Proceedings of the Annual General Donald R. Keith Memorial Conference
West Point, New York, USA
May 4, 2023
A Regional Conference of the Society for Industrial and Systems Engineering

Measuring the Reach and Impact of Information Warfare

Eda Hopkins, Jemel Jones, Joshua Mettling, Spencer Weisgram, Oliver Wood, and David Beskow

Department of Systems Engineering
United States Military Academy
West Point, NY 10996

Corresponding author's Email: eda.hopkins@westpoint.edu

Author Note: The views expressed herein are those of the authors and do not reflect the position of the United States Military Academy, the Department of the Army, or the Department of Defense.

Abstract: Information is a critical component of national power, with new information narratives emerging daily. While some narratives deserve attention by leaders, others lack enough traction to warrant attention. In the context of Information Warfare, leaders must understand key metrics of emerging narratives to determine their importance. Faced with the problem that leaders do not know if a social media story matters, we created a dashboard to aid leaders in identifying Information Warfare campaigns and their impacts. The methodology consisted of measuring four metrics of distinct narratives: reach, virality, propagation, and organic versus inorganic activity. Taken together, these metrics show how narratives spread, the speed of spread, authenticity, and impact. When packaged together in dashboard visualizations, these metrics enable informed decision making in strategic environments in the information dimension.

1. Introduction

New narratives and stories emerge across the information dimension daily, some of which have major implications for a nation state. Narratives may be in response to an organization's decisions, such as the United States' decision to kill or capture Osama Bin Laden. Other narratives may involve a campaign by a competitor, such as Iran's #CalExit campaign designed to encourage California to leave the United States. These are just a few examples of Information Warfare (IW). As stories and their resulting discussion evolve, some stories gain traction and spread, while others simply fizzle out. This paper provides quantitative metrics to apply to emerging narratives in the information dimension to help leaders determine which stories require attention and which can be largely ignored. We sought to measure the reach and impact of IW campaigns to aid leaders in determining whether they matter. The paper consists of four focus areas: propagation, virality, reach, and organic versus inorganic activity. These metrics provide leaders with a holistic view of campaigns. To understand information flow through social media networks and how it impacts populations, we conducted literature reviews on five topics in the context of social media: network propagation, virality, marketing, measures of effectiveness (MOEs), and human behavior.

2. Literature Review

"Network propagation" is a set of algorithms that integrate input data across nodes in a network (Charmpi, Chokkalingam, Johnen, & Beyer, 2021). Three concepts encompassed in propagation are most relevant to the paper. First, centrality measures identify which nodes form a network's core. Nodes with higher centrality measures have greater reach and influence and are indicative of network cores (Jiang, Wen, Yu, Xiang, & Zhou, 2017). It is important that we identify core nodes in this project, for they are the disinformation sources. Second, it is important to determine whether nodes are "real" users by distinguishing them from bots. Bots have the highest tweet publication rate, exhibiting strong uniformity between real and fake content to disguise fake content (Balestrucci & De Nicola, 2020). Third, content type helps predict propagation scale. For example, using hashtags and possessing indicators of "trustworthiness" (ex. containing more text) increase propagation scale (Li et al., 2020). By knowing which factors increase the likelihood of a tweet to spread, we can verify our results by asking, "given the content and nature of a post, does it make sense that it spread so far?" These three findings enhance our project understanding by illustrating that nodes and edges form a social media network; trending content surrounds one or more core nodes; humans and bots differ most notably via post frequency; and some posts will likely spread more based on their content.

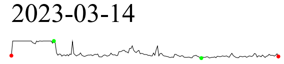
Propagation refers to how content spreads; virality refers to the tendency to spread. Virality is the volume and speed of message diffusion (Tsugawa & Ohsaki, 2017) and can significantly influence public opinion, so it is critical to understand how to measure virality. Degree of social impact, viral users, and speed of message diffusion all comprise virality. Viral media has a greater influence on public opinion, for the public tends to believe that virality indicates seriousness and trustworthiness (Lee-Won, Abo, Na, & White, 2016)). The most viral users include information aggregates, mainstream media, satirical users, and political activists (Hoang, Lim, Achananuparp, Jiang, & Zhu, 2011). Moreover, negative messages tend to have higher diffusion rates compared to positive and neutral messages (Tsugawa & Ohsaki, 2017). Overall, It is essential that we develop a measure of virality to help leaders measure and gauge this important metric.

Next, to understand the strategic mindset that actors employ in IW, we transition to marketing. Business marketing relies on proper utilization of social media to promote brand identity and longevity. IW actors are analogous to business marketers: propaganda must be "branded" and promoted. Thus, social media marketing strategies are relevant to our project. Marketers engage in active and passive marketing techniques. In the passive approach, they utilize social media as a market intelligence source, obtaining analytics online rather than engaging with customers directly (Constantinides, 2014). Marketers also consider what factors attract consumers to certain brands, known as motivation scale developments. One study concluded that customer motivations fall into five categories: brand affiliation, investigation, opportunity seeking, conversation, and entertainment (Enginkaya & Yılmaz, 2014). Twitter discourses tend to fall into these categories as well.

Marketing is an iterative process where content creators continually assess the effectiveness of their branding and campaigns via MOEs. The primary MOEs of social media marketing campaigns include analyses of user sentiment, keywords, and engagement metrics (likes, comments, and follows). Data analyst Chris Murdough emphasized the importance of consolidating outputs from data sources and analytical tools into a comprehensive dashboard. These can be used to improve how brands increase the effectiveness of marketing to customers through social media (Murdough, 2009). In researching MOEs, we began to shape our vision of a final dashboard, serving as tool to show whether a campaign (analogous to brand) was successful.

Finally, social media narratives, whether true or not, can permeate and influence adolescent and young adult populations; thus, it is critical to understand the effects of social media on behavior. Case studies have found positive relationships between social media consumption and political and protest behaviors(Valenzuela, Arriagada, & Scherman, 2012). Frequent consumption of social media content increases users' knowledge of public issues and social movements, thus enabling greater political participation (Valenzuela, 2013). As we evaluate social media metrics, we must continually evaluate the impact of the virtual information world on the real and tangible.

Table 1: Summary of Data Used for Evaluation

Metric	Data Related to International Security
Tweets	1,445,063
Users	350,569
earliest date	2022-08-25
latest date	2023-03-14
Daily Count (AUG22 - MAR23)	
% of replies	3.0%
% of retweets	34.8%
% of quotes	10.4%
% of original	51.7%
median/mean friends	74 | 822
median/mean followers	33 | 8739
% of likely bots	61.5% > 0.6 prob | 40% > 0.75 prob

3. Data

We chose to focus our research on Twitter Data associated with multi-actor security related discourse in Asia. This data was acquired from the Twitter streaming application programming interface (API) from August 2022 through March of 2023. This data had multiple narratives from multiple actors. A detailed description of the data is provided in Table 1. Of interest is the large decrease in daily volume in the Fall of 2022. This is caused by Twitter suspending 92% of the accounts participating in one of these information campaigns. Our analysis indicates that these accounts were likely bot and/or troll accounts associated with one of the actors.

Also of interest is the Daily Count trend line which guided identification of potential propaganda campaigns. Large spikes in tweet volume indicate trending stories (celebrities, natural disasters, etc.). Inferring that actors likely leverage these trends to propagate other potentially harmful narratives, we utilized daily count trends to focus on particular time frames in our analysis. This data gave us a number of narratives and information campaigns with which to evaluate importance.

4. Methodology

We set out to measure and portray *reach*, *virality*, *propagation*, and *organic vs. inorganic activity*. This paper will focus on defining algorithms to measure *reach*, *virality*, and *propagation*. We will leverage the "bot-hunter" model (Beskow & Carley, 2018) to measure organic vs. inorganic behavior.

Actors leverage both networks and narratives to conduct information operations (Beskow & Carley, 2019). Our analysis here is primarily focused on understanding the network aspects of these operations. *Reach*, *virality*, and *propagation* are all network measures. Computational methods to evaluate the semantic narrative are important but beyond the scope of this paper. We will define algorithms and metrics for *reach*, *propagation*, and *virality* below.

4.1. Measuring Reach

One of the most important metrics for evaluating the importance of emerging narratives is the size of the audience. The fundamental question is "How many people does this narrative potentially impact?" In our analysis, we measured two types of accounts related to a distinct online narrative: 1) participants and 2) potential audience.

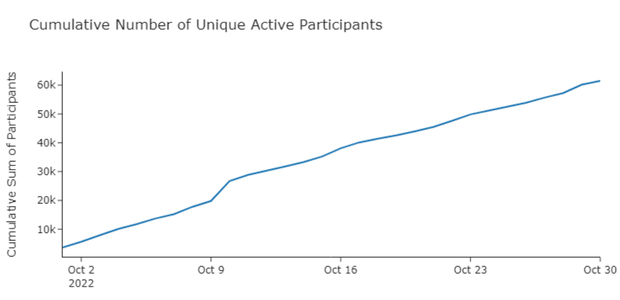
Participants are those individuals/accounts that are actively participating in the conversation at hand by producing or propagating (retweet, reply, or quote) content related to the distinct narrative. Our method measures and visualizes the cumulative growth of participation over time. This algorithm finds the first instance a unique user joins the conversation, and then calculates the cumulative sum of these unique users joining the conversation. Figure 1 demonstrates the computationally efficient algorithm for measuring cumulative participation of Twitter. This algorithm uses data frame operations to avoid a computationally expensive nested loops.

Algorithm 1 Measuring Cumulative Participation

```
parse Twitter JSON to dataframe
sort temporally in ascending order
drop duplicates by user ID, keep first instance
calculate cumulative sum, weight = 1
compute max value by temporal bin
```

Algorithm 2 Measuring Cumulative Potential Reach

```
parse Twitter JSON to dataframe
sort temporally in ascending order
drop duplicates by user ID, keep first instance
calculate cumulative sum, weight = follower count
compute max value by temporal bin
```



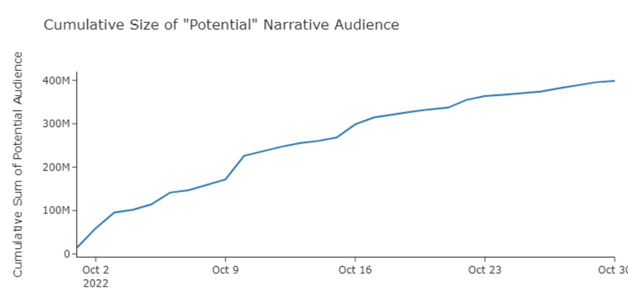


Figure 1: Algorithms to measure participation (left) and potential reach (right)

A slight change to the algorithm for measuring cumulative participation will enable measurement of *potential reach*. *Potential reach* is the ceiling for audience size; an audience is the individuals who have seen and read a narrative. In Twitter and other social media platforms, the content of those an account follows populates that account's feed. Therefore, the potential audience of a narrative is the sum of the followers of participants. Figure 1 demonstrates how to measure and visualize the cumulative sum of the potential audience by weighting the participation algorithm by follower count. In this case we weight the cumulative participation algorithm by *follower count* to create a computationally efficient method to compute potential reach (or potential audience size).

The one limitation to this approach is that the cumulative sum of participant followers may contain duplicate accounts. It is likely that an account follows more than one participant in the conversation. By summing the followers of participating accounts, we over count these individuals. However, given current rate limits on social media follower data, removing duplicates is not computationally feasible. Despite the duplicates, this metric still provides a "ceiling" for audience size, making it an insightful metric with minimal impact on leaders' decision making.

An important evaluation of a narrative is to compare the graph of participation to the graph of potential audience. A large increase in potential audience that is not associated with a corresponding spike in participation indicates that some accounts with a large number of followers (celebrities, politicians, athletes, etc) have joined the conversation.

4.2. Measuring Virality

Karen Nelson-Field defines virality as "...the phenomenon of Internet users passing along messages, images, and videos at an exponentially growing rate through their social networks, much like a virus spreads from person to person." (Nelson-Field, 2013). While *reach* focuses on the size of the participants and size of the audience, *virality* seeks to understand the speed and acceleration of the narrative. We will measure and visualize the velocity and the acceleration of the narrative in time.

Velocity measures a change over time. In our case we will measure the change in volume (of posts) over time. Velocity is inherently non-negative. This social media analysis measures and visualizes velocity by taking the count of posts (in our case Tweets) over time.

It is also important to measure and visualize the *acceleration* (or *deceleration*) of a narrative. Acceleration is the change in velocity over time. Computationally, acceleration is the difference in the ordered count of posts by temporal bin. The acceleration of a narrative can be negative (indicating that the narrative is decelerating or slowing down). Figure 3 provides a five week look at narrative acceleration. The spikes in Figure 3 indicate the rise and fall of Tweets throughout the day, affected by work/rest or sleep/wake cycles. When viewed over a larger time period, steeper spikes indicate abnormally rapid acceleration, which typically follow a major event occurring. We also see that during the highest spikes the velocity and jump up to 2000 tweets per hour.

4.3. Measuring Propagation

Propagation is the measure of a social media message or narrative spreading from one user/account to the next. Danah Boyd and Nicole Ellison define propagation as "..the degree to which information is disseminated through a network, as indicated by the number of people who receive the information, the frequency with which they receive it, and the extent to which they are connected to one another" (Boyd & Ellison, 2007). The following methodology visualizes conversation network over time to help leaders understand narrative propagation.

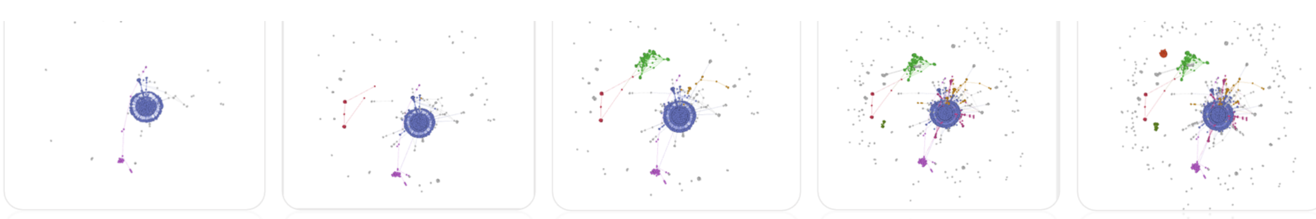


Figure 2: Dynamic Gephi File displaying network propagation of two Taiwan-China campaigns

We developed a method to create the multi-modal conversation network of Twitter (edges represent *retweet*, *reply*, *quote*, and *mention* actions) and visualize it over time. We developed Python code that creates this conversational network and then saves the network architecture and timestamps in a way for Gephi to show the cumulative narrative evolution over time. Figure 2 shows snapshots of a dynamic network evolving over time. We see the narrative move organically at first, then spurred

ISBN: 97819384964-4-8

by bots it jumps into some new communities. The nature by which the bots nearly simultaneously appear (in the dynamic network) indicate that they are not human users. Moreover, the dynamic network enables viewers to glean how quickly content originating from bots reach other users.

4.4. Measuring Organic vs. Inorganic Behavior

Leaders must distinguish between organic human online behavior and artificial bot behavior. A social media narrative with high reach, virality, and propagation metrics could be artificial with little human involvement or impact. We therefore want to measure the level of inorganic activity in any narrative. For our purposes we used the "bot-hunter" model (Beskow & Carley, 2018) to measure organic vs. inorganic behavior. This is a supervised machine learning model trained on a variety of data, to include labeled bot data and suspended bot accounts. The metrics of bot activity are provided in Table 1. We recommend that all other metrics are evaluated through the lens of organic vs inorganic metrics. For example, telling a leader a narrative has 100K participants that are 80% bot and 20% human if very different than telling a leader that a narrative has 100K participants that are 20% bot and 80% human.

5. Results and Conclusion

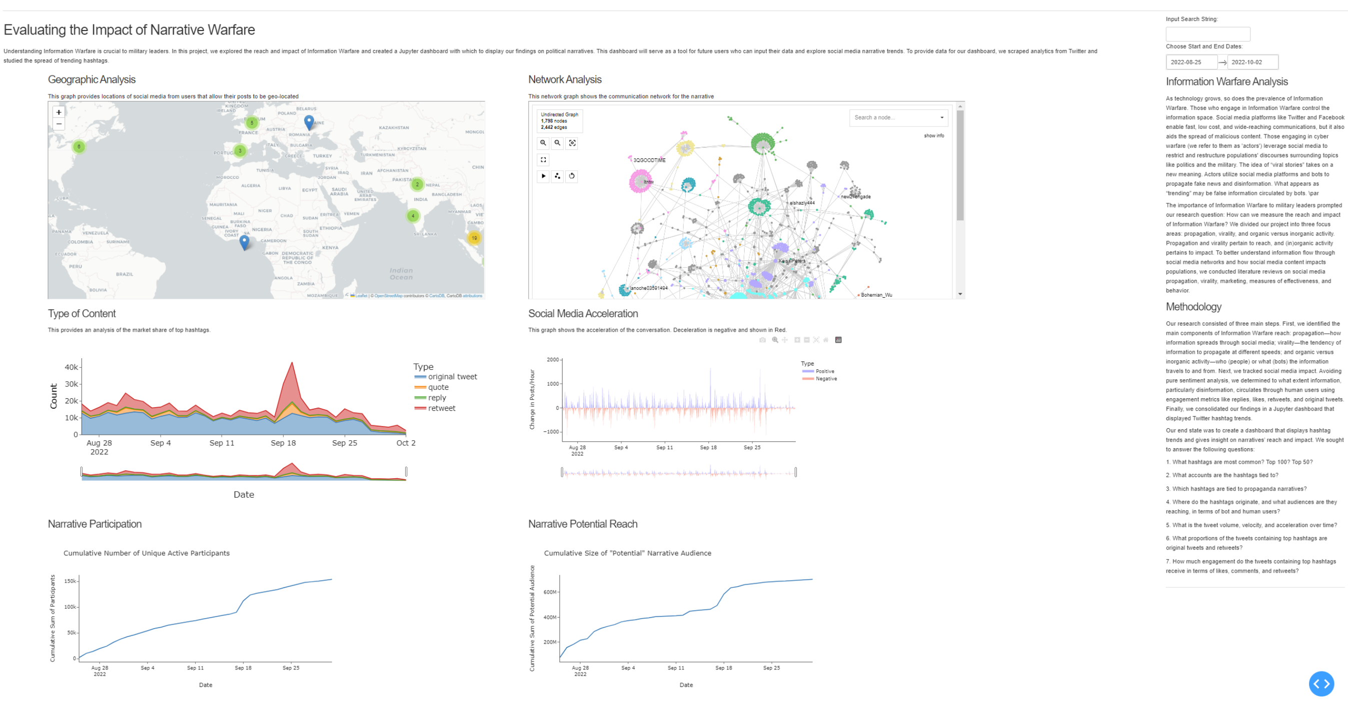


Figure 3: Screenshot of Dashboard Developed to Display Information Warfare Metrics

We combined the analysis into a comprehensive dashboard to evaluate distinct information narratives. This dashboard was developed with Plotly Dash as a Jupyter Dashboard and is available to data scientists to deploy inside of JupyterHub based Data Science Environments or to deploy as a stand-along web application. In addition to the metrics discussed here, we also provided an interactive map to identify those Tweets that are geolocated (< 1% of all tweets are geolocated). A screenshot of this application is provided in Figure 3.

Over the course of the study we monitored and evaluated several distinct narratives related to international security using the metrics of propagation, virality, and reach. We calculated the reach, virality, and network propagation metrics for each of the narratives. We found that some of them had a moderate impact and required leader attention. Others did not seem to get traction or seemed to lack human involvement (were largely inorganic bot conversations). These algorithms and related metrics are critical to help organizational leaders evaluate whether an emerging narrative is of strategic importance and deserving of their attention.

In summary, this research effort developed specific methods and algorithms to measure and visualize propagation, virality, and reach. We created a comprehensive dashboard for leaders to leverage these metrics on distinct social media narratives. Future work includes categorizing the types of audiences and participation as well as improvements on the organic vs. inorganic measurement methods. This research effort improves the ability of leaders to evaluate the importance of discrete narratives in the information dimension.

6. References

Balestrucci, A., & De Nicola, R. (2020). Credulous users and fake news: A real case study on the propagation in Twitter. In *2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)* (p. 1-8).

Beskow, D. M., & Carley, K. M. (2018). Bot conversations are different: Leveraging network metrics for bot detection in twitter. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 825–832).

Beskow, D. M., & Carley, K. M. (2019). Social cybersecurity: An emerging national security requirement. *Military Review*, *99*(2), 117-127.

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-mediated Communication*, *13*(1), 210–230.

Charmpi, K., Chokkalingam, M., Johnen, R., & Beyer, A. (2021). Optimizing network propagation for multi-omics data integration. *PLoS Computational Biology*, *17*(11), e1009161.

Constantinides, E. (2014). Foundations of social media marketing. *Procedia-Social and behavioral sciences*, *148*, 40–57.

Enginkaya, E., & Yılmaz, H. (2014). What drives consumers to interact with brands through social media? A motivation scale development study. *Procedia-Social and Behavioral Sciences*, *148*, 219–226.

Hoang, T.-A., Lim, E.-P., Achananuparp, P., Jiang, J., & Zhu, F. (2011). On modeling virality of twitter content. In *International conference on asian digital libraries* (pp. 212–221).

Jiang, J., Wen, S., Yu, S., Xiang, Y., & Zhou, W. (2017). Identifying propagation sources in networks: State-of-the-art and comparative studies. *IEEE Communications Surveys Tutorials*, *19*(1), 465-481.

Lee-Won, R. J., Abo, M. M., Na, K., & White, T. N. (2016). More than numbers: Effects of social media virality metrics on intention to help unknown others in the context of bone marrow donation. *Cyberpsychology, Behavior, and Social Networking*, *19*(6), 404–411.

Li, L., Zhang, Q., Wang, X., Zhang, J., Wang, T., Gao, T.-L., … Wang, F.-Y. (2020). Characterizing the propagation of situational information in social media during covid-19 epidemic: A case study on weibo. *IEEE Transactions on Computational Social Systems*, *7*(2), 556-562.

Murdough, C. (2009). Social media measurement: It's not impossible. *Journal of interactive advertising*, *10*(1), 94–99.

Nelson-Field, K. (2013). *Viral marketing: The science of sharing*. Oxford University Press.

Tsugawa, S., & Ohsaki, H. (2017). On the relation between message sentiment and its virality on social media. *Social Network Analysis and Mining*, *7*(1), 1–14.

Valenzuela, S. (2013). Unpacking the use of social media for protest behavior: The roles of information, opinion expression, and activism. *American behavioral scientist*, *57*(7), 920–942.

Valenzuela, S., Arriagada, A., & Scherman, A. (2012). The social media basis of youth protest behavior: The case of Chile. *Journal of Communication*, *62*(2), 299–314.