Thesis Proposal
Characterizing China's Information Campaigns

Charity S. Jacobs

April 20, 2023

Software and Societal Systems Department
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Kathleen M. Carley, Chair, Carnegie Mellon University
Patrick Park, Carnegie Mellon University
John Chin, Carnegie Mellon University
Matthew Benigni, U.S. Army Futures Command

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Societal Computing.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship (DGE 1745016), Department of Defense Minerva Initiative (N00014-15-1-2797), and Office of Naval Research Multidisciplinary University Research Initiative (N00014-17-1-2675). Any opinions, findings, conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation, the Department of Defense, or the Office of Naval Research.

Abstract

The information environment encompasses various organizations, individuals, and systems that deal with information, with three interconnected dimensions: physical, informational, and cognitive. Today's internet plays a crucial role in the information environment, where information can be disseminated widely through viral tweets, controlled via state censor protocols, or suppressed through government internet censorship. State actors are leveraging social networks for information campaigns, making it imperative to understand who is disseminating what to whom. However, existing research on state-sponsored campaigns, particularly related to China's activities, faces challenges due to proprietary data restrictions, resource constraints, and the anonymous nature of many public forums, resulting in a gap in identifying state-sponsored inauthentic behavior on social media.

To address this research gap, my thesis leverages computational and network science methods with political and sociological groundings to identify and characterize state actor campaigns within the information environment. At the singular campaign level, I use the network structure of a campaign to identify key actors and their functions across different types of information campaigns. I then employ information theory methods to analyze how China's dialogue towards regional countries in the area has shifted over time, examining the aggregated effects of campaign narratives. Furthermore, I expand the research by utilizing a social cybersecurity framework to map out how the People's Republic of China (PRC) uses information maneuvers to shape the information environment. To account for the technical means that state actors may use to manipulate web browser results, particularly for websites that are frequently used in information campaigns, I employ network structures of China's top website domains to identify common PRC search engine optimization techniques that artificially boost a given site's browser visibility. Lastly, I aggregate my findings by expanding a decision-making game designed for analysts who are studying how different types of actors manipulate the social media domain. I integrate key measurements in the synthetic creation of a PRC information campaign and validate generated data. These contributions enhance the understanding of state-sponsored information campaigns and provide valuable tools to researchers in the field.

Contents

Chapter 1

Introduction and Background

1.1 Thesis Research Goals

Information operations on social media platforms have gained significant attention in recent years due to their measured negative impacts around events of great or grave importance, such as national elections, protests, the COVID-19 pandemic, and conflicts such as Russia's ongoing war in Ukraine. A broad spectrum of actors wages these operations to influence and manipulate a targeted audience towards a desired outcome, potentially changing how they act, think, or engage with an idea or belief.

The increasingly connected way that society socializes, creates, and shares information has led to the unprecedented speed of information and misinformation propagation. State actors increasingly exploit and manipulate social media domains to promote their self-interest through computational propaganda. The attribution of these campaigns poses challenges and frequently is not feasible with the data available to researchers. To muddy the waters further regarding understanding *who* is doing *what*, the information environment is a complex ecosystem consisting of an "aggregate of individuals, organizations, and systems that collect, process, disseminate, or act on information."

The field of computational analysis on state-actor activity continues to grow in tandem with the growing awareness around information campaigns. However, the research community frequently stops analyzing inauthentic activity as it occurs on a single platform, leaving a more extensive landscape where actors can move from one social media platform to the next. This thesis argues for a computational framework that can be applied to the information environment to identify and characterize state-sponsored information campaigns at a micro-level and empirically build off how a state-actor wages information operations to understand larger narratives toward a desired end-state. We specifically analyze the People's Republic of China (PRC) in how they propagate narratives and propaganda, how those narratives target specific audiences, and ultimately how China's narratives and stories migrate and reach people through search results targeting the international community.

1.2 Literature Review

1.2.1 Information Operations on Social Media

Social media platforms such as Twitter have become beakers for experimentation by state actors against their own populations and external audiences, with evidence of organized campaigns found in over 80 countries in a 2021 survey [18]. These campaigns are engineered to imitate organic human behavior to persuade, manipulate, coerce, or crowd out targeted online audiences through information tactics such as distorting online discussion, bridging together, and exciting users towards a viewpoint or topic [11]. Oxford University researchers coined "computational propaganda" to describe how inauthentic methods, such as automated bots and human curation, can algorithmize and shape public opinion [81].

Inauthentic activity often comprises bot accounts controlled by automation software, troll farms that use professionalized groups coordinating activity, and puppet accounts that use fictitious personas. These, among others, have been well-documented methods for disseminating information and disinformation by imitating human social media interactions [33, 45, 60, 74]. Bots have been shown to effectively change public opinion and shift narratives on social media, creating a clear incentive for bad actors to manipulate the information domain [8, 13]. This fact is aided mainly due to the cost efficiency of bots versus paid troll users, paired with social media users' poor ability to determine if another account is authentic or a bot [31]. Additionally, bots are useful as an amplification device in spreading and disseminating information campaigns throughout a social network more efficiently than humans [81]. Social media users are easily swayed by information cascades or social contagion in which users may adapt their behavior based on how they perceive others are behaving or thinking a certain way [9]. Within the Twitter domain, tweets are easily manipulated via the retweeting function, which bot automation efficiently exploits [3]. While Twitter periodically takes down bot networks, especially those associated with state-sponsored activity, bots constantly adapt to circumvent bot detection measures [57].

1.2.2 China's bureaucracy for propaganda

China's underlying framework for information warfare draws significantly from its long history of developing and disseminating propaganda. Authoritarian countries have been shown to enable centralized power to inform and manage information campaigns more efficiently, and China may be the best example of this [17]. We use the philosopher Jason Stanley's definition of *propaganda* to describe the informational context of these campaigns, where it "fundamentally involves political, economic, aesthetic, or rational ideals, mobilized for a political purpose, used to either prop up or erode some ideal [76]."

China's ruling party, the Chinese Communist Party (CCP), maintains its organizational capability for propaganda thru a structured hierarchy of various departments, the main offices of which are the Central Propaganda Department (CPD) and the United Front, managed under the CCP's Central Committee, in addition to the Ministry of Foreign Affairs (MFA) under the State Council Information Office (SCIO) [29]. The CPD controls all media outreach and oversight of all propaganda systems, including oversight of China's largest media companies, such as China Global Television Network (CGTN) and Xinhua News. The United Front was developed for

soft power influence, specifically charged with increasing the CCP's support among domestic ethnic groups, political, social, and economic elites, social media influencers, and the Chinese international diaspora [29, 30]. The MFA serves as the general bureaucracy that liaisons with foreign governments and heavily participates in *public diplomacy*, which entails one country engaging and communicating with the population(s) of other countries in order to strategically communicate narratives and extend its soft power influence [41]. It is in large part through this organizational structure, disciplinary measures, and strict censorship and surveillance that the CCP maintains power so thoroughly in China [26].

China's propaganda machine operates in tandem with a robust system of legislation and technological components to censor its domestic information sphere. Perhaps most symbolic of China's iron-clad control of domestic discourse is the "Great Firewall" or the Golden Shield Project, completed in 2008 and designed to censor mainland China's access to foreign websites which are perceived to host negatively influential content to include politically sensitive information, pornographic material, and online gambling content [25]. The Great Firewall uses IP and domain name blocking techniques, keyword censorship, VPN detection, and increased penalties for VPN manufacturers to control the flow of information into China [25]. Most notably, for this thesis, the Chinese public cannot access the most popular websites worldwide, including social media platforms such as Facebook, Twitter, and Youtube, and news sites such as the New York Times, the Washington Post, and Fox News.

1.2.3 China's approach to information operations

Much of what is known about China's information campaigns is derived from social platforms releasing data on inauthentic activity network takedowns, observational studies, and qualitative research based on original Chinese government documents. China's information campaigns have typically promoted pro-China rhetoric and attacked perceived state enemies [70, 78]. Twitter suspended hundreds of thousands of accounts between 2019 and 2020, where subsequent analysis revealed these accounts were promoting PRC stances towards discussion around the Hong Kong protests and COVID-19 [61, 78].

Further analysis of China's sock puppet accounts reveal the use of both newly created and dormant accounts that displayed poorly developed account bios with very few followers, where batches of accounts are created on the same day with similar naming conventions [61]. Additionally, China has traditionally avoided using memes that often take on a life of their own and potentially undermine centralized messaging control [12]. Extensive computational journalism by media outlets like the New York Times, ProPublica, and the Australian Strategic Policy Institute has coordinated with Twitter in analyzing these networks. Key findings include evidence of the Chinese government outsourcing Twitter campaigns to private companies, using a layered ecosystem of state-sponsored accounts and influencers/foreign voices, and the prolific use of fake accounts for amplification [62, 63, 72].

Past research, specifically on China's MFA and Diplomatic Twitter accounts, determined that information diffusion occurred primarily from key state-sponsored outlets as a function of China's centralized censorship policies. However, this research did not explore other elements of information diffusion within the network beyond state-sponsored Twitter accounts [45]. Additionally, this research was conducted when there were only 14 state-sponsored accounts com-

pared to the present count of over 300. More recent research from Oxford University found that China's Twitter accounts are heavily amplified by a small number of super-spreading accounts that were later suspended. However, this research did not account for bot accounts or the type of content being spread [73].

1.2.4 Social CyberSecurity's Role

The emerging field of social cybersecurity plays a critical role in understanding and analyzing information operations that attempt to shape people's engagement with ideas and beliefs, using both social and computational frameworks [21]. The types of information communication that are of interest to researchers, such as propaganda, misinformation, and disinformation, do not exist in isolation from the tangible world, but rather exist within the complex interactions of informational, physical, and cognitive elements in the information environment. As a result, social cybersecurity provides an interdisciplinary approach to account for these complexities.

One notable framework in this field is the BEND framework, developed by researchers at Carnegie Mellon University, which builds upon Ben Nimmo's "4D" Approach used to describe Russian information warfare against Ukraine in 2014 [67]. The BEND framework expands upon the four tactics of dismiss, distort, distract, and dismay, and accounts for additional ways in which actors may try to influence a target audience (see Figure 1.1 for the 16 BEND maneuvers). This framework allows for the mapping of community and network maneuvers, as well as the narratives that actors use to alter informational content towards a target audience, using bot probability scores, psycholinguistic cues, and semantic stance with a social network [12, 15, 21].

In the context of global competition between China and the United States, China has been employing influence operations against the United States with the objective of weakening the US and exacerbating social and political divisions. This is often referred to as China's "divide and disintegrate" strategy [37]. In 2021, China started propagating narratives through its state-affiliated news sources that the United States is in irreversible decline economically and socio-culturally, with the ultimate goal of ending US global hegemony or even causing the collapse of the United States [14]. This belief was articulated by CCP Central Committee member Yang Jiechi in the 2021 Anchorage Summit when he dismissively told US diplomats that the United States does not have the qualification to speak to China from a position of strength [49].

China's efforts to achieve information dominance over the United States are largely carried out through social media manipulation, which can be effectively analyzed using the BEND model. The CCP employs the doctrine of "Mind Superiority," which involves using psychological warfare to shape or control the enemy's cognitive thinking and decision-making, as part of its "cognitive domain operations" strategy [37]. This concept aligns with the US and NATO concept of "information dominance" and can be seen as the desired end state of a successful influence operation. By influencing what is discussed in the US and how it is discussed, China aims to undermine US government policies and credibility, ultimately seeking to "divide and disintegrate" the United States through its influence operations. Social cybersecurity plays a vital role in understanding and countering these efforts by analyzing the complex maneuvers and narratives employed by state actors in the digital realm [21].

		Community	Narrative
Maneuver	Deliberate action to achieve a desired end state	Altering the connections within the social network	Altering content within the information network
Impact	Maneuver end-state	Change in the social network structure	Change in the content of a narrative or message
Positive	Maneuvers focused on creating growth or increase (real or perceived) in a metric for a network or narrative	**Back** **Build** **Bridge** **Boost**	**Engage** **Explain** **Excite** **Enhance**
Negative	Maneuvers focused on reduction or decrease (real or perceived) in a metric for a network or narrative	**Neutralize** **Nuke** **Narrow** **Neglect**	**Dismiss** **Distort** **Dismay** **Distract**

Figure 1.1: Overview of the BEND manueuvers [15]

Chapter 2

Data and Tools

2.1 Data

This thesis incorporates data across different types of information campaigns. It extends methods beyond social media data to capture a more holistic characterization of how China wages state-sponsored campaigns toward Western audiences. This data incorporates rich social network connections through the Twitter API, kinetic flight data from China's air incursions towards Taiwan, scraped media datasets from China's Ministry of Foreign Affairs, Global Times and People's Daily state-affiliated newspapers, and exported search engine optimization data from Ahrefs.

Table 2.1: Summary of datasets used in this thesis.

Data	Type	Size	Ch. 2	Ch. 3	Ch. 4	Ch. 5
Elections in the Asia-Pacific	Twitter	16M texts	✓	✓		✓
State Actor Timelines	Twitter	180K tweets	✓	✓	✓	✓
Democracy Summit	Twitter	7.2K tweets	✓			✓
Covert Operations	Twitter	1M tweets	✓			✓
Taiwan	Twitter	2M tweets	✓	✓		✓
Taiwan Kinetic	Flight	3 years		✓		
Scraped Data	News	1.2M		✓		
Ahrefs Data	SEO	1M			✓	

2.1.1 Elections Data

Elections in the Asia-Pacific provide a regional context for how China participates in other regional discourse and provides insight into geopolitical relations. Datasets are collected from the Philippines, , Taiwan, South Korea, and Japan for over 16 million tweets from 2020 to 2022.

2.1.2 State Actor Timelines

We used the Twitter V2 API to collect the timelines of approximately 350 Chinese government or media-affiliated accounts for 180K tweets. We collected the retweets to understand better the amplifying network supporting MOFA's public diplomacy efforts.

2.1.3 Democracy Summit

The United States hosted a two-day "Summit on Democracy" event in mid-December 2021 to address democracy-related challenges worldwide, inviting over one hundred countries to participate. In response, Chinese state-sponsored Twitter accounts began a hashtag campaign around this event with anti-US sentiment and messaging for its internal event called "Dialogue on Democracy" on December 2, 2021. China's hashtag campaign around the Summit on Democracy represents a time-constrained influence campaign with extensive state-sponsored support. We collected tweets using the hashtags #WhoDefinesDemocracy and #WhatisDemocracy, resulting in the collection of 7,798 tweets from September 1 - December 31, 2021.

2.1.4 Covert Operations

The observable data that researchers have makes attribution extremely difficult. Social media platforms maintain access to technical indicators that provide more certainty around attributing state-sponsored information campaigns. We include suspected campaigns in addition to verified campaigns to provide ground truth in our analysis of emerging covert information tactics.

- Critics of China: We collected 64,774 tweets from the Twitter API using the keywords Guo Wengui, Li Meng Yan, and Jiayang Fan. China uses its spammy bot network to target exiled Chinese citizens who do not support China's geopolitical interests. These tweets are not attributed but are likely PRC covert operations based on news reporting [23].
- China's Protest Lockdown: Protests erupted across China following the Urumqi fire that killed a family during China's covid lockdown policy [32]. A spammy network used potential hashtag-jacking tactics to block out Twitter stories about the protests. We collected over 3M tweets from November 28, 2022, to December 19, 2022, using hashtags of the Chinese cities, using the following hashtags: #chinaprotest2022, #上海[Shanghai], #乌鲁木齐路[Urumqi Road], #兰州[Lanzhou], #北京[Beijing], #北京线下[Beijing Offline], #拉萨[Lhasa] and #西宁[Xining].
- Disclosed and Validated Twitter Network Takedowns: We have over 5,000 accounts from 2019-2021 that Twitter attributed to the PRC from Twitter's Moderation Research Consortium [24]. This dataset includes account information and the tweets that these accounts participated in.

2.1.5 Taiwan

We collected tweets using the hashtag #Taiwan, resulting in the collection of over three million tweets from January 1, 2021 - October 31, 2022. We chose these dates to capture China's Twitter

dialogue before Russia invaded Ukraine, in addition to the months following US House Speaker Nancy Pelosi's visit to Taiwan in August. After filtering our data to tweets and retweets relevant to Chinese-flagged accounts, our dataset consisted of 114,719 tweets.

We also use a time-series aggregation of incident reports published by Taiwan's Republic of China Ministry of National Defense on China's People's Liberation Army (PLA) air activities within the Taiwanese Air Defense Identification Zone (ADIZ). These reports provide the number and type of PLA aircraft in the ADIZ when they cross the median strait boundary, along with location and time information [19, 20].

2.1.6 Scraped Datasets

This dataset was exported from the FOCUSdata project at the National Security Studies Department at New Jersey City University (NJCU), a collaborative project with the Rutgers University Center for Critical Intelligence Studies under a grant from the U.S. Office of the Director of National Intelligence (ODNI).

- Ministry of Foreign Affairs: 2,607 English-language speeches and communique from China's Ministry of Foreign Affairs from November 15, 2000, to December 31, 2022 [36].
- Global Times: 677,532 English-language articles from the state-affiliated Global Times outlet. This dataset spans from April 9, 2009, to December 31, 2022 [34].
- People's Daily: 544,940 English-language articles from the state-affiliated newspaper outlet People's Daily. This dataset spans from May 12, 2007, to December 31, 2022 [35].

2.2 Tools

2.2.1 Ora-Pro Software

Ora is a dynamic meta-network and visualization tool with interoperability to import other network extraction technologies to visualize and measure information diffusion, belief change, stance detection, and understand how networks change over time [4, 22]. Ora maintains an import function for Twitter data and maps Twitter's rich metadata fields to a meta-network structure. Critical data fields include the language metadata tag, sender of a tweet, whether a tweet is a retweet and thus not the original tweet, temporal data and self-reported geotags. We use Ora to rapidly calculate network metrics such as centrality measurements, community detection and clustering and exploratory data analysis.

2.2.2 Netmapper

Netmapper is part of Ora's interoperable pipeline for network analysis and extracts networks from unstructured text data in over 40 languages. The context-level attributes Netmapper provides include psycholinguistic cues and emotion [77], semantic networks [75], and social identity theory [42]. The linguistic cues enable us to analyze sentiment and emotionality from pronouns, emojis, and negative or positive language. These cues are foundational for generating a BEND model in Ora.

2.2.3 Bothunter

Bothunter is a bot detection algorithm that generates a prediction that an account is a bot that demonstrates automated activity and is likely controlled by software. Bothunter is a tier-based random forest classifier trained on past bot campaigns and maintains an accuracy above 90% [10]. To increase certainty around our bot classification for each Twitter account, we use the recommended bot probability score of .7, at which the bot classification label is most stable from flipping from one class to the other for outlying bot activity [66].

2.2.4 Ahrefs

Ahrefs is a search engine optimization (SEO) platform for research on traffic and optimizing websites. Ahrefs maintains its web crawler that processes 8 billion pages daily to update its indexed backlinks [1]. We use this tool to collect SEO data for particular websites of interest, collecting backlink data and how users get to sites via keyword search phrases.

2.2.5 BERT-Topic

We use the open-source Python library BERTopic to convert the unstructured text of tweets into document clusters for evaluating semantic text similarity [39]. We use this topic modeling algorithm to provide a high-level overview of China's strategic messaging. This library converts text into sentence embeddings using a pre-trained language model. After converting our tweets into sentence embeddings, the library uses methods such as UMAP and HDBSCAN algorithms to reduce data dimensionality and cluster similar documents. Finally, it creates topic clusters using a variant of the Term Frequency-Inverse Document Frequent (TF-IDF) algorithm.

Chapter 3

Research Plan

3.1 Characterizing China's influence campaigns (Chapter 2)

3.1.1 Research Questions

This chapter provides the foundational empirical basis and theoretical understanding of how the PRC wages individual information campaigns, covering the range of both attributed and unattributed campaigns. Additionally, we will map these campaigns to BEND information maneuvers to understand how China wages information tactics within a campaign. Our key research questions for this chapter are as follows:

- Can we characterize PRC information campaigns on social media using network measurements?
- How does China use information maneuvers within its campaigns?

3.1.2 Methods and Proposed Work

This chapter uses network measures to examine how individual information campaigns are structured and how information diffuses through social media network structures. We propose analyzing multiple datasets under two different types of campaigns; overt campaigns in which we examine how the PRC wages public diplomacy and its amplifier network, in addition to covert campaigns, in which there is intentionally no attribution to the PRC.

We will conduct an initial statistical and network analysis on an overt campaign dataset using the pipeline in Figure 3.1 where we first label all agents within a dataset as either PRC state-sponsored or state-affiliated or a probable bot account. Our Twitter datasets will be converted into meta-networks or a network of networks in which relationships are defined between Twitter users, shared hashtags, and interactions with tweets, such as retweeting or liking a tweet and the URLs that are embedded within a given tweet.

- Size: Number of actors within a network
- Density: Metric for denoting how connected a network is, represented by the fraction of edges between agents over all possible connections ratio [79].

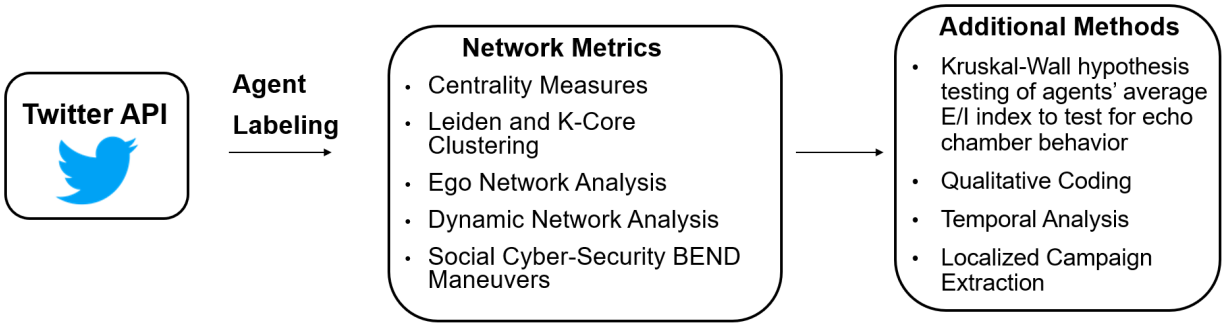

Figure 3.1: Pipeline for mapping network information maneuvers

- E/I Index: The External/Internal index is a community structure measure representing a node's isolation from an external community conversation with the rest of the network. This measure is calculated by the formula $E - I/E + I$ where E represents external links, and I represents internal links [53].
- In-Degree Centrality: Within a directed network, this represents the number of nodes adjacent *to* a given node [79].
- Out-Degree Centrality: With a directed network, this represents the number of nodes adjacent *from* a given node [79].

In published work [47], my initial findings analyzing a small overt campaign around the Democracy Summit indicate a few key factors to expand and investigate. First, *overt campaigns* may be defined as simple centralized information campaigns in which key personnel propagate tweets and amplifiers retweet. For Twitter data, a user with high In-Degree Centrality is generally characterized by a high retweet, reply, and quote frequency. In contrast, users with high Out-Degree Centrality will have tweets, replies, or quotes that other users frequently share.

Finding Key Superspreaders In a Twitter dataset, a user with high Out-Degree Centrality will have tweets, replies, or quotes that other users frequently share. Within our labeled state-sponsored accounts, ten accounts were responsible for 50% of all original tweets. This group can be captured by the following in which the out-degree centrality of node v in the network is given by:

$$C_{\text{out}}(v) = \frac{\text{outdeg}(v)}{|V| - 1}$$

where $\text{outdeg}(v)$ is the number of outgoing edges from node v and $|V|$ is the total number of nodes in the network. To identify nodes whose out-degree centrality is one standard deviation above the mean, denoted as $(\mu + \sigma)$, we can use the following equation:

$$N = \{v \mid C_{\text{out}}(v) > (\mu + \sigma)\}$$

where v is a node in V, and $C_{\text{out}}(v)$ is the out-degree centrality of node v.

Finding Key Amplifiers In a Twitter dataset, a user with high In-Degree Centrality is generally characterized by a high retweet, reply, and quote frequency. Similar to our outliers for

out-degree centrality, all six top amplifier accounts were probable bot accounts. We can find these accounts by the following: Let $G = (V, E)$ be a directed graph, where V is the set of nodes and E is the set of edges. The in-degree centrality of node v in the network is given by:

$$C_{\text{in}}(v) = \frac{\text{indeg}(v)}{|V| - 1}$$

where $\text{indeg}(v)$ is the number of incoming edges to node v and $|V|$ is the total number of nodes in the network. Our information campaign's main information propagators are those nodes whose in-degree centrality is one standard deviation above the mean, denoted as $(\mu + \sigma)$, we can use the following equation:

$$N = \{v \mid C_{\text{in}}(v) > (\mu + \sigma)\}$$

where v is a node in V, and $C_{\text{in}}(v)$ is the in-degree centrality of node v.

Finding Key Facilitators We define key facilitators as agents belonging to both in-degree and out-degree centrality outlier sets as defined above. We found one PRC diplomatic account and one media account that matched these parameters. Additionally, the diplomatic account became active before, during, and towards the end of the campaign, providing more evidence of its function as a facilitating account that creates and retweets content within an information campaign.

To identify nodes whose in-degree centrality and out-degree centrality are one standard deviation above the mean, denoted as $(\mu + \sigma)$, we can use the following equation:

$$N = \{v \mid C_{\text{in}}(v) > (\mu + \sigma) \text{ and } C_{\text{out}}(v) > (\mu + \sigma)\}$$

where v is a node in V, $C_{\text{in}}(v)$ is the in-degree centrality of node v, and $C_{\text{out}}(v)$ is the out-degree centrality of node v.

Localized Campaigns A given information campaign may represent international reach but may also contain smaller campaigns with a more targeted intended audience. We propose extracting sub-networks that may represent localized campaigns conducted through China's Consulate offices based on research regarding China's past information operations [48, 50]. By using our known labeled agents from the labeling process, we can find similar unlabeled agents by looking at outliers within our closeness centrality network that determines the closeness of a node to other nodes within a network. We examine the nodes at least one standard deviation above the mean and explore these accounts manually to determine if they represent other non-Chinese state-actor accounts. We also use the tweet language metadata tags to add tweets if the targeted audience uses a specific language. We extract the sub-meta network of all known state actors and follow-on tweets labeled for a given language.

We examined China's localized campaign strategies to target different populations, primarily through China's diplomatic offices. These campaigns frequently coordinated with local government, news outlets, and community members. While the impact of these campaigns was not measured, we found evidence of China using similar tactics in different countries to disseminate Chinese consulate messages with similar verbiage and rhetoric to local populations. This finding indicates a global strategy with localized sub-components for diffusing messaging at the international and local levels.

3.1.3 Challenges and Limitations

Our social media analysis is solely derived from our Twitter data captures. As such, our analysis reflects the PRC's social media strategies in this domain and may not reflect its campaigns taken on other platforms such as Weibo or localized social media platforms such as the PTT Bulletin in Taiwan. Attributing covert campaigns on social media platforms remains a technical challenge. We supplement our covert datasets with validated covert networks to account for our inability to attribute the datasets to the PRC fully.

Attribution challenges are a primary driving factor in this chapter, but distinguishing between attributed and unattributed campaigns is an open area of research, as state actors intentionally mask their origin or use proxies to disseminate information. Attribution requires analysis and triangulation of multiple data sources which are frequently not available through public API. As such, there remain strong limitations in accurately identifying the true origin of information campaigns.

Network analysis is a useful method for understanding social structures and dynamics, but it has its limitations. This method may not capture all relevant factors or interactions, and other qualitative or quantitative methods may be needed to complement the network analysis and provide a comprehensive understanding of China's influence campaigns. Additionally, these measurements are based on assumptions and simplifications of social network structures, and their validity and reliability may vary depending on the context and data being analyzed. We apply mixed methods for our research questions in this thesis to fill in these methods gaps with our network approach.

3.2 China's Narratives towards regional neighbors (Ch3)

3.2.1 Research Questions

This chapter builds off the last chapter, extending beyond *how* the PRC messages to further explore *what* the PRC messages. Specifically, we explore what themes emerge in the PRC's narratives using both social media and newspaper articles. Our data allows us to start understanding how these narratives have shifted over time in rhetoric oriented towards China's neighboring countries in the East Asia Pacific, e.g. South Korea, Japan, the Philippines, and Taiwan.

The Asia Pacific is currently undergoing a turbulent period of shifting security alliances and expanding spheres of influence toward China or the United States. Russia's invasion into Ukraine may seem far removed from the region, but it has had a rippling effect on security posturing. Japan is continuing to strengthen its relationship with NATO in the Indopacific [5] in addition to strengthening security cooperation and economic trade with South Korea [58]. There is a research gap in exploring regional narratives and how they shift over time. We ask the following questions in this chapter:

- Can we extract China's shifting narratives and dialogue about regional neighbors?
- What do these narratives tell us about the region's geopolitical trajectory using the BEND maneuvers?

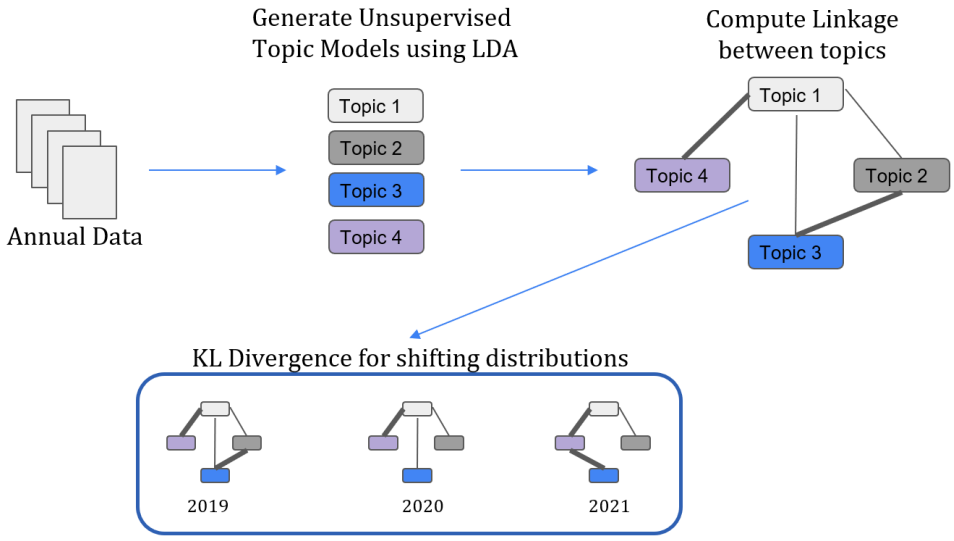

Figure 3.2: Creating and analyzing point-wise mutual information within a network structure

3.2.2 Methods and Proposed Work

Our base data for this chapter uses Twitter social media data and scraped news articles. We run a similar analysis on our two datasets to extract narratives and understand how the latent variables for our datasets shift over time. The pipeline in Figure 3.2 demonstrates how we can transform unstructured text, extract narratives and compare network structures over time. In this chapter, the planned work will: (a) expand our BEND analysis to understand how China uses information maneuvers toward specific countries, (b) expand our understanding of China's narratives and how they have shifted, (c) Provide a case study for the special case of how China uses narratives to support its offline kinetic activity.

Unsupervised Topic Modeling using LDA To capture the semantic context of our data, we build unsupervised topic models based on Latent Dirichlet allocation (LDA), which is a powerful tool for uncovering latent variables in observed data [16]. These topic models assume that a dataset comprises some number of topics and that each word in each document is mapped to a topic, effectively extracting word and phrase patterns and clustering them to best describe document groupings. The joint probability, p_{ij}, can be calculated using the following formula:

$$p_{ij} = \frac{\text{Number of documents where topic i and topic j co-occur}}{\text{Total number of documents or occurrences in the dataset}}$$

Precisely improving the quality of the topic quality output is tricky as not match human quality ratings [44]. A manual and iterative process for data cleaning, topic merging and parameter tuning will be how we increase topic quality [71].

Linkage Networks Once a topic model has been complete, we can generate the *linkage* between topics, where a link indicates the extent that two topics co-appear within a document, weighted by the joint probability or how often two topics would co-appear by chance [69]. This

linkage is called point-wise mutual information (PMI), and is a measure in information theory to quantify the association or dependency of two events [27, 59]. PMI can be defined as the following:

$$\text{PMI}(x, y) = \log_2 \left(\frac{P(x, y)}{P(x) \cdot P(y)} \right)$$

We can discover how highly interlinked nodes are clustered using the Louvain clustering algorithm to detect these groupings.

How do Narratives shift? We use a Bayesian definition of probability to measure surprising text in our data by measuring the distance between posterior and prior distributions with the Kullback-Leibler divergence (KL) [46]. KL divergence[54] is defined as the following:

$$\text{KL}(P \parallel Q) = \sum_x p(x) \cdot \log_2 \left(\frac{p(x)}{q(x)} \right)$$

where P and Q are two probability distributions being compared, $p(x)$ represents the data the model has been trained on, and $q(x)$ represents a new data distribution.

Lastly, we will use an innovative application of KL introduced by researchers [6] to measure how novel a document or narrative theme is with surprise. We can compare the surprise factor of a given document to prior documents in addition to documents that appear in the future of collected data, which is defined as transience. Using these measures will allow us to identify which documents signal a shift from past discussions and where narratives first emerge within our collected documents.

3.2.3 Challenges and Limitations

The primary challenge for this chapter will be the disparity between how frequently China discusses country *x* versus country *y*, in addition to producing coherent topic models through which we can compare distributions over time. Additionally, while KL Divergence has been shown to capture cognitive aspects of language shift, the signals in our data may not be strong enough to warrant finding major shifts in the PRC's narratives.

Lastly, while LDA is a commonly used technique for topic modeling, it has limitations, such as sensitivity to the choice of hyperparameters, reliance on bag-of-words representation which may lose contextual information, and potential for interpretability issues in interpreting the resulting topics. The interpretation of the topics generated by LDA may also be subjective and requires manual intervention for data cleaning, topic merging, and parameter tuning. This subjective element may introduce bias and affect the reliability of the findings.

3.3 Cross-Domain Activity: How does China use the internet to push narratives? (Ch 4)

3.3.1 Research Questions

Information operations thrive by manipulating a target audience to think, feel or act in a specific way. Web browser search results can also be manipulated by search engine optimization with both"white hat" ways in which websites adapt to be more discoverable by search engines and "black hat" ways that seek to manipulate and trick web browser search results [38]. There is a research gap regarding how SEO impacts information operations in reaching an intended target audience. However, emerging research reveals that state actors have used black hat techniques to drive web traffic to specific websites. We seek to expand this thread of research by answering the following:

- How do authoritarian governments such as China use SEO to promote traffic?
- What are the cross-domain patterns that emerge?

3.3.2 Methods and Proposed Work

In this chapter, we apply network methods to characterize how China uses black and white hat search engine optimization to steer internet traffic to sites. This approach creates a higher-level understanding of how the PRC narratives flow across platforms to users and what underlying technique they may use to drive traffic to any site.

Seed URLs We will use the proprietary SEO tool Ahrefs to conduct website analysis, discovery, and information extraction. We will initialize our website traffic ecosystem by first exporting SEO data for important sites of interest. Given the PRC's multifaceted approach to using both military, civilian, and soft power organizations to conduct influence operations, we will be using the following categories to collect website data:

- News: Top State-Affiliated News Sites e.g., Global Times, Xinhua, CGTN
- Government: Top Government Domains
- Social Media: Top Superspreader accounts for Twitter, Youtube, and Facebook
- Think Tanks: Most influential PRC-affiliated think tanks [56].

Generate Network Relationships and find Co-Amplification Using methods first introduced by Williams and Carley [80], we will generate networks for each group of websites exported from Ahrefs. We will generate *backlink* networks, also known as inbound links, to connect one website to another. We will also generate *key-phrase* networks. These words are added to online content to match a target audience's query results for more accurate search results.

Our backlink network will enable us to see if websites heavily use backlinks from low-quality domains, such as blog sites, that are easily propped up to generate traffic demand. This black hat SEO technique is called a *link scheme*, and it attempts to manipulate browser results with un-

natural links that may trick the browser into bumping up a website's ranking [40]. Additionally, common websites that provide many outbound links to multiple websites of interest will begin to show a strategy across multiple domains. We will pull the 1,000 top back-linking websites for each website by backlink volume. Source nodes will be the website of interest, target nodes are the domains and edge links will represent the number of backlinks between each node pair.

Our key phrase networks will also show each website's top key phrases to guide web traffic. We will use a Leiden clustering algorithm to group websites by how similar their connections to key phrases are. Apart from finding patterns for different types of websites, we are interested in finding unusual key phrases. Black hat exploitation of *data voids* has demonstrated that websites linked to conspiracy theories will exploit unusual key search terms to create a more direct line of traffic to a website. For each website, we will pull the top 1,000 critical phrases with the highest Google ranking for each domain, where source nodes are the websites, target nodes are the key phrases and edge links are a connection.

We measure co-amplification as measuring the overlap between any website of interest and the back-linking domains. We will explore these relationships using a LinkScore metric given by:

$$LS(i,j) = \frac{A(i,j) - E(i,j)}{SD(i,j)}$$

where $A(i,j)$ represents the actual number of links, $E(i,j)$ represents the expected number of links, and $SD(i,j)$ represents the standard deviation of the expected number of links between nodes i and j in the network.

3.3.3 Challenges and Limitation

We do not have access to the temporal properties of SEO network data that would enable coordination analysis. We can observe if co-amplification occurs, but we can not conclude that a coordinated effort is being conducted. Additionally, co-amplification is not an indication of causality. If key phrases are similar enough, there are valid reasons for backlink correlations between websites of interest.

Identifying other black hat SEO techniques will be challenging, in addition to making the distinction between manipulative and covert tactics. We will likely encounter predictable but challenging problems in defining this boundary for what constitutes likely covert black hat SEO.

Last and most important, the landscape of SEO and information operations is constantly evolving, with new techniques and strategies emerging over time. The research may face challenges in keeping up with the rapidly changing nature of SEO practices and information operations, and the findings may have limitations in capturing the most current and up-to-date strategies used by China or other actors.

3.4 Synthetically Generating PRC Information Operations (Ch 5)

The Information Environment may be defined as the " aggregate of individuals, organizations, and systems that collect, process, disseminate, or act on information—as consisting of physical, informational, and cognitive dimensions – Joint Concept for Operating in the Information Environment (JCOIE) [68].

However, the information environment (IE) is more complex than just the important actors, organizations, and systems that create information than we are interested in. There are countless other actors, organizations and systems creating volumes of noise that may not be relevant to the information environment in question. A 2022 report on general internet activity found that roughly 42% of all internet activity was created by bots or automated accounts [2].

Researchers at Carnegie Mellon University created a game for instructing analysts and decision makers on how to extract critical information from the IE. Project OMEN (Operational Mastery of the Information Environment) is a training game designed and developed to help analysts fight misinformation on social media. It is a practical and scalable game for learning about social media analytics for either analysts or decision makers [52]. The first OMEN scenario began with semi-synthetic training data based on the NATO Trident Juncture exercise in 2018 and the COVID-19 pandemic. Analysts were trained to use ORA and other network analysis tools and provided scenario data daily during the 5-day exercise to extract critical information on information maneuvers taking place. OMEN is an example of a decision-making game where players are presented with choices at different parts of the story-line, where they have the ability to affect the game's outcome based on their decisions [7].

This chapter progresses the game to incorporate scenarios based on state-actor tactics, techniques and procedures (TTPs) within the information environment. While our focus will be the PRC, these same methods will likely be expanded to also understand Russian and other state-actor TTPs within the information environment

3.4.1 Research Questions

The final chapter of this thesis integrates work from earlier chapters to synthetically generate the framework for PRC information maneuvers. This chapter contributes to ongoing efforts at automating information environment training scenarios with the scenario generation of state-sponsored information operations. Given what we now know about the types of information campaigns, types of key actor, and PRC patterns for information maneuvers, can we synthetically design a PRC information campaign?

3.4.2 Methods and Proposed Work

We propose extending current research on a scenario generator to facilitate state-sponsored information campaigns by creating templates for the scenario, types of actors, and types of information maneuvers based on preceding chapters. The scenario generator is currently an R-Shiny app and will serve as the front-end interface for a customer to input the parameters for a de-

sired wargame. If a user designates that the scenario will entail a PRC information campaign, our templates will automatically draw from known superspreader Twitter accounts and denote a distribution for a bot amplifier network. Additionally, we will evaluate the final synthetic information campaign output by other Office of Naval Research grant performers to validate their methods for producing synthetic content.

Type of Scenario Overall, OMEN's goal is to enable synthetic information campaign generation for the following scenario categories: Health, Election, Climate, Conflict/War, Collision/Accident, Military and Diplomatic events, where we generate the information environment for a given event. Key fields for a given template include an event taking place over a timeline, key actors to include adversarial actors and their potential allies, friendly actors and their allies, the location that an event takes place, and a scenario description for what major sub-events occur each day prior to, during, and after an event.

For this work, we will conduct research on possible scenarios that cover a spectrum of conflict, types of actors, and types of sub-events. Our final scenario may include one of the following:

1. Planned military exercise: There are a number of planned joint exercises that currently take place in the Indopacific region to include: US-Philippines Balikatan exercise [65] and other bilateral exercises with Japan [55], and South Korea [51].
2. Military Collision: As China continues increasing the extent of military encirclement around Taiwan during military drills, there is also an increased chance of a military accident or collision that could spike conflict. China simulated precision strikes on Taiwan in April 2023, with an increased amount of both Chinese and Taiwanese aircraft and maritime assets around the Taiwan Strait median line. This set of military drills was in response to the Taiwanese Presidents visit to the United States [64].
3. National Security Law: This scenario would extract information maneuvers from China's implementation of the Hong Kong's Security Law, which laid down the legal framework for expanding China's jurisdiction and prosecution authority over Hong Kong dissidents [43].
4. Military Invasion: This scenario would encompass the most complex and unpredictable information maneuvers around an actual invasion into Taiwan.

Our scenario for this chapter has no bearing on the actual likelihood of such an event occurring, but will be based on current geopolitical tensions and conflict.

Type of Actors This category covers main actors, background actors, organizations, and bot accounts. We may extend this category to cover other types of inauthentic activity that we see in either covert or overt types of campaigns to include use of influencers, bots that seek to drown out a given hashtag (also known as hashtag jacking), puppet accounts operated by humans, etc.

- PRC State-Sponsored Government Accounts: Based on collected accounts
- PRC State-Affiliated Media: Based on collected accounts
- Other State-Actors: United States, Philippines, South Korea, Japan, Russia, India

- Media Organizations: Regional, International, and State-Sponsored
- International: United Nations, NATO, ASEAN

Course of Action Development The Courses of Action (COA) template requires at least two COAs created for each day of the simulated exercise. Different examples of COAs might be to use the BEND *explain* maneuver to counter disinformation narratives or to use narratives that *excite* a targeted audience prior to a joint exercise if the public relations outreach is limited. These COAs will reflect available options a user can make to shape the information environment according to their messaging priorities.

Synthetic Data Evaluation We will conduct evaluation of the synthetic scenario data by comparing the output to what we have observed in past campaigns.

- Network Structure: Are Key Actors and Amplifier networks similar to other campaigns? How is information diffused?
- BEND Distribution: Does the data for each day reflect major information maneuvers within the scenario? Do the COAs line up to what is observed in the data?
- Inauthentic activity: Do bot networks and puppet accounts match PRC TTPs?
- Semantic Similarity: How are the state-actor tweets generated? How similar is the discourse to historical data? The open-source library HuggingTweets is a tweet generator that enables a user to fine-tune the pre-trained transformer on past tweets to generate or predict new text with a simple prompt [28].
- ORA Report Analysis: Does the scenario and prompt match what the user input? For each day in the data, do Key Actors get reflected in the data? Is inauthentic coordination present and discoverable within the data?
- COA Analysis: This portion of user validation may require an internal hackathon to analyze each COA scenario and ensure decision points for players mirror what occurs in the data.

3.4.3 Challenges and Limitations

The primary logistical challenge for this chapter is that beyond the user interface, the synthetic data is being generated by an external group. Our timeline may have challenges for any iterations that are required to fix the scenario data and ensure that our synthetic campaign mirrors state-sponsored TTPs. Additionally, our understanding of the generative process for the output data is determined by our collaborating partners.

The generation of synthetic information campaigns, even for research purposes, raises ethical concerns. Creating scenarios that involve state-sponsored information operations, such as those of the People's Republic of China (PRC), may potentially have real-world implications and impact the perception of actual events. Our work will ensure that the generated scenarios do not contribute to misinformation or manipulation of public perception.

Additionally, the accuracy and validity of synthetic scenarios generated through the proposed framework may be a challenge. The templates for the scenarios, types of actors, and types of information maneuvers are based on preceding chapters and existing knowledge. However, there

may be limitations in the accuracy and completeness of the templates, and the synthetic scenarios may not fully capture the complexity and nuances of real-world information campaigns. The templates for the scenarios, types of actors, and types of information maneuvers used in the framework may be based on existing data and knowledge, which may potentially contain biases based on the data collection and the accounts of interest. These biases may affect the accuracy and validity of the synthetic scenarios, and may impact the conclusions and insights derived from the generated scenarios.

Lastly, the proposed framework aims to generate synthetic information campaigns specifically for the PRC, but the generalizability of the findings to other state actors, such as Russia, may be a limitation. The tactics, techniques, and procedures (TTPs) used by different state actors may vary, and the framework may need to be adapted or modified to capture the unique characteristics of other state-sponsored information operations.

Chapter 4

Contributions

4.1 Theoretical Contributions

This thesis's theoretical framework is grounded in the information environment, which encompasses the complex ecosystem of organizations, individuals, and systems that collect, process, disseminate, or act on information. We recognize the three interconnected dimensions of the information environment, namely the physical, informational, and cognitive components. This theoretical perspective provides a wide lens to understand the complexities of the modern information landscape, particularly on the internet, where information can be propagandized, manipulated, and disseminated rapidly globally.

Furthermore, this thesis draws on past research on state-sponsored campaigns, particularly in the context of China's activities, to highlight the limitations of existing approaches that rely on collaboration with social media platforms for data access and attribution. This theoretical foundation underscores the need for alternative methodologies to overcome proprietary data restrictions, scarcity of resources, and human interest in obtaining data on information campaigns. We propose a mixed methods approach that integrates network centrality methods, overarching narrative analysis with information theory measurements, and a network application of how China uses underlying SEO to address these limitations and comprehensively understand state-sponsored inauthentic behavior in the information environment.

Moreover, the theoretical framework of the thesis acknowledges the dynamic nature of the information environment, particularly concerning China's information campaigns toward regional areas. We analyze the shifting dynamics of China's dialogue in the information domain over time and emphasize the importance of considering temporal changes in communication patterns and strategies. This theoretical perspective contributes to a nuanced understanding of China's information campaigns and provides a foundation for analyzing the motivations and tactics employed by state actors in the information environment.

Overall, the theoretical framework of the thesis is grounded in the concept of the information environment and draws on past research to highlight the limitations of existing approaches. The proposed mixed methods approach and emphasis on temporal dynamics contribute to a comprehensive understanding of state-sponsored inauthentic behavior in the information environment, providing theoretical insights that inform the methodology and analysis undertaken in this thesis.

4.2 Methodological Contributions

This thesis makes several methodological contributions to literature around state-sponsored information campaigns. We introduce a basic framework using network centrality methods for identifying key agents within an information campaign, including outliers and other metrics that reveal the functions and roles of different agents. Pointwise mutual information networks, on the other hand, provide insights into the linguistic patterns and semantic associations in the communication of state actors, particularly in their dialogue towards regional countries, Russia, and the United States. This mixed methods approach allows for a comprehensive characterization of state-sponsored campaigns in the information environment, capturing both structural and content-related aspects of information campaigns.

This thesis proposes an innovative approach to analyze China's potential use of search engine optimization (SEO) techniques to boost its websites' visibility in browsers artificially. We explore how China may manipulate search engine rankings to amplify its online presence by analyzing different categories of government-affiliated websites and social media domains. This novel methodological approach provides insights into the potential use of SEO as a tactic in state-sponsored information campaigns, contributing to the understanding of how governments can manipulate the information environment through technical means.

Lastly, this thesis integrates empirical findings into a decision-making game for analysts to identify inauthentic state-sponsored information maneuvers in the information environment. This practical contribution provides a valuable resource for researchers and analysts in the field, offering guidance and insights into the methodology employed in the thesis. The scenario generator and its references will serve as a comprehensive reference for researchers interested in studying state-sponsored information campaigns, providing critical findings for discovering information operations in the wild.

In summary, the thesis makes methodological contributions through its mixed methods approach that integrates network centrality methods and pointwise mutual information networks, its innovative analysis of search engine optimization techniques, and the extension of a synthetic training environment for analysts. These methodological contributions enhance the understanding of state-sponsored information campaigns in the information environment and provide valuable tools for researchers in this field.

4.3 Academic Contributions

Work for this thesis has resulted in a series of publications at journals and conferences. My initial framework for Chapter 1 on characterizing an overt PRC information campaign was published in the *SBP-BRIMS* conference proceedings [47] with an extension on localized campaign discovery pending in the *Computational and Mathematical Organization Theory*. Our work while conducting literature review in understanding PRC doctrine was published in *Small Wars Journal* [48], with follow-on empirical work pending a submission to *Security Dialogue*. We will continue to prioritize high-impact venues that value interdisciplinary social science work and insight.

Chapter 5

Proposed Timeline

The work schedule for this dissertation is as follows in figure 5.1. In the spring of 2023, I will extend my analysis for Chapter 2 to analyze covert datasets and additional overt datasets for characterizing PRC information campaigns. The topic modeling and linkage networks component to Chapter 3 will also be completed for the SBP-BRiMS conference. By the end of the summer of 2023, I will have Chapter 2 finalized. Additionally, I will expand the BEND analysis in Chapter 3 for the social media datasets. During this period, I will also begin working with OMEN developers to develop a computational template for automated campaign generation for training scenarios. In the fall of 2023, I will begin finalizing Chapter 3's narrative comparison between the social media and scraped datasets. Additionally, I will begin preparing the SEO networks for Chapter 4 analysis. By the winter of 2023, Chapter 4 will be nearing completion, in addition to any last changes for the OMEN scenario generator. The beginning of the spring semester mainly serves as a buffer period for any last analysis or components that need to be completed. My goal is to defend this thesis by April 2024.

	Spring 23	Summer 23	Fall 23	Winter 23	Spring 24
Chapter 2: Characterizing PRC Campaigns		Complete			
Extend Overt/Begin Covert Campaigns					
Social Cyber Security Maneuver Analysis					
Chapter 3: PRC Narratives			Complete		
Topic Modeling/Linkage Networks over time					
Social Media and Social Cyber Maneuver					
Comparison of Social Media and News Narratives					
Chapter 4: The PRC's Cross Domain Activity				Complete	
Seed Site Backlink Network Generation					
Analysis of potential link schemes					
Key Phrase Networks					
Extend other metrics for correlationg activity					
Chapter 5: Synthetic Campaign					Complete
Scenario Research					
Aggregate Key Information Manuever Findings					
Integrate PRC Tactics into Scenario Generator					
Finalize Thesis Document					March 2024
Thesis Defense					April 2024

In-Progress

Completing

Figure 5.1: Proposed Timeline for Dissertation Milestones

Bibliography

[1] What is Ahrefs? | Help Center - Ahrefs. URL https://help.ahrefs.com/en/articles/78203-what-is-ahrefs. 2.2.4

[2] 2022 bad bot report | evasive bots drive online fraud | imperva. URL https://www.imperva.com/resources/resource-library/reports/bad-bot-report/. 3.4

[3] Luca Maria Aiello, Martina Deplano, Rossano Schifanella, and Giancarlo Ruffo. People are strange when you're a stranger: Impact and influence of bots on social networks. In *Proceedings of the international AAAI conference on web and social media*, volume 6, pages 10–17, 2012. 1.2.1

[4] Neal Altman, Kathleen M Carley, and Jeffrey Reminga. Ora user's guide 2020. *Carnegie-Mellon Univ. Pittsburgh PA Inst of Software Research International, Tech. Rep*, 2:2, 2020. 2.2.1

[5] No Author. Japan welcomes increased NATO involvement in indo-pacific region. URL https://www.japantimes.co.jp/news/2023/04/06/national/hayashi-nato-involvement/. 3.2.1

[6] Alexander TJ Barron, Jenny Huang, Rebecca L Spang, and Simon DeDeo. Individuals, institutions, and innovation in the debates of the french revolution. *Proceedings of the National Academy of Sciences*, 115(18):4607–4612, 2018. 3.2.2

[7] Benoit Bediou, Don M. Adams, Richard E. Mayer, Elizabeth Tipton, and C. Shawn Green. Effects of strategy-based video games on cognitive performance: A meta-analysis. *Journal of Cognitive Enhancement*, 2(4). 3.4

[8] Matthew C Benigni, Kenneth Joseph, and Kathleen M Carley. Bot-ivistm: Assessing information manipulation in social media using network analytics. In *Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining*, pages 19–42. Springer, 2019. 1.2.1

[9] George Berry, Christopher J. Cameron, Patrick Park, and Michael Macy. The opacity problem in social contagion. *Social Networks*, 56:93–101, 2019. ISSN 0378-8733. doi: https://doi.org/10.1016/j.socnet.2018.09.001. URL https://www.sciencedirect.com/science/article/pii/S0378873317303465. 1.2.1

[10] David M Beskow and Kathleen M Carley. Bot-hunter: a tiered approach to detecting & characterizing automated activity on twitter. In *Conference paper. SBP-BRiMS: International conference on social computing, behavioral-cultural modeling and prediction and*

behavior representation in modeling and simulation, volume 3, 2018. 2.2.3

[11] David M Beskow and Kathleen M Carley. Social cybersecurity: An emerging national security requirement. *Military Review*, 99(2):117–127, 2019. ISSN 0026-4148. 1.2.1

[12] David M Beskow and Kathleen M Carley. Characterization and comparison of russian and chinese disinformation campaigns. *Disinformation, misinformation, and fake news in social media: emerging research challenges and opportunities*, pages 63–81, 2020. 1.2.3, 1.2.4

[13] Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 us presidential election online discussion. *First monday*, 21(11-7), 2016. 1.2.1

[14] Jude Blanchette and Seth G. Jones. Beijing's new narrative of u.s. decline. URL `https://opensource.csis.org`. 1.2.4

[15] Janice T. Blane. Social-cyber maneuvers for analyzing online influence operations. 1.2.4, 1.1

[16] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. In *Journal of Machine Learning Research*, volume 3, pages 993–1022, 2003. 3.2.2

[17] Samantha Bradshaw and Philip N Howard. Challenging truth and trust: A global inventory of organized social media manipulation. *The computational propaganda project*, 1:1–26, 2018. 1.2.2

[18] Samantha Bradshaw, Hannah Bailey, and Philip N Howard. *Industrialized disinformation: 2020 global inventory of organized social media manipulation*. Computational Propaganda Project at the Oxford Internet Institute, 2021. 1.2.1

[19] G. Brown and B. Lewis. Taiwan ADIZ violations. URL `https://docs.google.com/spreadsheets/d/1qbfYF0VgDBJoFZN5elpZwNTiKZ4nvCUcs5a7oYwm52g/edit#gid=2015900050`. 2.1.5

[20] Chris Buckley and Amy Qin. In a surge of military flights, china tests and warns taiwan. ISSN 0362-4331. URL `https://www.nytimes.com/2021/10/03/world/asia/china-taiwan-flights-airspace.html`. 2.1.5

[21] Kathleen M Carley. Social cybersecurity: an emerging science. *Computational and Mathematical Organization Theory*, pages 1–17, 2020. 1.2.4

[22] L Richard Carley, Jeff Reminga, and Kathleen M Carley. Ora & netmapper. In *International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation. Springer*, volume 3, page 7, 2018. 2.2.1

[23] Flora Carmichael. How a fake network pushes pro-china propaganda. URL `https://www.bbc.com/news/world-asia-china-58062630`. 2.1.4

[24] Twitter Transparency Center. Twitter moderation research consortium. URL `https://transparency.twitter.com/en/reports/moderation-research.html`. 2.1.4

[25] Sonali Chandel, Zang Jingji, Yu Yunnan, Sun Jingyao, and Zhang Zhipeng. The golden shield project of china: A decade later—an in-depth study of the great firewall. In *2019*

International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), pages 111–119, 2019. doi: 10.1109/CyberC.2019.00027. 1.2.2

[26] John J Chin. The longest march: Why china's democratization is not imminent. *Journal of Chinese Political Science*, 23:63–82, 2018. 1.2.2

[27] Kenneth Church and Patrick Hanks. Word association norms, mutual information, and lexicography. *Computational linguistics*, 16(1):22–29, 1990. 3.2.2

[28] Boris Dayama. Hugging tweets: Train a model to generate tweets. URL https://wandb.ai/wandb/huggingtweets/reports/HuggingTweets-Train-a-Model-to-Generate-Tweets--VmlldzoxMTY5MjI. 3.4.2

[29] Larry Diamond and Orville Schell. *China's influence and American interests: Promoting constructive vigilance*. Hoover Press, 2019. 1.2.2

[30] Renee DiResta, Carly Miller, Vanessa Molter, John Pomfret, and Glenn Tiffert. *Telling China's Story: The Chinese Communist Party's Campaign to Shape Global Narratives*. Stanford Internet Observatory, 2020. 1.2.2

[31] Richard M Everett, Jason RC Nurse, and Arnau Erola. The anatomy of online deception: What makes automated text convincing? In *Proceedings of the 31st Annual ACM symposium on applied computing*, pages 1115–1120, 2016. 1.2.1

[32] Emily Feng. How a deadly fire in xinjiang prompted protests unseen in china in three decades. URL https://www.npr.org/2022/11/26/1139273138/china-protests-covid-lockdown-urumqi-beijing. 2.1.4

[33] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *Communications of the ACM*, 59(7):96–104, 2016. 1.2.1

[34] Scott Fisher and Graig Klein. China's global times articles from 9 apr 2009 to 31 dec 2022, . URL https://dataverse.harvard.edu/dataverse/focusdataproject. 2.1.6

[35] Scott Fisher and Graig Klein. China's people's daily articles from 12 may 2007 to 31 dec 2022, . URL https://dataverse.harvard.edu/dataverse/focusdataproject. 2.1.6

[36] Scott Fisher and Graig Klein. Chinese ministry of foreign affairs articles from nov. 2000 - dec. 2022, . URL https://dataverse.harvard.edu/dataverse/focusdataproject. 2.1.6

[37] Kerry K Gershaneck. *Political Warfare:*. Marine Corps University Press, 2020. 1.2.4

[38] Dipayan Ghosh and Ben Scott. Digital deceit: the technologies behind precision propaganda on the internet. 2018. 3.3.1

[39] Maarten Grootendorst. The algorithm - BERTopic. URL https://maartengr.github.io/BERTopic/tutorial/algorithm/algorithm.html. 2.2.5

[40] Venkat N Gudivada, Dhana Rao, and Jordan Paris. Understanding search-engine optimization. *Computer*, 48(10):43–52, 2015. 3.3.2

[41] Falk Hartig. How china understands public diplomacy: The importance of national image

for national interests. *International Studies Review*, 18(4):655–680, 2016. 1.2.2

[42] David R Heise. Affect control theory: Concepts and model. In *Analyzing Social Interaction*, pages 1–34. Routledge, 2016. 2.2.2

[43] Javier C. Hernández. Harsh penalties, vaguely defined crimes: Hong kong's security law explained. ISSN 0362-4331. URL https://www.nytimes.com/2020/06/30/world/asia/hong-kong-security-law-explain.html. 3

[44] Alexander Hoyle, Pranav Goel, Denis Peskov, Andrew Hian-Cheong, Jordan Boyd-Graber, and Philip Resnik. Is automated topic model evaluation broken?: The incoherence of coherence. URL http://arxiv.org/abs/2107.02173. 3.2.2

[45] Zhao Alexandre Huang and Rui Wang. Building a network to "tell china stories well": Chinese diplomatic communication strategies on twitter. *International Journal of Communication*, 13:2984–3007, 2019. 1.2.1, 1.2.3

[46] Laurent Itti and Pierre Baldi. Bayesian surprise attracts human attention. *Vision research*, 49(10):1295–1306, 2009. 3.2.2

[47] Charity S. Jacobs and Kathleen M. Carley. #WhoDefinesDemocracy: Analysis on a 2021 chinese messaging campaign. In *Social, Cultural, and Behavioral Modeling: 15th International Conference, SBP-BRiMS 2022, Pittsburgh, PA, USA, September 20–23, 2022, Proceedings*, pages 90–100. Springer-Verlag. ISBN 9783031171130. doi: 10.1007/978-3-031-17114-7_9. URL https://doi.org/10.1007/978-3-031-17114-7_9. 3.1.2, 4.3

[48] Charity S Jacobs and Kathleen M Carley. Taiwan: China's gray zone doctrine in action. *Small Wars Journal*, 2022. 3.1.2, 4.3

[49] Martin Jacques. US can't accept painful fact that china is now its equal. URL https://www.globaltimes.cn/page/202103/1219181.shtml. 1.2.4

[50] Elsa B Kania. Chinese military innovation in the ai revolution. *The RUSI Journal*, 164 (5-6):26–34, 2019. 3.1.2

[51] Hyung-Jin Kim. US, south korea announce largest field exercises in 5 years. URL https://apnews.com/article/south-north-korea-us-drills-nuclear-66f94a64982e255b23ae6ac7860a5e2c. 1

[52] Catherine King, Christine Sowa Leipird, and Kathleen M. Carley. Project OMEN: Designing a training game to fight misinformation on social media. URL http://reports-archive.adm.cs.cmu.edu/anon/isr2021/abstracts/21-110.html. 3.4

[53] David Krackhardt and Robert N Stern. Informal networks and organizational crises: An experimental simulation. *Social psychology quarterly*, pages 123–140, 1988. 3.1.2

[54] Solomon Kullback. *Information theory and statistics*. Courier Corporation, 1997. 3.2.2

[55] Jonathan Lehrfeld. U.s. and japan team up for indo-pacific training exercise. URL https://www.airforcetimes.com/news/your-air-force/2023/01/23/us-and-japan-team-up-for-indo-pacific-training-exercise/. 1

[56] Cheng Li. China's new think tanks: Where officials, entrepreneurs, and scholars interact.

China Leadership Monitor, 29:1–21, 2009. 3.3.2

[57] Luca Luceri, Ashok Deb, Silvia Giordano, and Emilio Ferrara. Evolution of bot and human behavior during elections. *First Monday*, 2019. 1.2.1

[58] Jean Mackenzie. South korea and japan: A milestone meeting of frenemies. URL https://www.bbc.com/news/world-asia-64962733. 3.2.1

[59] Christopher Manning and Hinrich Schutze. *Foundations of statistical natural language processing*. MIT press, 1999. 3.2.2

[60] Stephen McCombie, Allon J Uhlmann, and Sarah Morrison. The us 2016 presidential election & russia's troll farms. *Intelligence and National Security*, 35(1):95–114, 2020. 1.2.1

[61] Carly Miller, Vanessa Molter, Isabella Garcia-Camargo, and Renée DiResta. Sockpuppets spin covid yarns: An analysis of prc-attributed june 2020 twitter takedown, 2020. 1.2.3

[62] Paul Mozur, Muyi Xiao, Gray Beltran, and Jeff Kao. China unleashed its propaganda machine on peng shuai's #MeToo accusation. her story still got out. URL https://www.propublica.org/article/china-unleashed-its-propaganda-machine-on-peng-shuais-metoo-accusation-her-story-still-got-out. 1.2.3

[63] Steven Lee Myers, Paul Mozur, and Jeff Kao. How bots and fake accounts push china's vision of winter olympic wonderland. URL https://www.propublica.org/article/how-bots-and-fake-accounts-push-chinas-vision-of-winter-olympic-wonderland. 1.2.3

[64] Sky News. China and taiwan ships in stand-off near sensitive buffer zone. URL https://news.sky.com/story/china-and-taiwan-ships-stand-off-near-sensitive-buffer-zone-12853403. 2

[65] Kelly Ng and Joel Guinto. US and philippines begin largest-ever drills after china exercises. URL https://www.bbc.com/news/world-asia-65236459. 1

[66] Lynnette Hui Xian Ng, Dawn C Robertson, and Kathleen M Carley. Stabilizing a supervised bot detection algorithm: How much data is needed for consistent predictions? *Online Social Networks and Media*, 28:100198, 2022. 2.2.3

[67] Ben Nimmo. Anatomy of an info-war: how russia's propaganda machine works, and how to counter it. *Central European Policy Institute*, 15:1–16, 2015. 1.2.4

[68] Joint Chiefs of Staff. Joint concept for operating in the information environment (JCOIE). URL https://www.jcs.mil/Portals/36/Documents/Doctrine/concepts/joint_concepts_jcoie.pdf. 3.4

[69] Chloe Perry and Simon DeDeo. The cognitive science of extremist ideologies online. *arXiv preprint arXiv:2110.00626*, 2021. 3.2.2

[70] Samantha C Phillips, Joshua Uyheng, and Kathleen M Carley. Competing state and grassroots opposition influence in the 2021 hong kong election. In *Social, Cultural, and Behavioral Modeling: 15th International Conference, SBP-BRiMS 2022, Pittsburgh, PA, USA, September 20–23, 2022, Proceedings*, pages 111–120. Springer, 2022. 1.2.3

[71] ME Roberts, BM Stewart, and D Tingley. Navigating the local modes of big data: the case of topic models. comput. *Soc. Sci*, 4, 2016. 3.2.2

[72] Fergus Ryan, Ariel Bogle, Albert Zhang, and Jacob Wallis. # stopxinjiang rumours: the ccp's decentralised disinformation campaign. 2021. 1.2.3

[73] Marcel Schliebs, Hannah Bailey, Jonathan Bright, and Philip N Howard. China's public diplomacy operations: understanding engagement and inauthentic amplifications of prc diplomats on facebook and twitter. 2021. 1.2.3

[74] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19(1): 22–36, 2017. 1.2.1

[75] John F Sowa. Semantic networks. *Encyclopedia of artificial intelligence*, 2:1493–1511, 1992. 2.2.2

[76] Jason Stanley. *How propaganda works*, pages 52–53. Princeton University Press, 2015. 1.2.2

[77] Yla R Tausczik and James W Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1): 24–54, 2010. 2.2.2

[78] Tom Uren, Elise Thomas, and Jacob Wallis. *Tweeting through the great firewall: preliminary analysis of PRC-linked information operations on the Hong Kong protests*. Australian Strategic Policy Institute, 2019. 1.2.3

[79] Stanley Wasserman and Katherine Faust. *Social Network Analysis: Methods and Applications*, pages 126–127. Cambridge University Press. 3.1.2

[80] Evan M. Williams and Kathleen M. Carley. Search engine manipulation to spread pro-kremlin propaganda. doi: 10.37016/mr-2020-112. URL `https://misinforeview.hks.harvard.edu/article/search-engine-manipulation-to-spread-pro-kremlin-propaganda/`. 3.3.2

[81] Samuel C Woolley and Philip Howard. Computational propaganda worldwide: Executive summary. 2017. 1.2.1