United States with the aims of lifting the COVID-19 safety restrictions which prevented many people from working and recreating. The majority of these protests had a few hundred members, while the largest had an approximate attendance of 3000 people. The protests were widely regarded as dangerous due to their disregard of social distancing and mask protocols. Despite this, they were encouraged on Twitter by President Trump. A number of websites and social media groups were created to organize the protests in different states. These websites and groups had the same goal, and very similar messaging, despite there being no stated association. This raises the question: was there a coordinated effort to initiate these protest across state lines or were they organic? The United States has previously seen political protests orchestrated by outside actors unbeknownst to the protest’s attendees, indicating that the coordination theory would not be unprecedented. This work looks at a small sub-set of the problem: detecting sets of Twitter users inorganically coordinating to further the anti-lockdown campaign.

We define coordinated behavior on Twitter between two users as many instances of a Tweet-behavior, i.e. tweeted hashtag, tweeted link, user mention, etc., within a small predetermined time window. This behavior is inconspicuous, as it can only be discovered through comparison of users. Previous works have used a similar definition of coordination, and proposed a number of behaviors that may be coordinated [2, 3, 5]. It is worth noting that non-network based approaches have been explored recently, though, the effectiveness of this method in detecting multi-view coordination remains to be seen [4]. Instances of coordination between users are recorded as weighted edges in a network, allowing for automated discovery of coordinating communities.

Our work builds on the existing literature in three ways. First, we give a scalable algorithm for constructing the coordination networks which is able to capture all links, while the existing method was only able to capture 50% of the links. Second, we introduce a weighting scheme similar to TF-IDF, in order to find more meaningful coordination than simple spam accounts. Third, we utilize multi-view clustering to account for multiple types of coordination at the same time. Previous works acknowledge the importance of different types of behaviors, but did not give a methodology for analyzing them simultaneously.

The dataset we analyzed was collected by searching for keywords and hashtags with “openup”, “reopen”, “operationgridlock”, and “liberate”. Additionally, all of the US State abbreviations were appended to each of the search terms, i.e. “liberateNY.” The dataset was collected from 4/1/2020 to 6/22/2020. In total, the dataset contains roughly 3.6 million unique users and roughly 9.9 million tweets. We further only consider users who tweet a hashtag or a URL 5 or more times in the dataset. This filtering lowers the amount of connections in the coordination network happening simply due to random chance.

In prior work, temporal coordination networks were constructed through bipartite projection [3]. In this procedure, the timeline is segmented based on a fixed window. Then, a higher-order target node set is constructed by taking the Cartesian product of the original nodeset and the set of time segments. That is, an example node might be (#reopen, 1). Where “1” indicates the first time segment.

If our window was 15 minutes, all hashtags within the first 15 minutes of collection would have "1" as their second higher-order node element. Those in the next 15 minutes would have a second higher-order node element of 2, and so on. This framework is easy to implement, and results in a bipartite network from users to behavior-time segment pairs. Then the network can simply be folded to obtain a user-to-user network, where connections are weighted by the number of behavior-time segment pairs two users shares.

There is a problem with this approach, however. Consider, again, the 15 minute time segment. If User A tweets #reopen at minute 14, and User B tweets #reopen at minute 16, no link will be created between the users, despite their tweets being only 2 minutes apart. In fact, under a uniform distribution of behaviors, this approach will only find 50% of the connections between users.

A simple sliding window approach will successfully obtain all of the connections. However, a direct implementation of this strategy will be intractable for large datasets and/or long time windows, because the algorithm scales quadratically with the number of tweets in a window.

We obtain all of the connections in a scalable manor as follows: The sequence of tweets containing a specific behavior is obtained and sorted in chronological order. For example, all of the tweets containing "#reopen." Then, a sliding window the size of our time segment is moved across the sequence of tweets, recording pairs of users falling in the window by adding them to the edgelist if they have not been seen or incrementing their weight if they have. The process is then repeated for all of the remaining behaviors, resulting in a weighted user to user coordination network. By studying each behavior individually, the number of tweets in any given time window is drastically reduced, making the quadratic behavior of the sliding window much less problematic.

Finally, we also consider an alternate weighting scheme. In the above procedure, many links will be created between users of popular hashtags simply because of their popularity. However, if a set of users all tweet a hashtag many times *before* it goes viral, that is a very important behavior we would like to find. To satisfy both of these constraints we normalize the weighting increment by the number of tweets falling within the current time window. Thus, a shared popular behavior might be weighted heavily, but only if it is not popular at the time in question.

This process is applied to hashtags and URLs with a 5-minute time window, resulting in a multi-view coordination network between users. We first apply multi-view clustering to find highly coordinated communities [1]. The multi-view cluster with highest density was a small number of Scientology users, spreading information about Scientology. Deeper analysis of the multi-view approach showed that the hashtag and link views are quite different. Much of the structure in the link coordination network was based on users repeatedly sharing a pro-lockdown article from the Washington Post titled "34 days of pandemic: Inside Trump's desperate attempts to reopen America." This article outlines the failures of the administration's pandemic response. Many hubs in the link-view coordination network were repeatedly replying to tweets with this story, presumably to signal evidence-based opposition to the reopen movement. Other instances of clusters in the link-view were constructed around satire posts, which have tendencies to gather many tweets in a short period of time, which can look like coordinated behavior. One reason that the link view is so different from the hashtag view is the collection strategy. By selecting tweets based on keywords and hashtags, we are restricted to link-based coordination which contains the keywords in the URL, or is posted with a hashtag in our search set.

The hashtag-view, however, shows a much more complete story. First, the coordinated hashtag-based clusters show far more anti-lockdown users than the link-based clusters, specifically around the hashtags #reopenNC, and #reopenCalifornia. Looking further into the long-term behavior of the users exhibiting highly-correlated in the reopen dataset shows that some of the most central users in the coordination network are currently on Twitter spreading misinformation about the 2020 election results. Specifically, tweeting pictures of tables with incorrect voter registration numbers, and using these incorrect numbers to claim that key swing states like Wisconsin received more votes than register voters, which it did not.

While a number of the most-coordinated users did appear to be spreading the same messaging, there are instances of strong links between users with opposing viewpoints. One anti-protest user is a medium-sized hub in the hashtag view, with many pro-protest neighbors. This is in agreement with the literature suggesting that use of a hashtag is not necessarily an endorsement of that hashtag, regardless of how strong the stance of that hashtag may be. This highlights the need for more

sophisticated coordination detection methods, which will be carried forward in two ways. First, through the use of targeted sentiment to separate the use of a hashtag in different ways. The second and more general part of future work, is the use of higher-order behaviors to detect more precise coordination behaviors. For example, if a user tweets a specific hashtag *and* mentions a specific user in the same tweet, that is a more specific behavior than a hashtag alone, and is thus a more targeted type of coordination. These methods of narrowing down which behaviors can count as potentially coordinated should lead to more meaningful detection methods.

Acknowledgments

This work was supported in part by the Office of Naval Research (ONR) Award N000141512797 Minerva award for Dynamic Statistical Network Informatics, and the Center for Computational Analysis of Social and Organization Systems (CASOS). Thomas Magelinski was also supported by the IDEaS Center as a Knight Fellow. 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.

References

- [1] I. J. Cruickshank. *Multi-view clustering of social-based data*. PhD thesis, Carnegie Mellon University, 2020.
- [2] L. Nizzoli, S. Tardelli, M. Avvenuti, S. Cresci, and M. Tesconi. Coordinated behavior on social media in 2019 uk general election. *arXiv preprint arXiv:2008.08370*, 2020.
- [3] D. Pacheco, P.-M. Hui, C. Torres-Lugo, B. T. Truong, A. Flammini, and F. Menczer. Uncovering coordinated networks on social media. *arXiv preprint arXiv:2001.05658*, 2020.
- [4] K. Sharma, E. Ferrara, and Y. Liu. Identifying coordinated accounts in disinformation campaigns. *arXiv preprint arXiv:2008.11308*, 2020.
- [5] L. Vargas, P. Emami, and P. Traynor. On the detection of disinformation campaign activity with network analysis. *arXiv preprint arXiv:2005.13466*, 2020.