2022-03

UNDERSTANDING THE ROLE OF OVERT AND COVERT ONLINE COMMUNICATION IN INFORMATION OPERATIONS

Dimitrov, Boyan I.

Monterey, CA; Naval Postgraduate School

http://hdl.handle.net/10945/69632

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

UNDERSTANDING THE ROLE OF OVERT AND COVERT ONLINE COMMUNICATION IN INFORMATION OPERATIONS

by

Boyan I. Dimitrov

March 2022

Thesis Advisor:	Timothy C. Warren
Second Reader:	Shannon C. Houck

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE		*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC, 20503.		
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** March 2022	**3. REPORT TYPE AND DATES COVERED** Master's thesis
4. TITLE AND SUBTITLE UNDERSTANDING THE ROLE OF OVERT AND COVERT ONLINE COMMUNICATION IN INFORMATION OPERATIONS		**5. FUNDING NUMBERS**
6. AUTHOR(S) Boyan I. Dimitrov		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.		
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.		**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

This thesis combines regression, sentiment, and social network analysis to explore how Russian online media agencies, both overt and covert, affect online communication on Twitter when North Atlantic Treaty Organization (NATO) exercises occur. It explores the relations between the average sentiment of tweets and the activities of Russia's overt and covert online media agencies. The data source for this research is the Naval Postgraduate School's licensed Twitter archive and open-source information about the NATO exercises timeline. Publicly available lexicons of positive and negative terms helped to measure the sentiment in tweets. The thesis finds that Russia's covert media agencies, such as the Internet Research Agency, have a great impact on and likelihood for changing the sentiment of network users about NATO than do the overt Russian media outlets. The sentiment during NATO exercises becomes more negative as the activity of Russian media organizations, whether covert or overt, increases. These conclusions suggest that close tracking and examination of the activities of Russia's online media agencies provide the necessary base for detecting ongoing information operations. Further refining of the analytical methods can deliver a more comprehensive outcome. These refinements could employ machine learning or natural language processing algorithms that can increase the precision of the sentiment measurement probability and timely identification of trolls' accounts.

14. SUBJECT TERMS information operations, social media, Russian media outlets, Russia's covert online activity, IRA, Internet Research Agency, sentiment analysis, Twitter			**15. NUMBER OF PAGES** 79
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

THIS PAGE INTENTIONALLY LEFT BLANK

UNDERSTANDING THE ROLE OF OVERT AND COVERT ONLINE COMMUNICATION IN INFORMATION OPERATIONS

Boyan I. Dimitrov
Podpolkovnik, Bulgarian Air Force
Master of Radio and Television Engineering , Air Force Academy, Bulgaria, 1999
MIB, "St. Cyril and St. Methodius" University of Veliko Tarnovo, Bulgaria, 2011

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN DEFENSE ANALYSIS
(IRREGULAR WARFARE)

from the

NAVAL POSTGRADUATE SCHOOL
March 2022

Approved by: Timothy C. Warren
Advisor

Shannon C. Houck
Second Reader

Carter Malkasian
Chair, Department of Defense Analysis

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis combines regression, sentiment, and social network analysis to explore how Russian online media agencies, both overt and covert, affect online communication on Twitter when North Atlantic Treaty Organization (NATO) exercises occur. It explores the relations between the average sentiment of tweets and the activities of Russia’s overt and covert online media agencies. The data source for this research is the Naval Postgraduate School’s licensed Twitter archive and open-source information about the NATO exercises timeline. Publicly available lexicons of positive and negative terms helped to measure the sentiment in tweets. The thesis finds that Russia’s covert media agencies, such as the Internet Research Agency, have a great impact on and likelihood for changing the sentiment of network users about NATO than do the overt Russian media outlets. The sentiment during NATO exercises becomes more negative as the activity of Russian media organizations, whether covert or overt, increases. These conclusions suggest that close tracking and examination of the activities of Russia’s online media agencies provide the necessary base for detecting ongoing information operations. Further refining of the analytical methods can deliver a more comprehensive outcome. These refinements could employ machine learning or natural language processing algorithms that can increase the precision of the sentiment measurement probability and timely identification of trolls’ accounts.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION....................1

II. LITERATURE REVIEW....................5
A. INFORMATION OPERATIONS AND SOCIAL MEDIA....................5
B. METHODS OF STUDYING INFORMATION OPERATIONS....................11

III. INFORMATION WARFARE: THE BEAR AND THE DRAGON....................15
A. INFORMATION CONFRONTATION, INFORMATION WARFARE, AND INFORMATION OPERATIONS....................15
B. SOCIAL MEDIA IN INFORMATION WARFARE....................20

IV. RESEARCH METHODS....................29
A. HYPOTHESES....................29
B. DATA AND METHODS....................29
1. Datasets....................29
2. Sentiment Analysis and Text Processing of Social Media Data....................30
3. Network Topography Measures....................31
4. Dependent Variables....................31
5. Independent Variables....................32
6. Control Variables....................33
C. REGRESSION ANALYSIS....................34

V. REGRESSION RESULTS....................35
A. RUSSIAN ONLINE INFLUENCE ON SOCIAL MEDIA SENTIMENT....................35
1. Findings....................35
2. Analysis of the Findings....................38
B. NETWORK MEASURES AND RUSSIAN ONLINE INFLUENCE....................42
1. Analysis of the Findings....................44
C. DISCUSSION....................46

VI. CONCLUSIONS....................51

LIST OF REFERENCES....................55

INITIAL DISTRIBUTION LIST 61

LIST OF FIGURES

Figure 1. IRA and Russian Media Tweets vs. Sentiment in the NATO Conversation 39

Figure 2. Substantive Effects of IRA and Russian Media Tweets on the Probability of Positive and Negative Sentiment 40

Figure 3. NATO Exercises and IRA Tweets vs. Probability of Positive Sentiment 41

Figure 4. NATO Exercises and Distance to Russia vs. Probability of Positive Sentiment 42

Figure 5. NATO Exercises and IRA Tweets vs. Network Size 44

Figure 6. NATO Exercises, IRA Tweets, and Russian Media Tweets vs. Degree Centralization 45

Figure 7. NATO Exercises, IRA Tweets, and Russian Media Tweets vs. Clustering Coefficient....46

Figure 8. Russian Military Exercises and Distance to Russia vs. Probability of Positive Sentiment 49

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

Table 1. Sentiment Regression Models: Goodness of Fit 35

Table 2. Sentiment Regression Models: Base Terms 36

Table 3. Sentiment Regression Models: Multiplicative Interaction Terms 38

Table 4. Network Regression Models 43

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AIC	Akaike’s Information Criteria
BIC	Bayesian Information Criteria
BLM	Black Lives Matter
GDP	Gross Domestic Product
IO	Information Operation
IRA	Internet Research Agency
IRC	Information-Related Capabilities
LDA	Latent Dirichlet Allocation
MAE	Mean Absolute Error
MoDRF	Ministry of Defence of the Russian Federation
NATO	North Atlantic Treaty Organization
NPS	Naval Postgraduate School
ORA	Organizational Risk Analyzer
RMSE	Root Mean Squared Error

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

This thesis would not be possible without the tireless work of the teachers and instructors from the Defense Analysis Department of the Naval Postgraduate School. I am particularly grateful to my advisors Professor T. Camber Warren and Professor Shannon Houck, for their dedication and patience during my research. In addition, I am also thankful to the people in the International Graduate Program Office for the exceptional support that they provide for the international students.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

In recent years, social media has become significantly important as a source of information and means of communication. Many people worldwide have an account on at least one social media platform. Such platforms created a new environment, which enormously simplified exchanging ideas, messages, and knowledge. However, it also becomes a channel for manipulation, deception, and proliferation of extremist and radical ideologies. The popularity of such online services creates opportunities for state and non-state actors to manipulate a society's values and beliefs. The online media platforms become both a target and source of information campaigns to manipulate public perception. Governments worldwide work actively to develop online capabilities for influence to strengthen their power and accomplish their political agendas. Examples of such influence are the Russian government's online campaign against Ukraine in 2014[1] and campaigns by ISIS[2] and the Taliban[3] using social media to communicate their messages to target audiences.

These examples also demonstrate that the main actors in information operations operate overtly or covertly when influencing the targeted audiences. Official media agencies or journalists known for their affiliation to such outlets can openly deliver part of the narratives in such operations. However, this affiliation can affect their credibility and diminish the effects of the information operation. In contrast, media organizations that use covert methods can avoid being attributed to the source of the campaign, primarily by using fake online accounts. Therefore, they can infiltrate different communities to affect their online communication. Most importantly, they do not just push narratives that are in line

[1] Ulises A. Mejias and Nikolai E. Vokuev, "Disinformation and the Media: The Case of Russia and Ukraine," *Media, Culture & Society* 39, no. 7 (October 2017): 8–11, https://doi.org/10.1177/0163443716686672; Yevgeniy Golovchenko, Mareike Hartmann, and Rebecca Adler-Nissen, "State, Media and Civil Society in the Information Warfare over Ukraine: Citizen Curators of Digital Disinformation," *International Affairs* 94, no. 5 (September 1, 2018): 2, https://doi.org/10.1093/ia/iiy148.

[2] Imran Awan, "Cyber-Extremism: Isis and the Power of Social Media," *Society* 54, no. 2 (April 2017): 5, https://doi.org/10.1007/s12115-017-0114-0.

[3] Vincent Bernatis, "The Taliban and Twitter: Tactical Reporting and Strategic Messaging," *Perspectives on Terrorism* 8, no. 6 (2014): 6, http://www.jstor.org/stable/26297291.

with the official position of their state or organization. They can employ more aggressive methods and deliberately disseminate misleading, false, or divisive information to create in the targeted society disorientation and social confrontation. Consequently, both overt and covert media organizations significantly contribute to the overall effect of information operations. Their role in such activities needs exploring and constant monitoring to employ proper countermeasures. National security agencies need to develop the capacity to counter such online information campaigns. Successful and efficient counteraction depends on the reliable identification of the sources of information attacks, their target audiences, and their strategic messages.

However, researchers usually employ a single analytical method when exploring online information operations. For example, some of the most common methods include temporal analysis, data mining, or social network analysis.[4] This lack of an integrated approach delivers only partial results about the overall picture in the information campaign and prevents proper countermeasure planning. Thus, this thesis demonstrates that combining statistical analysis, social network analytic tools, text mining, and sentiment analysis can improve this approach. It shows that these methods provide the necessary results to build models that can explore the role of the different actors in information operations. This thesis's more comprehensive approach can increase the probability of timely identification of ongoing information operations, their audiences, main actors' probable location(s), and contribution to the overall effect of the operation. This research has great potential for improvement by adding new methods to the analysis. For example, machine learning algorithms can increase the precision of sentiment measurement. Additionally, dynamic network analysis and natural language processing can identify the most important topics in the online conversation and who contributes to these topics.

This thesis focuses on the online conversation on Twitter about the North Atlantic Treaty Organization (NATO), by examining the effects of military exercises over the

[4] Kai Shu et al., *Fake News Detection on Social Media: A Data Mining Perspective*, vol. 19, 2017, https://arxiv.org/pdf/1708.01967.pdf; Xinyi Zhou et al., "Fake News: Fundamental Theories, Detection Strategies and Challenges," in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (WSDM '19: The Twelfth ACM International Conference on Web Search and Data Mining, Melbourne VIC Australia: ACM, 2019), 836–37, https://doi.org/10.1145/3289600.3291382.

2013–2014 time period. It explores the relations between the average sentiment of messages in this online conversation and the activities of Russia’s overt and covert online media agencies. The thesis’s most important findings are that the messages generated by Russia’s covert agencies have a more substantial impact on the probability of sentiment change than those of the overt Russian media. Furthermore, the study finds that both the overt and the covert Russian actors more effectively influence positive rather than negative sentiments. Nevertheless, the evidence also shows that sentiment during NATO exercises becomes more negative as the activity of Russian covert media organizations increases. Furthermore, this thesis finds that during NATO exercises the trend toward negative sentiments increases if the sender’s location is closer to the borders of the Russian Federation. Finally, this study also finds that activity by the Internet Research Agency (IRA), a covert Russian social media influencer agency, and Russian media during NATO exercises increases the daily network size of users engaged in the NATO conversation, and this network becomes more interconnected. When NATO exercises occur, Russian media activities tend to increase the centralization of the daily network. By contrast, the IRA seems to operate as a more “egalitarian” network of small teams with a similar number of ties.

THIS PAGE INTENTIONALLY LEFT BLANK

II. LITERATURE REVIEW

A. INFORMATION OPERATIONS AND SOCIAL MEDIA

The major actors on the international political stage, such as the United States, Russia, and China, have similar views about confrontation that unfolds in the form of information. Whether they define it as an information operation or information warfare, its primary goals are to influence the adversary's political and military leadership's decision-making process, reduce the enemy's morale, and tarnish the population's confidence in its leaders and institution. All three countries acknowledge that this type of confrontation has multidimensional characteristics because it targets physical, informational, and cognitive aspects of the information environment. The Joint Chiefs of Staff's JP3-13 provides details the U.S. concept of information operations,[5] while I. N. Panarin presents the Russian view and provides that country's definitions for information warfare (*Russ. informacionnoe protivoborstvo*) and its primary goals.[6] In turn, Paul Charon and Jean-Baptist Jeangène Vilmer[7] provide extensive details about China's concept of information operations, examine multiple case studies, and explain how Beijing studies and exploits the lessons learned from the Russian and U.S. military campaigns. In addition to the work of Charon and Jeangène Vilmer, Nathan Beauchamp-Mustafaga and Michael Chase[8] and Elsa Kania[9] can provide complementary insight into the Chinese concept of "the Three Warfares"—psychological warfare, public opinion warfare, and legal warfare.

[5] The Joint Chiefs of Staff, *Joint Publication 3-13: Information Operations*, JP 3-13 (Washington, DC: Joint Chief of Staff, 2012), https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_13.pdf.

[6] I. N. Panarin, *Informatsionnaia Voina i Geopolitika*, [Information Warfare and Geopolitics] Velikii Put′ (Moskva: Pokolenie, 2006).

[7] Paul Charon and Jean-Baptist Jeangène Vilmer, *Chinese Influence Operations: A Machiavellian Moment* (Paris, France: Ministry for the Armed Forces, Institute for Strategic Research, 2021), https://www.irsem.fr/report.html.

[8] Nathan Beauchamp-Mustafaga and Michael S. Chase, *Borrowing a Boat Out to Sea: The Chinese Military's Use of Social Media for Influence Operations*, Policy Papers (Washington, DC: Johns Hopkins University School of Advanced International Studies, 2019).

[9] Elsa Kania, "The PLA's Latest Strategic Thinking on the Three Warfares," *China Brief* 16, no. 13 (August 2016), https://jamestown.org/program/the-plas-latest-strategic-thinking-on-the-three-warfares/.

Researchers examining the organization and capabilities of the Russian media propaganda machine explain that it has overt and covert components. The former encompasses the vast network of media outlets and news agencies whose editorial policies are under the Kremlin's direct control.[10] Monica Hanley and Andrey Kuzichkin provide valuable details about the organization of the Russian media landscape, the major media holding companies, and who owns them. The authors show that these media networks receive funds from some of the largest Russian state or private corporations such as Gazprom, VneshTorgBank, and Severstal. The authors also describe the specialized administrative structures that support the Russian media operations and the role of the Russian presidential administration in this process.

Similarly, Tod Helmus finds that the Kremlin uses its nexus of media companies in congruence with covert assets such as online trolls and bots.[11] He explores the difference in Moscow's tactics in the countries close to its borders and those far away. In the 'far abroad' states, Helmus observes that the Kremlin identifies radical or extremist groups from the conservative or liberal spectrum, tries to reinforce their extreme political views, and then provokes a conflict between them. In contrast, in the 'near abroad,' the targets are typically Russian ethnic and language minorities, Orthodox religious communities, or those with a shared memory about historical events.[12] In this light, Vasile Rotaru examines Moscow's "soft power" in the countries bordering Russia and provides details on the potential role of such groups in Russian influence operations.[13]

[10] Monica Hanley and Andrey Kuzichkin, *Russian Media Landscape:Structures, Mechanisms, and Technologies of Information Operations* (Riga, Latvia: NATO Strategic Communications Centre of Excellence, 2021), https://stratcomcoe.org/publications/russian-media-landscape-structures-mechanisms-and-technologies-of-information-operations/215.

[11] Todd C. Helmus, *Russian Social Media Influence: Understanding Russian Propaganda in Eastern Europe*, Research Report (Rand Corporation), RR-2237-OSD (Santa Monica, CA: RAND Corporation, 2018), https://www.rand.org/pubs/research_reports/RR2237.html.

[12] Helmus, 14–22.

[13] Vasile Rotaru, "Forced Attraction?: How Russia Is Instrumentalizing Its Soft Power Sources in the 'Near Abroad,'" *Problems of Post-Communism* 65, no. 1 (January 2, 2018): 37–48, https://doi.org/10.1080/10758216.2016.1276400.

The covert component of the Russian information operation capabilities includes groups such as the IRA that have used concealed actions to disseminate disinformation or employ divisive tactics among their target audiences. These IRA activities have attracted the attention of the scientific community worldwide. In particular, two events have primarily captured the researchers' focus: the 2016 U.S. presidential elections and the protest movement #BlackLivesMater (BLM).

Ahmer Arif et al.[14] demonstrate that during the BLM protests, the IRA used "Twitter and other online platforms to infiltrate politically active online communities." An important finding is that the IRA agents focused their efforts to manipulate the pro-BLM audience and those against the protest movement. Moreover, the Russian trolls' activities went beyond simply spreading disinformation on social media, as they were able to "connect to the cultural narratives, stereotypes, and political positions" of their target audiences." The authors point out that the IRA activities rely on three different components. The first is "the affordances of the online environment," or the inherent characteristics of Twitter and the other social media that define how the users use them. The second is "the social structures and behaviors of the online crowd," or the targeted audience's subgroups, hierarchy, or leaders and how they interact with each other. The third is "the improvised performances of agents that seek to leverage that crowd," or the actual actions to influence the target audience.

Similar findings are presented by Darren Linvill and Patrick Warren, who categorize IRA accounts into five general types: Right Troll, Left Troll, News Feed, Hashtag Gamer, and Fearmonger.[15] These authors observe that "within each type, accounts were used consistently, but the behavior across types was different, both in terms of

[14] Ahmer Arif, Leo Graiden Stewart, and Kate Starbird, "Acting the Part: Examining Information Operations Within #BlackLivesMatter Discourse," *Proceedings of the ACM on Human-Computer Interaction* 2, no. CSCW (November 2018): 1–27, https://doi.org/10.1145/3274289.

[15] Darren L. Linvill and Patrick L. Warren, "Troll Factories: Manufacturing Specialized Disinformation on Twitter," *Political Communication* 37, no. 4 (February 2020): 447–67, https://doi.org/10.1080/10584609.2020.1718257.

'normal' daily behavior and in how they responded to external events."[16] They compare the Internet Research Agency's organization to industrial enterprises that use "interchangeable parts" with specialized functions to achieve political goals.

Philip Howard et al. reveal how the IRA's activities sought to polarize U.S. society during the 2016 presidential election by engaging primarily far-conservative and far-liberal communities on social media. Specifically, the authors used qualitative and quantitative analysis to demonstrate that IRA operations targeted "African American voters to boycott elections or follow the wrong voting procedures" and incited "extreme right-voters to be more confrontational." [17]

Congressional hearings questioning social media companies' officials[18] regarding the Kremlin meddling in the U.S. election process and court indictments against IRA-linked individuals[19] also provide an important source of detailed information about the IRA's structure, leadership, and activities. During the hearings, a Twitter representative revealed that more than 47% of the IRA accounts were automated. In addition, he confirmed that Russian media company RT purchased ads that promoted election-related content that Twitter identified as being "inflammatory or low-quality."[20] The court indictment details the IRA's owners, management, departments, activities, and budget.

[16] Linvill and Warren, 447.

[17] Philip N. Howard et al., "The IRA, Social Media and Political Polarization in the United States, 2012–2018," *Oxford, UK: University of Oxford, Computational Propaganda Research Project, 2019*, October 2019, 48, https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1004&context=senatedocs.

[18] *Russia Investigative Task Force: Testimony before Permanent Select Committee on Intelligence*, 115th Cong. (2017) (statement of Sean J. Edgett Acting General Counsel, Twitter, Inc), accessed February 4, 2022, https://docs.house.gov/meetings/IG/IG00/20171101/106558/HHRG-115-IG00-Wstate-EdgettS-20171101.pdf.

[19] The U.S. vs. The Internet Research Agency LLC, 1:18-cr-00032-DLF, filed February 16, 2018, https://www.justice.gov/file/1035477/download.

[20] *Russia Investigative Task Force: Testimony before Permanent Select Committee on Intelligence*, 115th Cong. (2017) (statement of Sean J. Edgett Acting General Counsel, Twitter, Inc), https://docs.house.gov/meetings/IG/IG00/20171101/106558/HHRG-115-IG00-Wstate-EdgettS-20171101.pdf, 13.

Information operations have different manifestations, and they often result in "distraction, distortion, dismay, and disruption"[21] in the public perceptions of specific communities or the society in a given country as a whole. David Beskow and Kathleen Carley's research demonstrates that these activities could be message-driven or network-driven.[22] Message-driven and network-driven activities differ in their targets. The former aims at a public perception by spreading inaccurate or divisive information.[23] The latter aims to affect group dynamics by manipulating the target's social ties. For example, network-driven messages focus on removing a group's less radical opinion to transform it into a polarized echo chamber.[24] As a result, the main goal of the researchers studying information operations is to examine the dynamics of these processes and "to classify adversarial actors and their activities, assess and predict their impact, and design effective strategies for intervention and building the resilience of online communities."[25]

In addition, studies of information operations conducted through the internet agree that the dissemination of disinformation or fake news is their most recognizable feature, and how to detect them is the central question that needs an answer. The researchers approach this problem from different perspectives. For example, the data-oriented approach explores online communication by analyzing messages and their social context. Knowledge perspective fake-news detection compares verified news articles to knowledge

[21] Ben Nimmo, "Anatomy of an Info-War: How Russia's Propaganda Machine Works, and How to Counter It," *Central European Policy Institute* 15 (2015), https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/.

[22] David M. Beskow and Kathleen M. Carley, "Social Cybersecurity: An Emerging National Security Requirement," *Military Review*, March-April (2019), https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/MA-2019/Beskow-Carley-Social-Cyber.pdf.

[23] W. Lance Bennett and Steven Livingston, "The Disinformation Order: Disruptive Communication and the Decline of Democratic Institutions," *European Journal of Communication* 33, no. 2 (April 2018): 122–39, https://doi.org/10.1177/0267323118760317; Mejias and Vokuev, "Disinformation and the Media."

[24] R. Kelly Garrett, "Echo Chambers Online?: Politically Motivated Selective Exposure among Internet News Users," *Journal of Computer-Mediated Communication* 14, no. 2 (January 2009): 265–85, https://doi.org/10.1111/j.1083-6101.2009.01440.x; Elmie Nekmat, "Prosocial vs. Trolling Community on Facebook: A Comparative Study of Individual Group Communicative Behaviors," *International Journal of Communication* 12 (2018): 1–22, http://ijoc.org.

[25] Joshua Uyheng et al., "Interoperable Pipelines for Social Cyber-Security: Assessing Twitter Information Operations during NATO Trident Juncture 2018," *Computational and Mathematical Organization Theory* 26, no. 4 (December 2020): 3, https://doi.org/10.1007/s10588-019-09298-1.

in a trustworthy database.[26] The knowledge perspective fake-news detection method compares verified news articles to knowledge in a trustworthy database, while style-based detection captures "the differences in writing styles between fake and accurate news."[27] Another approach, the propagation perspective, explores the dissemination path of the news and indirectly detects fake news by assessing the credibility of the headlines, publishers, comments, and users.

Data mining, machine learning, and natural language processing methods are the major tools in implementing these approaches. Each tool differs from the others in its strategies, datasets, and techniques.[28] They are described in more detail in the following section of this thesis.

Other researchers correctly observe that combining these approaches may yield better results. Joshua Uyheng et al.[29] put the concept of social cyber-security and interoperability at the center of their work. Social cyber-security is "a multidisciplinary and multimethodological field that studies how to preserve the Internet as a free and open space for the exchange of information." [30] Carley and Beskow further examine how technology change and decentralization of information flow enables the emergence of social cyberthreats.[31] They explain that "technology has waived the requirement for physical proximity to influence society; and, the decentralization of information flows has reduced the cost of entry."[32]

[26] Kai Shu et al., *Fake News Detection on Social Media: A Data Mining Perspective*, vol. 19, 2017, https://arxiv.org/pdf/1708.01967.pdf.

[27] Xinyi Zhou et al., "Fake News: Fundamental Theories, Detection Strategies and Challenges," in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (WSDM '19: The Twelfth ACM International Conference on Web Search and Data Mining, Melbourne VIC Australia: ACM, 2019), 837, https://doi.org/10.1145/3289600.3291382.

[28] Zhou et al., "Fake News."

[29] Uyheng et al., "Interoperable Pipelines for Social Cyber-Security."

[30] Kathleen M. Carley et al., "Social Cyber-Security," in *Social, Cultural, and Behavioral Modeling*, ed. Robert Thomson et al., vol. 10899, Lecture Notes in Computer Science (Cham, Switzerland: Springer International Publishing, 2018), https://doi.org/10.1007/978-3-319-93372-6_42.

[31] Beskow and Carley, "Social Cyber-Security."

[32] Beskow and Carley, 122.

The analysis of "online content polluters, such as bots and trolls,"[33] has an important place in the field of social cyber-security. They both spread aggressive or disruptive messages. It should be noted that in contrast to trolls who are humans, bots are automated accounts.[34] The success of these "online polluters" has both technological and sociopsychological aspects. Thus, social cyber-security researchers have to combine methods in both computational social sciences.

Several prior Naval Postgraduate School (NPS) theses have also focused on social media to study influence in the information environment. For example, Eric Chan explored Russian influence operations and the historical background of IRA activity.[35] Meanwhile, James Morales provided valuable information on how to use tweets' sentiment analysis for assessing the anti-American mood in Pakistan and Japan. Particularly interesting is his study on how public perceptions change due to U.S. exercises in Japan.[36] Greg Selph et al. examine the relation between the civil conflicts in Nigeria, the Philippines, and Pakistan and the change in sentiment of tweets. The authors provide a valuable assessment of the significance of sentiment analysis for studying social media.[37]

B. METHODS OF STUDYING INFORMATION OPERATIONS

The most frequently used methods for analyzing information operations are dynamic network analysis, natural language processing, and machine learning. Machine

[33] Kyumin Lee, Brian David Eoff, and James Caverlee, "Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter," in *Fifth International AAAI Conference on Weblogs and Social Media*, Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (College Station, TX: Texas A&M University, 2011), 8; Emilio Ferrara, *Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election* (Los Angeles, CA: University of Southern California, Information Sciences Institute, 2017), 33, https://ssrn.com/abstract=2995809.

[34] David M. Beskow and Kathleen M. Carley, "It's All in a Name: Detecting and Labeling Bots by Their Name," *Computational and Mathematical Organization Theory* 25, no. 1 (March 2019): 24–35, https://doi.org/10.1007/s10588-018-09290-1.

[35] Elvis M. Chan, "Fighting Bears and Trolls: An Analysis of Social Media Companies and U.S. Government Efforts to Combat Russian Influence Campaigns during the 2020 U.S. Elections" (master's thesis, Naval Postgraduate School, 2021), 25–59, http://hdl.handle.net/10945/68309.

[36] James Morales, "Assessing Anti-American Sentiment through Social Media Analysis" (master's thesis, Naval Postgraduate School, 2016), 45, http://hdl.handle.net/10945/51587.

[37] Gregory R. Selph, Michael H. Crain, and Andrew Anderson, "Measuring Sentiment Response to Collective Violence through Social Media" (master's thesis, Naval Postgraduate School, 2018), 26, http://hdl.handle.net/10945/66275.

learning algorithms are very efficient in the automatic identification of fake accounts and bots in online conversations.[38] These algorithms are applied to aggregated datasets to discover patterns in bots' messaging behavior or relationships with other social network nodes. Uyheng et al. claim that "given a reasonably large dataset of a labeled bot and non-bot accounts, predictive models can be trained to discriminate between each type of account with decent accuracy (≥ 90%) across various contexts."[39] Other studies demonstrate that detecting trolling and opinion manipulation in news community forums and Twitter can use online messages' textual features.[40] When studying online information operations, another critical task is identifying the topics or central messages of the targeted conversations. Techniques such as topic modeling allow researchers to capture the core ideas and reveal critical aspects of a disinformation campaign. For example, the Latent Dirichlet Allocation (LDA) algorithm can extract topics from a corpus of texts using natural language processing.[41] Finally, dynamic network analysis provides quantitative methods to test hypotheses about users' influence in an online conversation. It analyzes how their characteristics and behavior change over time and how they interact. For example, on Twitter, the study of users' features and their ties (how they retweet, reply, or are mentioned) can deliver "more complex insights into online discourse."[42]

Uyheng et al.[43] convincingly demonstrate that these methods can discover patterns in Twitter users' behavior attributed to online information campaigns during NATO

[38] Fred Morstatter et al., "Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose," *ICWSM 2013*, June 2013, http://arxiv.org/abs/1306.5204; SiHua Qi, Lulwah AlKulaib, and David A. Broniatowski, "Detecting and Characterizing Bot-Like Behavior on Twitter," in *Social, Cultural, and Behavioral Modeling: 11th International Conference, SBP-BRiMS 2018 Washington, DC, USA, July 10–13, 2018 Proceedings 123*, n.d.

[39] Uyheng et al., "Interoperable Pipelines for Social Cyber-Security," 3.

[40] Carley et al., "Social Cyber-Security"; Todor Mihaylov, Georgi Georgiev, and Preslav Nakov, "Finding Opinion Manipulation Trolls in News Community Forums," in *Proceedings of the Nineteenth Conference on Computational Natural Language Learning* (Proceedings of the Nineteenth Conference on Computational Natural Language Learning, Beijing, China: Association for Computational Linguistics, 2015), 310–14, https://doi.org/10.18653/v1/K15-1032.

[41] David M. Blei, Andrew Ng, and Michael Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research* 3 (January 2003): 22.

[42] Kathleen M. Carley et al., "Toward an Interoperable Dynamic Network Analysis Toolkit," *Decision Support Systems* 43, no. 4 (August 2007): 1324–47, https://doi.org/10.1016/j.dss.2006.04.003.

[43] Uyheng et al., "Interoperable Pipelines for Social Cyber-Security," 5.

exercises. They apply LDA to determine what topics are prevalent in Twitter communication when NATO exercises occur and how they change over time. Next, the researchers explore the features of Twitter accounts to classify them as bots or not. For this purpose, they use a random forest machine learning model such as Bothunter.[44] In parallel, the authors employ different neural networks trained on a dataset of user descriptions for role identification (e.g., if the users are news agencies, reporters, and others) and for location prediction. In the end, they use an Organization Risk Analyzer (ORA) for individual and network drill-down to identify influential users, characterize the Twitter conversation's overall structure, and visualize the results. Uyheng et al.'s approach shows that combining computational tools can better analyze information operations in social media, as this approach can gain more complex insights into participants and their messages than single tools can.

Nonetheless, these results could be improved further by integrating different analytical tools or methods, such as sentiment analysis (especially in languages other than English), regression analysis, and network topology metrics' calculation. The change in the sentiment of tweets can play the role of a dependent variable in regression models to establish which factors have a statistically significant influence. The calculation of different topology metrics of the users' social network can reveal how its characteristics, such as centralization or interconnectedness, change under specific factors. This thesis complements the earlier analysis of the Russian covert and overt media agencies' role in influencing the information environment during NATO exercises in Europe. It combines regression, sentiment, and social network analyses to determine systematically how significant political-military events such as Allied exercises affect social media users' perceptions.

[44] David M. Beskow and Kathleen M. Carley, *Bot-Hunter: A Tiered Approach to Detecting & Characterizing Automated Activity on Twitter*, Xx (Pittsburgh, PA: Carnegi Mellon University, 2018), 1, https://www.researchgate.net/publication/326606376_Bot-hunter_A_Tiered_Approach_to_Detecting_Characterizing_Automated_Activity_on_Twitter.

THIS PAGE INTENTIONALLY LEFT BLANK

III. INFORMATION WARFARE: THE BEAR AND THE DRAGON

A. INFORMATION CONFRONTATION, INFORMATION WARFARE, AND INFORMATION OPERATIONS

The development of information technologies has led many military and political leaders worldwide to value the significance of influence and control over the information environment. Moreover, the widely accepted understanding is that this influence and control can go beyond its traditional force multiplier role in conventional military operations and become central in non-military confrontation.

In the United States, the Department of Defense defines information operations as disruptive actions aimed at the decision-making process of potential adversaries. These actions include employing information-related capabilities (IRC) to gain advantages in the information environment's three distinct dimensions: physical, informational, and cognitive.[45] JP 3-13 further specifies that the IRCs affect the ability of the targeted individual or group "to collect, process, or disseminate information before and after decisions are made."[46] Beyond the offensive aspect, information operations incorporate defensive actions as well. By employing IRC, the information operations aim to protect their own forces in the three-dimensional information environment from hostile activities that can undermine their own chain of command and the decision-making process. Thus, in its essence, an information operation seeks to change the adversary's situational awareness and diminish its information capabilities, in order to ultimately disrupt the adversary's decision-making process.

The Russian political elite, the military, and the academic community attach equally high importance to the process of influence and confrontation in the information environment. As the Chief of General Staff of the Russian Armed Forces, General Valery Gerasimov, points out that information warfare creates wide asymmetric opportunities to

[45] The Joint Chiefs of Staff, *JP 3-13*, 3.

[46] The Joint Chiefs of Staff, I–3.

reduce the enemy's combat potential.[47] In Russian documents on the topic, these activities are part of the broader concept of information warfare (*informacionnoe protivoborstva, Russ. информационное противоборство*). Igor Panarin postulates that the main goals of information warfare are to undermine the information security of the hostile state; to damage the integrity (stability) of its governmental and military control system; and to effectively influence the information provided to its leadership, political elite, and the systems that form the public's opinion, perception, and decision making. Panarin states that the ultimate goal of every state in information warfare is to achieve information superiority in the world information environment. Although he does not define the distinct dimensions of the information environment, he distinguishes two types of information warfare—technical and psychological. The first refers to the situation in which the targeted systems exchange and process information, and the latter focuses on the political elite's and public's cognitive capabilities and the alteration of public opinion. However, this differentiation demonstrates that the Russian concept also recognizes that information warfare can target different dimensions—physical or cognitive.[48]

The Ministry of Defense of the Russian Federation shares this definition. It specifies that, in wartime, information resources and capabilities become specific means to suppress enemy activity and deprive it of the opportunity to resist.[49] The Russian military leadership also recognizes the multidimensional character of information warfare. It can have both physical and informational-psychological impacts on the adversary's force. The former is the destruction of the enemy's information, radio-electronic, and computer networks, and the latter is the psychological influence on the population and the personnel of the armed forces of the opposing sides. The Russian concept incorporates both defensive and offensive aspects that must be considered in their unity. As both Panarin and the

[47] Valery Gerasimov, "The Science's Value is in the Prediction." *Military Industrial Courier* no. 8(476), 02/27-03/05/2013, https://vpk-news.ru/sites/default/files/pdf/ VPK_08_476.pdf, accessed February 1, 2022.

[48] Panarin, *Informatsionnaia Voina i Geopolitika [Information Warfare and Geopolitics]*, 172.

[49] *Enciclopedia of the Ministry of Defence of the Russian Federation*, s.v. "information warfare," accessed February 1, 2022, https://encyclopedia.mil.ru/encyclopedia/dictionary/details.htm?id=5221@morfDictionary.

military experts underline, it is equally important to protect one's own decision-making process and to influence the adversary's command and control system in information warfare.

The contemporary Russian information warfare concept uses and adapts the Soviet Union's "active measures" such as *dezinformatsiya* or the intentional spread of disinformation or false, inaccurate, and controversial information to mislead the target audience and shape adversaries' public opinion.[50] Arif et al. point out that such strategies are "ideologically fluid," and thus they are suitable to sow discord among very diverse political groups. They involve "harnessing existing public discontent by amplifying reductive social interpretations that confirm existing beliefs, support desired conclusions, or prompt certain strong emotions regarding groups of people and events."[51]

Beijing has carefully analyzed the campaigns of the United States in the Gulf War, Kosovo, Iraq, Afghanistan, and Ukraine, and what role information operations played in them.[52] It uses these lessons learned to adjust and update its concept for information warfare. The foundation of modern Beijing's concept are the theories of "Unrestricted War" and the "Three Warfares." The former posits that, in future conflict, the state has to implement "all means, including armed force or non-armed force, military and non-military, and lethal and non-lethal means to compel the enemy to accept one's interests."[53] The latter is part of so-called "political warfare," which includes the integrated usage of psychological warfare, public opinion warfare, and legal warfare. These three warfare types constitute "a discursive power over an adversary—that is, the power to control

[50] More details about Soviet disinformation campaigns can be found in Ladislav Bittman, *The KGB and Soviet Disinformation: An Insider's View* (Washington, DC: Pergamon-Brassey's, 1985); Peter Pomerantsev and Michael Weiss, *The Menace of Unreality: How the Kremlin Weaponizes information, Culture and Money* (New York: Institute of Modern Russia, 2014); and Alvin A. Snyder, *Warriors of Disinformation: American Propaganda, Soviet Lies, and the Winning of the Cold War: An Insider's Account* (New York: Arcade Publishing, 2014.

[51] Arif, Stewart, and Starbird, "Acting the Part," 3.

[52] Beauchamp-Mustafaga and Chase, *Borrowing a Boat Out to Sea*, 34–36.

[53] Liang Qiao and Xiangsui Wang, *Unrestricted Warfare*, 2nd ed. (Wuhan, China: Chongwen, 2011), 41.

perceptions and shape narratives that advance Chinese interests and undermine those of an opponent."[54]

Beauchamp-Mustafaga and Chase explain that public opinion warfare includes using "various media means and information resources" to "create a favorable public opinion environment for political initiative and military victory.[55]" This type of warfare has a relatively permanent nature and occurs both in wartime and peacetime. It finds expression in "long-term infiltration into the objects of the society and culture's deep structure, changing the awareness and conviction of the enemy masses."[56] Charon et al. specify that China's general idea of public opinion refers to the terms "public emotion" and "public opinion." The first term is "subjective interpretation of certain social realities,"[57] and therefore, it is an individual's perception. At the same time, the second emphasizes "the socio-political attitudes generated by social interactions, and thus it is the collective majority opinion."[58] Hence, control over public opinion is inextricably tied with control over public emotions. Charon and Jeangène Vilmer describe the core of public opinion warfare as the "cognitive orientation of the masses, to excite their emotions and to constrain their behavior."[59]

Beauchamp-Mustafaga and Chase point out that there is a difference between public opinion warfare and psychological warfare. The latter is "more focused and concentrated in wartime."[60] According to their analysis of strategic documents of the People's Liberation Army, Charon and Jeangène Vilmer assert that Beijing distinguishes four types of psychological warfare: "coercion," "mystification," "division," and "defense."[61] The first seeks to force the adversary to adopt a particular behavior; the

[54] Beauchamp-Mustafaga and Chase, *Borrowing a Boat Out to Sea*, 8.

[55] Beauchamp-Mustafaga and Chase, 9.

[56] Kania, "The PLA's Latest Strategic Thinking on the Three Warfares," 2.

[57] Charon and Jeangène Vilmer, *Chinese Influence Operations*, 30.

[58] Charon and Jeangène Vilmer, 30.

[59] Charon and Jeangène Vilmer, 48.

[60] Beauchamp-Mustafaga and Chase, *Borrowing a Boat Out to Sea*, 8–9.

[61] Charon and Jeangène Vilmer, *Chinese Influence Operations*, 49–50.

second spreads confusion and misleads; the third uses the adversary's weaknesses and domestic disagreements to hinder its decision-making process, ruin fighters' morale, and diminish public confidence. The fourth protects one's own troops' morale from the enemy's influence.[62] Related to psychological warfare is the concept for "cognitive domain operations." As Beauchamp-Mustafaga and Chase explain, China considers these operations as the next step in the evolution of warfare, "moving from the natural and material domains—land, maritime, air, even electromagnetic—into the ephemeral, namely the human mind."[63] Such operations target the enemy's cognitive thinking and decision-making processes with the means of psychological warfare to disrupt. Their ultimate goal is to achieve "mind superiority."[64]

The third component of the "Three Warfares" concept, legal warfare, corresponds to the strategic use of the law. Its essence is to exploit the legal provisions of international and national laws to provide legitimacy for Chinese demands or policies.[65] Its goal is to "attain normative superiority" that can justify the use of force during a conflict, or when confrontation ends "to retain any gains or to claim its due."[66]

The "Three Warfares" concept demonstrates that China, like the United States and Russia, also sees information warfare as a multidimensional phenomenon and recognizes the importance of operations in the information and cognitive dimensions. Beijing combines contemporary theories such as cognitive-domain operations, discursive power, and political warfare with older concepts such as "active measures," which China inherited from its historical ties with the former Soviet Union. This combination, together with the lessons learned from its main allies and adversaries, allows China to adapt its strategies and doctrines successfully to exploit the benefits and weaknesses of today's information environment.

[62] Charon and Jeangène Vilmer, 49.

[63] Beauchamp-Mustafaga and Chase, *Borrowing a Boat Out to Sea*, 10.

[64] Charon and Jeangène Vilmer, *Chinese Influence Operations*, 31.

[65] Charon and Jeangène Vilmer, 31.

[66] Charon and Jeangène Vilmer, 51.

In conclusion, all three major actors in the international political system adopt strategies and develop capabilities to gain an advantage over the information environment. As described in the literature just reviewed, the ultimate goals of information operations are to hinder the decision-making process of the adversary's political leadership and military commanders, affect the morale of the population and armed forces of the hostile nation, and protect a country's own forces and political system from malicious influence. Thus, information operations have both defensive and offensive aspects.

A common understanding is that the information environment has different dimensions that define two specific directions of confrontation. The first targets the physical dimension and the technical systems that transmit, receive, and store information; the second aims at psychological impacts on the target audience by manipulating the information and cognitive dimensions. In recent years, with the increase in online connectivity and rapid expansion of social media such as Facebook, Instagram, and Twitter, the latter dimension has grown more significant.

B. SOCIAL MEDIA IN INFORMATION WARFARE

This section explores the main characteristics of the information operations conducted on social media, and on Twitter in particular, by the adversaries of the United States and NATO. Although the previous chapter makes clear that China has significant capabilities and ambitions to dominate the global information environment, until recently, Beijing's focus has been predominantly on influencing its close neighborhood (Hong Kong, Taiwan), the Pacific region, and North America. This thesis, however, focuses primarily on the activities of various entities connected to the Russian Federation as they relate to the possible effects of hostile information operations conducted during NATO exercises in Europe. Samantha Bradshaw and Philip Howard point out that the social media information environment is complex and constantly evolving. In this environment, various actors pursue their political goals by creating and disseminating narratives designed to provoke a specific reaction in the targeted audience. At the same time, these audiences also are a diverse community, and they react differently to these narratives. Often, they contribute to the further development of the narrative by adding links to other content such as video, pictures, or music.[67]

Various Russian companies and organizations are involved in social media information operations. Some of them operate openly, such as the television channel RT (formerly Russia Today) and the news agency Sputnik, while others work covertly, such as the IRA. This combination allows the Kremlin to run "large-scale and complex information operations" with actors "at varying levels of attribution."[68]

The overt companies comprise a vast network that connects to domestic and foreign audiences through online channels on different social media platforms and at the same time creates more traditional content for TV channels, newspapers, and news agencies. These two forms of communication cannot be entirely separated because the traditional media products are also disseminated online.[69] A 2021 report by the NATO Center of Excellence for Strategic Communication reveals that the RT news service is central in Moscow's information operations. This news service was created in 2005 under Russia Today's brand, but later it changed its name to RT. The non-profit organization ТВ-Новости/TV-News controls RT; however, RT's funding comes entirely from the Russian state budget. In 2020, the subsidy to TV-News was more than 360 million U.S. dollars. Around 83% of these funds are dedicated to media content production. Margarita Simonyan has been in charge of the main editorial office of RT since 2005.[70] She and her husband have tight connections to members of the presidential administration, such as its First Deputy Chief of Staff, Alexei Gromov. He is one of RT's creators and maintains the Kremlin's control over the work of the largest newspapers, television channels, and news agencies.[71] RT also has a YouTube channel, which was the first "to reach 1 billion views … and later it was

[67] Philip N. Howard and Samantha Bradshaw, "The Global Organization of Social Media Disinformation Campaigns," *The Journal of International Affairs, SIPA* 71 (September 2018): 9, https://jia.sipa.columbia.edu/global-organization-social-media-disinformation-campaigns.

[68] Helmus, *Russian Social Media Influence*.

[69] Helmus, 26.

[70] Hanley and Kuzichkin, *Russian Media Landscape*, 22.

[71] Hanley and Kuzichkin, 11.

also the first to reach 10 billion views."[72] In addition, RT has a vast network of Twitter accounts in different languages, such as @RT_COM, @RT_America, @RTUKnews, @de_rt_com, and @RTarabic_Bn. Each of these accounts has between several thousand and several million followers. Simonyan herself has 531.7 thousand Twitter followers. In a congressional hearing, one Twitter official testified that "@RT_COM and @RT_America together spent $516,900 in advertising in 2016. As a result, they were able to promote 1,912 Tweets and generate approximately 192 million impressions[73] across all ad campaigns.[74]

Another Russian media group with significant input into information operations is Rossiya Segodnya. It is state-owned, and in 2020 the company received 100 million U.S. dollars from the state budget.[75] Rossiya Segodnya is a holding company that controls influential media outlets such as RIA Novosti, Sputnik, InoSMI, Baltnews, and others. These news agencies publish their information products on various online platforms, including Twitter.

Russia's network of existing official media channels significantly contributes to Moscow's capability to organize and conduct information operations. Its main contributions are to receive legally directed funds from the Russian state for media production and to operate legally in different countries. In addition to creating content in several different languages, this arrangement helps such media outlets to access broad audiences. As Hanley and Kuzichkin point out, RT and Sputnik use this opportunity and customize their media products *"*depending on the audience and the strategic objective."[76] The authors explain that for its Latin American audience RT produces leftist content in the

[72] Hanley and Kuzichkin, 22.

[73] Twitter defines impressions as viewing the content of the tweet. For details see *Russia Investigative Task Force: Testimony before Permanent Select Committee on Intelligence*, 3.

[74] *Russia Investigative Task Force: Testimony before Permanent Select Committee on Intelligence*, 13.

[75] Hanley and Kuzichkin, *Russian Media Landscape*, 25.

[76] Hanley and Kuzichkin, 9.

Spanish language, while its products for Eastern European countries emphasize national or historical ties to Russia or praise the strength of the Russian economy.[77]

Complementing these public media organizations, the Kremlin's information operations arsenal has companies that use covert methods, including fake accounts and divisive and fake information posting. A notorious example is the Internet Research Agency, a Russian corporation based in Saint Petersburg. Various authorities and scientists have revealed the corporation's interference in the 2016 U.S. Presidential election[78] and the Black Lives Matter movement's protests.[79] According to Howard et al., the IRA has been using Twitter since 2009. The initial focus was on the Russian audience, and the tweets were primarily in the Russian language. The company began targeting English-speaking users at a slow pace in 2013. By the beginning of 2014, however, its engagement with English-speaking users had increased and grew significantly by the end of the year.[80]

Although the company had direct links to the Russian government, the IRA primarily received funding for its operations from two companies, Concord Management and Consulting and Concord Catering. Both were under the control of Nikolay Prigozhin,[81] who is also the owner of Wagner (a private military company) and has close connections to the Kremlin elite.[82] The IRA's activities also received funding as part of a large interference operation named "Project Lakhta" that targeted audiences in the United States and Russia, France, and other countries. As of 2016, the IRA's monthly budget was over 1.25 million U.S. dollars.[83]

[77] Hanley and Kuzichkin, 9.

[78] *The U.S. vs the Internet Research Agency LLC*, 1:18-cr-00032-DLF.

[79] Arif, Stewart, and Starbird, "Acting the Part."

[80] Howard et al., "The IRA, Social Media and Political Polarization in the United States, 2012–2018," 10.

[81] *The U.S. vs the Internet Research Agency* at 6–7.

[82] Nathaniel Reynolds, *Putin's Not-So-Secret Mercenaries: Patronage, Geopolitics, and the Wagner Group*, *The Return of the Global Russia* (Washington, DC: Carnegie Endowment for International Peace, 2019), 9–12, https://carnegieendowment.org/files/GlobalRussia_NateReynolds_Vagner.pdf.

[83] *The U.S. vs. the Internet Research Agency* at 7.

The IRA's organizational structure was complex, and it had hundreds of employees "ranging from creators of fictitious personas to technical and administrative support."[84] The company had management and specialized graphic design, data analysis, search engine optimization, finance, information technology, and translation departments. The latter allowed its operators to post information on social media in different languages and expand its global reach.[85] The IRA's staff members worked two 12-hour shifts to match regular users' activity in different time zones, as they did during the 2016 U.S. presidential election. Their tasks were "to create social media accounts that appeared to be operated by U.S. persons…[and]… create political intensity through supporting radical groups, users dissatisfied with [the] social and economic situation and oppositional social movements."[86] The employees were very productive: they created 36,746 accounts related to the 2016 presidential election. They posted approximately 1.4 million election-related tweets from September through November 2016, and these tweets received approximately 288 million impressions.[87] The IRA's management was responsible for receiving and evaluating the impact of the organization's online social media operations and the size of the online U.S. audiences reached through IRA's posts. They measured "different types of engagement with the posts (such as likes, comments, and reposts), changes in audience size, and other metrics."[88]

These Russian organizations, both the overt and covert ones, involved in information operations have the capability to engage global audiences on various social media platforms and in different geographical locations. Online, such organizations operate on Twitter, Facebook, Instagram, and geographically, they can influence audiences both far and near their borders.[89] However, these targeted social groups in the 'far abroad' and

[84] *The U.S. vs. the Internet Research Agency* at 6.

[85] *The U.S. vs. the Internet Research Agency* at 6.

[86] *The U.S. vs. the Internet Research Agency* at 14.

[87] *Russia Investigative Task Force: Testimony before Permanent Select Committee on Intelligence*, 10.

[88] *The U.S. vs. the Internet Research Agency* at 15.

[89] Helmus, *Russian Social Media Influence*, 32–39.

‘near abroad’ have different characteristics, forcing the Kremlin to adapt its influence strategies.

The so-called ‘near abroad’ includes the former Soviet republics and most East European states. Such countries usually have sizable Russian minorities, Russian-speaking citizens, or groups that share common memories or interpretations of history. These communities are both targets and conduits of Russian “soft power” and influence.[90] Moscow’s most important narratives in these countries include tropes about traditional (Russian) conservative values, such as family and orthodoxy, a shared fear of violent revolutions, and the West’s betrayal and moral degradation.[91]

In contrast to the ‘near abroad,’ Russian information operations in the ‘far abroad’ countries focus on radical political groups, social movements, and religious groups. Andrew Weisburd, Clint Watts, and Jim Berger argue that Moscow’s information campaigns fall into four categories: political, financial, social, and conspiracy. The political and financial campaigns aim to slander political leaders, undermine governmental institutions’ credibility, or erode trust in the financial system or experts. The social objectives of such operations seek to “undermine the fabric of society.”[92] Finally, the dissemination of conspiracy theories promotes images of “global calamity while questioning the expertise of anyone who might calm those fears.”[93]

Ben Nimmo formulates four specific tactics that the Kremlin uses to achieve these goals: dismiss, distort, dismay, and distract. The dismiss tactics deny the truth of the facts that contradict Moscow’s narratives or tarnish the credibility of the source of these facts. The distorted tactics mix cherry-picked facts, lies, and disinformation. Since the ultimate goal is to sow doubts in the designated audiences, the truthfulness of the dismissed facts or blatant falsehood of the distorted message is not essential. All these tactics need only to

[90] Rotaru, “Forced Attraction?,” 3–10.

[91] Helmus, *Russian Social Media Influence*, 10.

[92] Andrew Weisburd, Clint Watts, and Jim Berger, “Trolling for Trump: How Russia Is Trying to Destroy Our Democracy,” War on the Rocks, November 6, 2016, https://warontherocks.com/2016/11/trolling-for-trump-how-russia-is-trying-to-destroy-our-democracy/.

[93] Weisburd, Watts, and Berger.

disorient the targeted population or social groups. The dismay tactics intimidate and induce fear and anxieties in the audience, usually through harassment and personal threats. Finally, the distraction techniques turn the "attention away from the activities of Russia and its allies by launching accusations elsewhere."[94] Overall, as Peter Pomerantsev and Michael Weiss observe, "the aim of this new propaganda is not to convince or persuade, but to keep the viewer hooked and distracted, passive and paranoid, rather than agitated to action."[95]

Linvill and Warren explored the implementation of these tactics by the IRA from June 19, 2015, to December 31, 2017, and "identified five categories of IRA-associated Twitter handles, each with unique patterns of behaviors."[96] These categories are Right Troll, Left Troll, News Feed, Hashtag Gamer, and Fearmonger. The Right Troll accounts posted nativist and right-leaning populist messages. The Left Troll handles published "socially liberal messages, with an overwhelming focus on cultural identity."[97] The News Feed accounts "overwhelming presented themselves as the U.S. local news aggregators,"[98] while the Hashtag Gamer handles promoted hashtag games, some of which had socially or politically divisive hashtags. The fifth category, Fearmonger trolls, posted news about crisis events, usually fabricated. The authors observed that Left and Right Trolls were more active than the other types, but their activity varied significantly on a day-to-day basis. The News Feed accounts tweeted at a relatively consistent rate throughout the explored period. By contrast, the Hashtag Gamer handles were "very active during and after the election season, but by the summer of 2017, they [were] nearly silent."[99] All five troll types changed their behavior under specific political circumstances. For example, when John Podesta's email was leaked to the press for the first time (8:30 pm UTC on October 7,

[94] Nimmo, "Anatomy of an Info-War: How Russia's Propaganda Machine Works, and How to Counter It."

[95] Peter Pomerantsev and Michael Weiss, *The Menace of Unreality:How the Kremlin Weaponizes Information, Culture and Money* (New York, NY: The Institute of Modern Russia, 2014), 11, https://www.almendron.com/tribuna/wp-content/uploads/2015/08/The_Menace_of_Unreality_Final.pdf.

[96] Linvill and Warren, "Troll Factories," 451.

[97] Linvill and Warren, 452.

[98] Linvill and Warren, 452.

[99] Linvill and Warren, 453.

2016), the News Feeds trolls continued to post messages at their regular low rate. Meanwhile, the Hashtag Gamers' activity which was very high in the late hours of October 5, 2016, went silent in the following days. In contrast, the Left and Right trolls became active on October 6, 2016, and tweeted at a high rate until October 7, 2016.[100]

In conclusion, the Russian Federation has the capability to organize and implement online information operations through its covert and overt assets. RT news network and Sputnik news agency are among the most powerful tools in the Russian online arsenal. Under the Kremlin's strict control, they deliver and popularize narratives to their vast audience in countries in the near and the far abroad. Additionally, Moscow has created companies that use covert techniques to engage carefully selected social groups in order to manipulate their perceptions and emotions. In such cases, the main goal is to sow distrust, disorient targeted groups' value systems, and amplify social divisions. An example of this type of covert organization is the Internet Research Agency, which had hundreds of employees, and its leadership had close connections to the Kremlin's political elite. Studies of the IRA's activities and tactics revealed possible patterns in its operations on Twitter. Most importantly, research has delineated the types of accounts that IRA operators used for their tasks. These accounts revealed distinct posting behaviors and shared specific information.

Russia's overt and covert organizations have developed a global reach and can target audiences in various languages. Typically, in the 'near abroad' countries, the target is local Russian-speaking communities, Russian ethnic minorities, and other groups that share cultural, historical, or religious identity with Russia. For these audiences, Moscow's narratives include positive interpretations of the Russian political and cultural model and negative depictions of the Western political, financial, and social systems. The 'far abroad' narratives aim at damaging social cohesiveness and confidence in institutions. Since their goal is to reinforce the most radical opposing opinions, they may also include positive and negative messages regarding a particular topic.

[100] Darren L. Linvill and Patrick L. Warren, *Troll Factories: The Internet Research Agency and State-Sponsored Agenda Building* (Clemson, SC: Clemson University, 2018), 11, http://pwarren.people.clemson.edu/Linvill_Warren_TrollFactory.pdf.

The next chapter details the research methods used in this thesis to examine user activity on one particular social media platform, Twitter, between July 2013 and August 2014, to identify factors related to IRA and Russian media efforts specifically influencing sentiments in tweets about NATO.

IV. RESEARCH METHODS

A. HYPOTHESES

This research examines one year of user activity on Twitter from July 2013 to August 2014. The purpose is to identify the significant factors that influence sentiments related to tweets on a particular topic—in this case, all tweets in which the text included the acronym NATO.

The first hypothesis is that the overall sentiment in the online conversation on NATO will become more negative if there is an increase in IRA and Russian media activity (Hypothesis 1). The second hypothesis is that while NATO exercises are occurring the sentiment of the online conversation related to NATO will become more negative as the activity of Russian media organizations, whether covert or overt, increases. Third, this thesis hypothesizes that the change in sentiment during NATO exercises will be more negative in a country closer to the borders of the Russian Federation (Hypothesis 3). Fourth, this thesis also hypothesizes that IRA and Russian media activity during NATO exercises will increase the size of the daily user network for the NATO conversation, and that network will become more centralized and interconnected (Hypothesis 4).

B. DATA AND METHODS

To test the hypotheses formulated in Section A this thesis examines specific datasets and applies different methods, such as regression and sentiment analysis, text mining, and social network analysis.

1. Datasets

This thesis uses the NPS-licensed Twitter archive as the primary data source. It contains a random sample of ten percent of the tweets posted globally between August 1, 2013, and July 31, 2014.

A search query was constructed based on the acronym NATO (in the English, French, Chinese, and Cyrillic alphabets) and a list of handles connected to IRA and Russian online media outlets. The U.S. House of Representatives Permanent Select Committee on

Intelligence published these handles in June 2018.[101] The query's results are stored in two datasets. In the first, the results are aggregated in "country-date" units. In the second dataset, each row represents a separate tweet with its text, and each message is represented as a "point" at a latitude/longitude/time location.

2. Sentiment Analysis and Text Processing of Social Media Data

Social media sentiment analysis relies on dictionaries containing words with predetermined positive or negative sentiment scores. In this research, two separate sentiment lexicons[102] are used, one with negative words and one with positive words. These dictionaries are in the English language; therefore, the text of every tweet needs translation before further processing.

Since this research dataset contains information about the tweet's language, the initial dataset was split into smaller sub-datasets according to the tweet's language. The result was 45 distinct sub-datasets organized by "tweet ID" and "text." These sub-datasets were translated using the Google Translate website at the next stage. After translation, each tweet's text went through text processing, including tokenization and cleaning of redundant symbols and words such as punctuation signs, special characters, or symbol combinations (e.g., @, RT, and URLs).

Next, cleaned tokens (words) were compared to the content of the positive and negative dictionaries. Depending on which lexicon contained a match for a word, each word received a sentiment score accordingly. Each word's score was calculated with Formula 1 to distinguish between a positive and a negative sentiment:

$$S_{word} = Positive\ score - Negative\ score \tag{1}$$

Words with a positive sentiment were coded as 1, and the negative as -1. If there was no match or the word was found in both dictionaries, the sentiment score was 0 or neutral.

[101] *The Internet Research Agency's Handles as of June 2018*, U.S. House of Representatives Permanent Select Committee on Intelligence (Washington, DC, 2018), https://intelligence.house.gov/uploadedfiles/ira_handles_june_2018.pdf.

[102] Minqing Hu and Bing Liu, *Mining and Summarizing Customer Reviews*, Research Track Paper (Chicago, IL: University of Illinois at Chicago, 2004).

Finally, the overall sentiment of the tweet was calculated using Formula 2 as the sum of the sentiment scores of its tokens:

$$S_{tweet} = \sum_{i=1}^{n} S_{word_i} \tag{2}$$

3. Network Topography Measures

After the sentiment score of the tweets was calculated, edge lists were created for every day in the research period. Two Twitter accounts have a tie between them if one of them retweeted, mentioned, or quoted the other's tweet. The attribute list includes username, language, followers, country name, date when the user tweeted, and two flags indicating whether the user was an IRA member or a Russian official media outlet. Using R-package "igraph,"[103] it was possible to calculate the network size, density, local clustering coefficient, and degree centralization score for all users in 365 social networks.

4. Dependent Variables

The mean sentiment of the tweets in the NATO conversation (as defined earlier) was the dependent variable used to test Hypotheses 1 through 3. It is aggregated to a "country-date" unit of analysis. It is an ordered categorical variable with three levels. The first corresponds to negative sentiment, the second to neutral sentiment, and the third to positive sentiment.

Other dependent variables based on network structure, such as network size, local clustering coefficient, and degree centralization, were used to test Hypothesis 4. The network size is based on the number of accounts in the network. It provides information on how many users participated in the explored NATO online conversation. The change in network size during the research period can provide information on how the network of participants was growing or shrinking.[104] The local clustering coefficient is an interconnectedness measure known as average ego-network density. It provides

[103] Gabor Csardi and Tamas Nepusz, "The Igraph Software Package for Complex Network Research," *InterJournal Complex Systems* (2006): 1695, https://igraph.org.

[104] Daniel Cunningham, *Understanding Dark Networks: A Strategic Framework for the Use of Social Network Analysis* (Lanham, MD: Rowman & Littlefield, 2016), 86.

information on how interconnected the actors of the ego network are.[105] The degree centralization score "uses the variation of the actor's [degree] centrality within the network to measure the level of centralization."[106] If the degree centralization is high, the users' degree centrality scores significantly vary. Therefore, if the score is high, one or a few accounts are significantly more active than the others.[107]

5. Independent Variables

The independent variables cover five categories: spatial effects, network topography, levels of online activity, and the dates of NATO and Russian military exercises.

The distance to the Russian border is the only spatial variable. It represents the shortest distance between the Russian state border and every tweet's location, measuring 0 for tweets originating inside Russia.

Edge density, the only social network topography measure, is an independent variable in the models that test Hypotheses 1 through 3. It introduces information about the behavior of daily user accounts into the model. Its change reflects the emergence or disappearance of the network's ties (as defined earlier).

The Twitter-related variables count the number of tweets originating from IRA or Russian overt accounts by day.

Finally, the exercise-related variables provide information measuring the dates on which NATO or Russian military exercises occurred, coded as 1 for the dates of the exercises and 0 for all other dates. Jan Brzezinski and Nicholas Varangis summarized NATO and Russian military activity over this period, and provided information about their timeline, location, and participants.[108] There were four NATO military drills in Eastern

[105] Cunningham, 100.

[106] Cunningham, 87.

[107] Cunningham, 87.

[108] Jan J. Brzezinski and Nicholas Varangis, "The NATO-Russia Exercise Gap," *Atlantic Council* (blog), February 23, 2015, https://www.atlanticcouncil.org/blogs/natosource/the-nato-russia-exercise-gap/.

Europe and Scandinavia in the research period. "Steadfast Jazz" took place in Poland and Lithuania (November 2–9, 2013), Norway hosted "Cold Response" (March 7–21, 2014), Estonia hosted "Spring Storm 14/Steadfast Javelin I" (May 5–23, 2014), and "Saber Strike 2014" took place in the Baltic States (June 9–20, 2014). The Russian Federation had three large-scale exercises in the research period. They took place in the Western Military District (Zapad-13, September 17–26, 2013), in the Western and Central Military Districts (February 26–March 3, 2014), and the Central Military District (June 21–28, 2014).

6. Control Variables

The control variables in the tested model are the total number of tweets, the country's population, gross domestic product per capita, political regime type, and access to the internet.

Gross domestic product (GDP) per capita and population provide information about the relative size and prosperity of each country. Access to the Internet provides information about the degree to which the population can use online media as a source of information. The World Bank Development Indicators[109] are the source of these three variables.

The political regime variable is derived from the Polity5 Project dataset.[110] The model uses the revised combined polity score (Polity2), adapted for time-series analysis. It ranges from 10 (strongly democratic) to -10 (strongly autocratic). Introducing this variable in the model provides an opportunity to examine how the ruling regime or the type of government influences the effectiveness of online information operations.

The total number of tweets variable provide information about the overall daily activity in Twitter

[109] World Bank, "World Development Indicator," accessed February 27, 2022, https://datatopics.worldbank.org/world-development-indicators/.

[110] Center for Systemic Peace, "Polity5 Project, Political Regime Characteristics and Transitions, 1800–2018," April 23, 2020, https://www.systemicpeace.org/inscrdata.html.

C. REGRESSION ANALYSIS

The initial analysis presented here utilizes ordered logit regression models. Its goal is to test the formulated hypotheses concerning the relationship among online sentiments, military exercises, and the activity of Russian overt and covert agents. The regression analysis tests four different models. It starts with a simple additive form and gradually introduces different multiplicative interaction terms. These terms examine different relations among the independent variables in accordance with the formulated Hypotheses 1 to 3. A log transformation is applied to all independent variables in the model except *NATO Exercise*, *Russian Exercise*, and *Political Regime* in order to reduce the skew in their values. The full version of the sentiment model is as follows:

$$\begin{aligned}
&\text{Sentiment} = \\
&\quad \beta_0 + \beta_1\text{NATO Exercise} + \beta_2\text{Russian Exercise} + \beta_3\text{IRA Tweets} \\
&\quad + \beta_4\text{Distance to Russia} + \beta_5\text{Russian Media Tweets} + \beta_6\text{Edge Density} \\
&\quad + \beta_7\text{Political Regime} + \beta_8\text{GDP} + \beta_9\text{Population} \\
&\quad + \beta_{10}\text{Access to Internet} + \beta_{11}\text{Total Tweets} \\
&\quad + \beta_{12}(\text{Russian Exercise} * \text{Distance to Russia}) + \beta_{13}(\text{NATO Exercise} * \text{Distance to Russia}) \\
&\quad + \beta_{14}(\text{Russian Exercise} * \text{IRA Tweets}) + \beta_{15}(\text{NATO Exercise} * \text{IRA Tweets}) \\
&\quad + \beta_{16}(\text{Political Regime} * \text{IRA Tweets}) + \beta_{17}(\text{NATO Exercise} * \text{Russian Media Tweets})
\end{aligned}$$

In addition, six log-linear models are used to test Hypothesis 4. Each pair of models tests a different dependent variable—network size, degree centralization, or clustering coefficient—with the first model providing a baseline additive specification and the second model including multiplicative interaction terms. The full version of each model is given by:

$$\begin{aligned}
&\text{Network Metric} = \\
&\quad \beta_0 + \beta_1\text{IRA Tweets} + \beta_2\text{Russian Media Tweets} + \beta_3\text{Distance to Russia} + \beta_4\text{NATO Exercise} \\
&\quad + \beta_5\text{Political Regime} + \beta_6\text{GDP} + \beta_7\text{Population} + \beta_8\text{Access to Internet} + \beta_9\text{Total Tweets} \\
&\quad + \beta_{10}(\text{NATO Exercise} * \text{IRA Tweets}) \\
&\quad + \beta_{11}(\text{NATO Exercise} * \text{Russian Media Tweets}) \\
&\quad + \beta 2(\text{Political Regime} * \text{IRA Tweets})
\end{aligned}$$

V. REGRESSION RESULTS

A. RUSSIAN ONLINE INFLUENCE ON SOCIAL MEDIA SENTIMENT

1. Findings

Tables 1, 2, and 3 show the results of four ordered logit models and their performance. The first table includes the goodness of fit scores for the different models, the second presents the base term coefficients, and the third shows the multiplicative interactive terms in the tested regression models. The analysis of models' goodness of fit scores reveals that the models achieve similar overall error rates, shown by the mean absolute error (MAE) and the root mean squared error (RMSE), while Model 4 performs slightly better than the others according to the Akaike's Information Criteria (AIC) scores. Thus, Model 4 is used for further analysis.

Table 1. Sentiment Regression Models: Goodness of Fit

	Sentiment			
	Model 1	Model 2	Model 3	Model 4
	(1)	(2)	(3)	(4)
Observations	57,377	57,377	57,377	57,377
MAE	0.222	0.222	0.222	0.222
RMSE	0.341	0.341	0.341	0.341
AIC	51,658	51,634	51,618	51,617
BIC	51,774	51,769	51,771	51,787
Log Likelihood	-25,816.011	-25,802.100	-25,792.176	-25,789.443

The results in Table 2 show that there is a positive (0.065) statistically significant relationship ($p < 0.01$) between the number of IRA tweets and the average sentiment of the NATO conversation. Russia's overt media activities also have positive relationship (0.051) to sentiment of the tweets across country-day, though with a lower level of statistical significance ($p < 0.05$). The edge density of the network of users is the third coefficient that is also positively related to sentiment of the tweets. Finally, the control variables are statistically significant, although on a different level. All of them have a p-value<0.01, except political regime with a p-value<0.1. All their coefficients except the country's access to the internet are negative.

Table 2. Sentiment Regression Models: Base Terms

	Sentiment			
	Model 1	Model 2	Model 3	Model 4
	(1)	(2)	(3)	(4)
NATO Exercise	-0.001	-0.152	-0.099	-0.365
	(0.038)	(0.137)	(0.137)	(0.252)
Russian Exercise	0.080	0.071	0.807***	0.807***
	(0.052)	(0.052)	(0.200)	(0.200)
IRA's Tweets	0.039***	0.052***	0.052***	0.065***
	(0.010)	(0.010)	(0.011)	(0.013)
Russian Media Tweets	0.062**	0.060**	0.060**	0.051**
	(0.025)	(0.025)	(0.025)	(0.026)
Distance to Russia	0.005	0.0003	0.004	0.004
	(0.003)	(0.004)	(0.004)	(0.004)
Edge Density	0.320***	0.326***	0.325***	0.325***
	(0.020)	(0.021)	(0.021)	(0.021)
Total Tweets	-0.107***	-0.107***	-0.107***	-0.107***
	(0.009)	(0.009)	(0.009)	(0.009)
Political Regime	-0.015***	-0.015***	-0.015***	-0.008*
	(0.002)	(0.002)	(0.002)	(0.004)
GDP per capita	-0.165***	-0.164***	-0.164***	-0.163***
	(0.022)	(0.022)	(0.022)	(0.022)
Population	-0.180***	-0.180***	-0.180***	-0.180***
	(0.014)	(0.014)	(0.014)	(0.014)
Access to Internet	0.173***	0.171***	0.171***	0.171***
	(0.026)	(0.027)	(0.027)	(0.027)
1| 2	-5.579***	-5.648***	-5.595***	-5.580***
	(0.174)	(0.175)	(0.175)	(0.176)
2| 3	0.110	0.043	0.100	0.115
	(0.170)	(0.171)	(0.172)	(0.173)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3 provides information about the multiplicative interaction terms included in Models 2, 3, and 4. The first two terms test Hypothesis 2, or whether there is a negative effect on sentiment arising from the combination of NATO exercises and the activity of Russia's overt and covert media actors. The first interaction term between *NATO Exercise* and *IRA Tweets* is negative (-0.106) and statistically significant (p < 0.01). The second term, which reflects the influence of Russia's overt media outlets on sentiment during NATO exercises, is not statistically significant. Therefore, it is impossible to draw firm conclusions concerning the direction and the strength of the relationship. The third term, which represents how the users' distance from the Russian border influences the sentiment of their tweets when a NATO exercise occurs, is positive and statistically significant, indicating that sentiment is stronger at longer distances. The fourth and fifth terms test the effects of *IRA Tweets* and *Distance to Russia* on sentiment during Russian military exercises rather than NATO exercises. Of these, only the interaction between *Distance to Russia* and *Russian Exercise* is statistically significant (p < 0.01), with a negative (-0.054) coefficient, implying that Russian exercises generate more negative effects on sentiment as the distance from Russia grows. Finally, the sixth term tests how IRA activities and the country's political regime type generate combined effects on sentiment. It is negative (-0.003) and statistically significant, although to a lesser degree (p < 0.1), implying that IRA activities generate more negative impacts on sentiments in democracies than in autocracies.

Table 3. Sentiment Regression Models: Multiplicative Interaction Terms

	Sentiment			
	Model 1	Model 2	Model 3	Model 4
	(1)	(2)	(3)	(4)
NATO Exercise*IRA Tweets		-0.105***	-0.105***	-0.104***
		(0.029)	(0.029)	(0.029)
NATO Exercise*Russian Media Tweets				0.108
				(0.086)
NATO Exercise*Distance to Russia		0.031***	0.027***	0.027***
		(0.008)	(0.008)	(0.008)
Russian Exercise*IRA Tweets			0.005	0.005
			(0.033)	(0.033)
Russian Exercise*Distance to Russia			-0.055***	-0.055***
			(0.012)	(0.012)
IRA Tweets *Political Regime				-0.003**
				(0.002)
Note:			*p<0.1; **p<0.05; ***p<0.01	

2. Analysis of the Findings

Hypothesis 1: *The overall sentiment of the NATO online conversation will become more negative if there is an increase in IRA and Russian media activity.*

The results from the regression models do not support Hypothesis 1. Figure 1, showing estimates derived from Model 4, demonstrates that increased IRA activity heightens the probability that the average sentiment will be positive (blue line) and lowers the probability for negative sentiment (red line) in the graph at the left. The second graph on the right shows that the that the tweets by overt Russian media have a similar effect on sentiment, though that effect is slightly weaker.

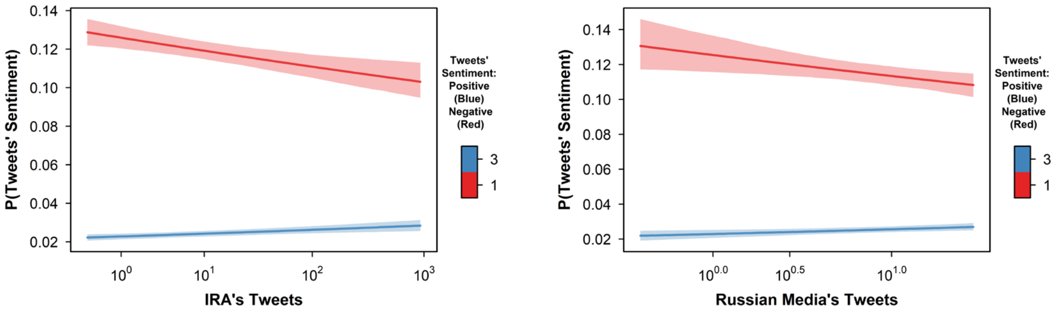

Figure 1. IRA and Russian Media Tweets vs. Sentiment in the NATO Conversation

The colored shading around the lines in the graphs shown in Figure 1 represent the 95% confidence interval of the respective variables and to what degree the slope of the lines in the graphs can vary within this interval. The comparison of the two graphs in Figure 1 reveals that the slopes of the lines are different. Both lines in the graph at left, depicting the IRA's tweets, have a steeper slope than those in the graph at right, depicting the effects of tweets by overt Russian media, which means that the IRA's activity has a more substantial effect on the sentiment than does the activity of overt Russian media.

Figure 2 further supports this conclusion by presenting the substantive effects of IRA and Russian media tweets on positive or negative sentiment probability. For convenience and better comparison, the graphic on the right in Figure 2 presents the absolute values of these effects. From Figure 2, it is easy to compare of the effect magnitudes, which are, in fact, negative (given the red lines on the graphs in Figure 1 have negative slopes).

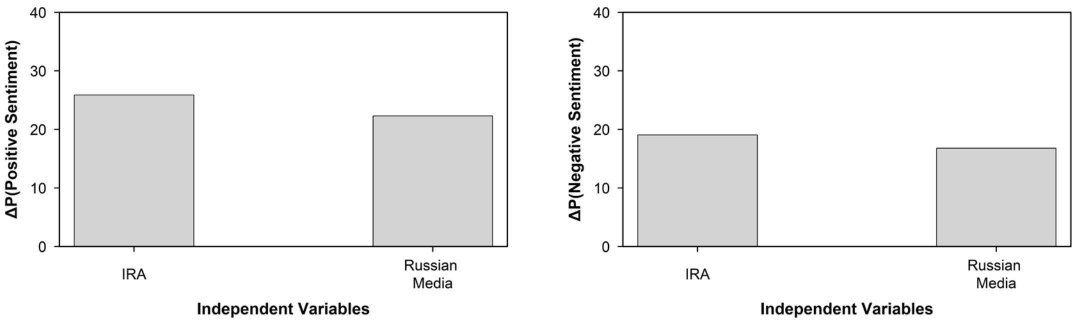

Figure 2. Substantive Effects of IRA and Russian Media Tweets on the Probability of Positive and Negative Sentiment

The graphics in Figure 2 demonstrate two important findings. First, the IRA's tweets have a stronger impact on the probability of sentiment change than tweets by the overt Russian media. Second, both the overt and the covert Russian actors more effectively generate positive sentiment and are counter-intuitively associated with decreases in negative sentiment.

Hypothesis 2: *The sentiment of the online NATO conversation during NATO exercises will become more negative with the increase of the activity of Russian media organizations, whether covert or overt.*

The regression models provide evidence that is partially supportive of Hypothesis 2. The first and second interaction terms in Table 3 provide the necessary information to examine how the Twitter activity of Russia's covert and overt media organizations affects the online conversation about NATO during NATO exercises. The first term *NATO Exercises*IRA Tweets* is statistically significant ($p < 0.01$) and negative (-0.106), implying that IRA activities generate more negative effects during NATO exercises. Figure 3 demonstrates this effect graphically. When NATO exercises take place, the increase in IRA's online activity results in a decreased probability for positive sentiment among network users. Conversely, when there is no NATO exercise, the probability for positive sentiment about NATO among network users increases when the number of the IRA tweets increases.

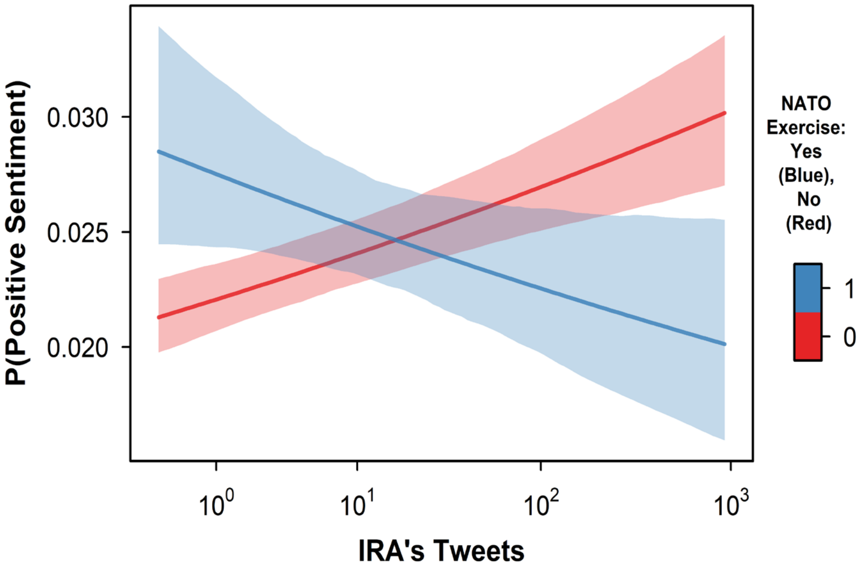

Figure 3. NATO Exercises and IRA Tweets vs. Probability of Positive Sentiment

In contrast, the term *NATO Exercise*Russian Media Tweets* is not statistically significant. This finding suggests that in the researched period, during NATO exercises, the Twitter accounts of overt Russian media probably were not as active in the online NATO conversation. This is contrary to the prediction of Hypothesis #2, which expected both overt and covert Russian media to have similarly negative effects during NATO exercises.

Hypothesis 3 *The change of sentiment during NATO exercises will be more negative if the country is closer to the borders of the Russian Federation.*

The regression results support this hypothesis. The multiplicative interaction term *NATO Exercise*Distance to Russia* is statistically significant ($p < 0.01$) and positive (0.027). Figure 4 shows that when NATO exercises take place, the probability that the average sentiment in the online conversation about NATO will be positive is lower in countries that are closer to the borders of Russia. In contrast, when there is no NATO exercise, the likelihood for positive sentiment decreases in countries that are further from Russia.

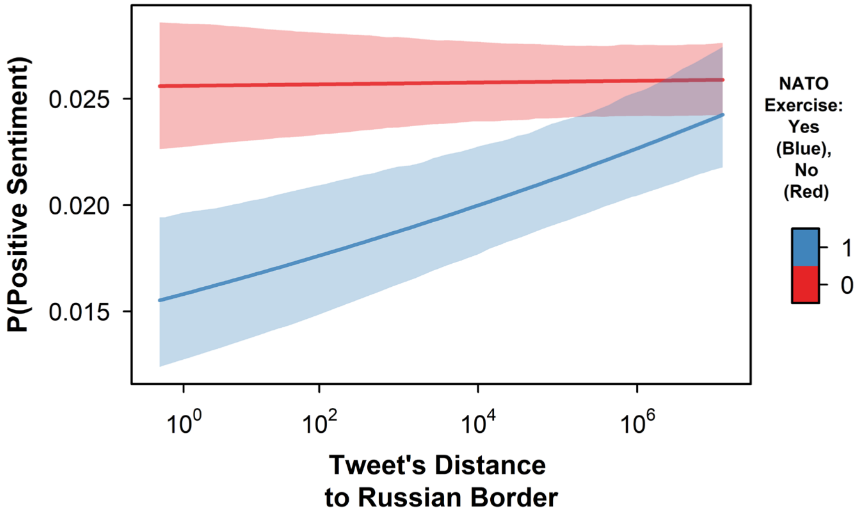

Figure 4. NATO Exercises and Distance to Russia vs. Probability of Positive Sentiment

B. NETWORK MEASURES AND RUSSIAN ONLINE INFLUENCE

Based on the models' performance scores in Table 4, it is possible to conclude that models 6, 8, and 10 have the best goodness of fit. They each have lower AIC and Bayesian Information Criteria (BIC) scores than the other three paired models presented in the table, indicating that the addition of the interaction terms decreases the models' errors.

The IRA and Russian media coefficients are statistically significant ($p < 0.01$) for all three topology measures. Those for clustering coefficient and network size are positive, for both overt and covert media activity, suggesting that these Russian entities' actions make the daily networks larger and more interconnected. The degree centralization coefficient for IRA activities is also statistically significant ($p < 0.01$). However, it is negative, which indicates that the daily network becomes less centralized with the increasing activity of Russia's covert media organizations.

The interaction term *NATO Exercise*IRA Tweets* is statistically significant ($p < 0.01$) for all three models. It is negative in the models for network size (-0.143) and clustering (-0.064), and it is positive in the case of degree centralization (0.027). The term

*NATO Exercise*Russian Media Tweets* is statistically significant and positive only in the degree centralization (p < 0.05) and clustering (p < 0.01) models.

Table 4. Network Regression Models

	Network Size		Degree Centralization		Clustering Coefficient	
	Linear	Linear	Linear	Linear	Linear	Linear
	(5)	(6)	(7)	(8)	(9)	(10)
IRA's Tweets	0.167***	0.179***	-0.049***	-0.051***	0.471***	0.471***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.007)	(0.007)
Russian Media Tweets	0.430***	0.349***	-0.005	0.011***	0.449***	0.274***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.018)	(0.019)
Distance to Russia	0.002***	0.002***	-0.001	-0.0004	0.004*	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Total Tweets	-0.009***	-0.008***	0.002	0.001	-0.018**	-0.015**
	(0.002)	(0.002)	(0.002)	(0.001)	(0.007)	(0.007)
Political Regime	-0.001	-0.0004	0.0002	0.0001	-0.001	-0.001
	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.002)	(0.002)
GDP per capita	-0.073***	-0.064***	0.018***	0.016***	-0.121***	-0.106***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.016)	(0.016)
Population	0.012***	0.010***	-0.002	-0.002	0.022**	0.019*
	(0.003)	(0.003)	(0.002)	(0.002)	(0.011)	(0.011)
Access to Internet	0.113***	0.099***	-0.026***	-0.023***	0.191***	0.167***
	(0.006)	(0.006)	(0.004)	(0.004)	(0.020)	(0.020)
NATO Exercise		0.984***		-0.287***		0.375**
		(0.049)		(0.037)		(0.172)
NATO Exercise*IRA Tweets		-0.143***		0.027***		-0.064***
		(0.006)		(0.005)		(0.022)
NATO Exercise*Russian Media Tweets		-0.018		0.033**		0.346***
		(0.019)		(0.014)		(0.065)
Constant	3.986***	4.052***	-3.124***	-3.136***	-9.010***	-8.784***
	(0.034)	(0.032)	(0.024)	(0.024)	(0.113)	(0.113)
Observations	57,377	57,377	57,377	57,377	57,377	57,377
MAE	139.775	127.017	0.017	0.017	0.006	0.006
RMSE	216.542	209.942	0.030	0.030	0.011	0.011
AIC	117,498	111,696	80,989	80,405	257,323	255,768
BIC	117,588	111,812	81,079	80,521	257,412	255,885
Log Likelihood	-58,739	-55,835	-40,485	-40,189	-128,651	-127,871
Note:						*p<0.1; **p<0.05; ***p<0.01

1. Analysis of the Findings

Hypothesis 4: *IRA and Russian media activity during NATO exercises will increase the size of the daily network of users in the NATO conversation, and it will become more centralized and interconnected.*

The results from these models are partially supportive of Hypothesis 4. Figure 5 demonstrates that an increase in the number of IRA tweets is associated with a larger size of the daily user network, which supports Hypothesis 4. The blue line in the graphic is less steep than the red, which means that the IRA's actions have a weaker effect on the network size when there is a NATO exercise than when there is none. The higher starting point of the blue line is expected because the dataset for this research is based on the online NATO conversation. Thus, when such military drills occur, more Twitter accounts are likely to become active and join the online NATO communication. The multiplicative interaction term between NATO exercises and Russian media tweets is not statistically significant; thus, it is impossible to determine if there is a consistent conditional relationship between these variables. However, the coefficient for *IRA Tweets* is statistically significant ($p < 0.01$) and positive (0.349). Thus, over the entire research period, both on exercise days and non-exercise days, increased levels of *Russian Media Tweets* were associated with a larger network size in the NATO conversation, which also supports Hypothesis 4.

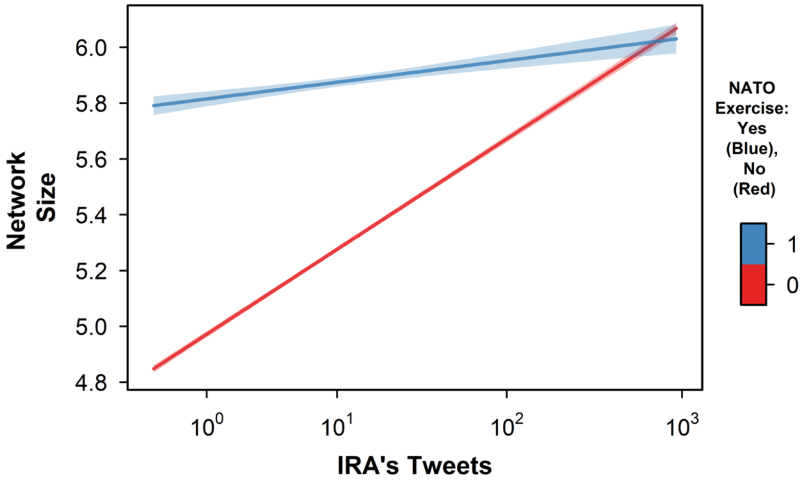

Figure 5. NATO Exercises and IRA Tweets vs. Network Size

In contrast, Model 9's results show that the online behavior of the IRA and Russia's overt media have different effects on the centralization of the daily networks. The results are partially supportive and partially contrary to Hypothesis 4. The blue lines on both graphics in Figure 6 lie below the red ones, which demonstrates that during NATO exercises, the centralization of the networks is lower than during periods without such military drills. However, the online behavior of the different types of Russian actors has an opposite effect on network centralization. When the IRA increases its online activity, centralization decreases. By contrast, when overt Russian media outlets are more active, centralization increases. Therefore, only findings about the effect of Russia's overt media organizations on network centralization support Hypothesis 4.

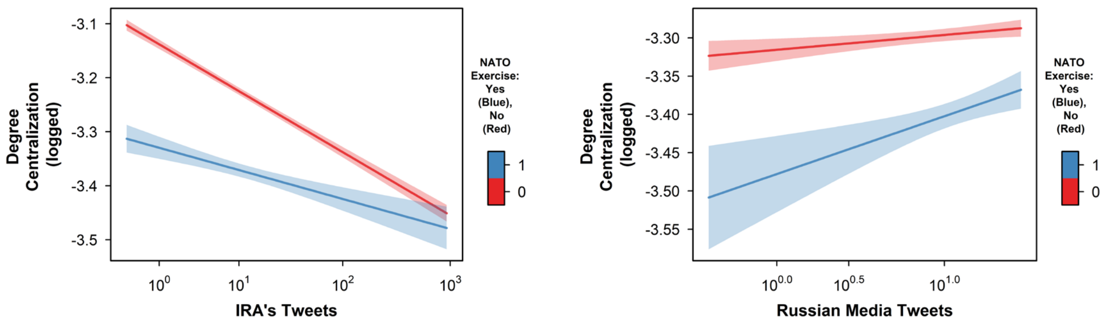

Figure 6. NATO Exercises, IRA Tweets, and Russian Media Tweets vs. Degree Centralization

The clustering coefficient is a measure of the network interconnectedness. It provides information on the degree to which the accounts in the explored daily networks have second-order ties. Figure 7 illustrates Model 10's results, which support Hypothesis 4. The clustering coefficient increases when Russia's covert and overt media organizations are more active. NATO exercises also lead to a further increase of interconnectedness in the daily networks.

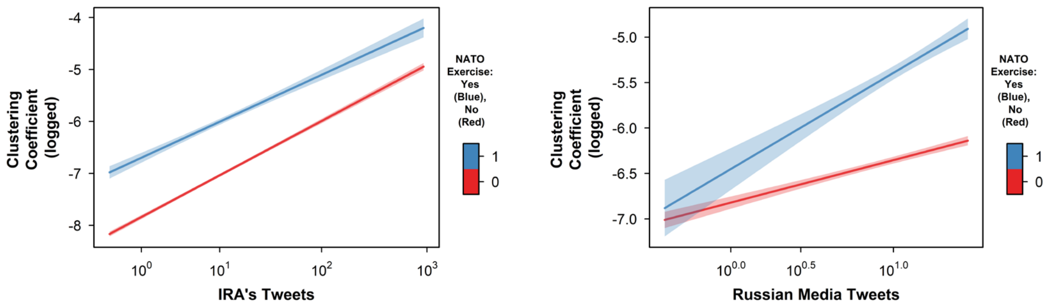

Figure 7. NATO Exercises, IRA Tweets, and Russian Media Tweets vs. Clustering Coefficient

C. DISCUSSION

In the examined period, July 2013 to August 2014, the regression analysis demonstrates that there is a statistically significant relationship between the sentiment of the online conversation about NATO and the online activities of Russia's overt and covert media organizations. At that time, the IRA was in its initial stage of formation. As Howard et al. point out, the agency used Twitter as a "training ground for the political polarization efforts."[111] The goal of IRA operatives then was to create "beachheads" in social networks by attracting followers, infiltrating online media platforms groups, and gaining credibility for its fake accounts. The prevalence of such preparatory actions could be one of the reasons why we observe the IRA's activity increasing the probability for positive sentiment in the research period, contrary to the expectations of Hypothesis 1.

Closer examination of the effects reveals that the influence potential of covert organizations is higher than that of overt media (Figure 2, Chapter V). This finding is logical because covert agencies can use fake identities or accounts to avoid attribution, infiltrate various online conversations, and implement tactics such as the "4D Approach" described by Nimmo.[112] At the same time, overt media outlets and the journalists who work for them cannot avoid this attribution, and the targeted audience can perhaps,

[111] Howard et al., "The IRA, Social Media and Political Polarization in the United States, 2012–2018," 10.

[112] Nimmo, "Anatomy of an Info-War: How Russia's Propaganda Machine Works, and How to Counter It."

therefore, more easily rebuff some of the information. This effect could explain why the regression model indicates that overt actors are weaker influencers than covert ones.

The findings about the effects of IRA and Russian media tweets during NATO exercises demonstrate that only the IRA's online activities have a statistically significant impact on sentiment. In this case, the model shows that the increase in the number of IRA tweets results in a lower probability for positive sentiment among network users. At the same time, the multiplicative interaction term combining NATO exercises with the online activity of Russia's overt media is not statistically significant. A possible interpretation of this regression result is that Russian media outlets did not target the NATO audience when the exercises took place. Another finding of the model supports this interpretation. The interaction term combining Russian exercises and IRA tweets is also not statistically significant. This result implies that the IRA's leadership probably did not make consistent use of its capabilities to manipulate the Twitter audience and influence their attitudes in favor of Russian military activities.

Although the model's results show that covert actors are more efficient in Twitter influence campaigns than overt media, the latter also have some advantages. First, they legally operate on social media platforms under their genuine identities. In this manner, the platforms cannot easily restrict their activities as long as the accounts adhere to the platforms' established rules and policies. In contrast, covert media agencies rely on fake accounts that act in violation of these regulations, and they could be banned or suspended at any time. As with all covert assets, their activity is efficient only as long as they operate hidden from their target. Second, official media companies, especially in the case of Russia, are quasi-private, and they can receive generous state funding to establish vast networks and create quality content. In return, they can be expected to coordinate and align their information policy with the ruling political elite.[113]

Both types of organizations can produce diverse online content adapted to specific social groups and geographical regions. The regression model indicates that there is a statistically significant relationship between the prevailing sentiment of tweets and their

[113] Hanley and Kuzichkin, *Russian Media Landscape*, 13–30.

authors' distance from the Russian border, conditioned by whether there is a NATO exercise or not (Figure 4, Chapter V). It demonstrates that in the vicinity of the Russian border, the probability for positive tweets is much lower when a NATO exercise takes place than when there is none. The farther from Russia the tweet originates, the greater the probabilities for positive sentiment—regardless of whether a NATO exercise is in progress. This result suggests that the Twitter users closer to the Russian border are more sensitive about NATO exercises. One of the possible reasons for this effect is that the dataset includes, by definition, a significant number of Russian accounts. In addition, as the discussion in Chapter III revealed, in the countries from the so-called 'near abroad' there are many communities that share Moscow's attitude towards NATO. These social groups are more susceptible to Russian "soft power" because of their common language, ethnicity, religion, or shared history.[114] The regression results also support this explanation. Figure 8 presents a situation where a Russian military exercise occurs, showing that the probability for positive sentiment is significantly higher near the border of Russia than when there is no exercise. In the absence of an exercise, the probability is almost constant as a function of distance to Russia.

[114] Rotaru, "Forced Attraction?," 3–9.

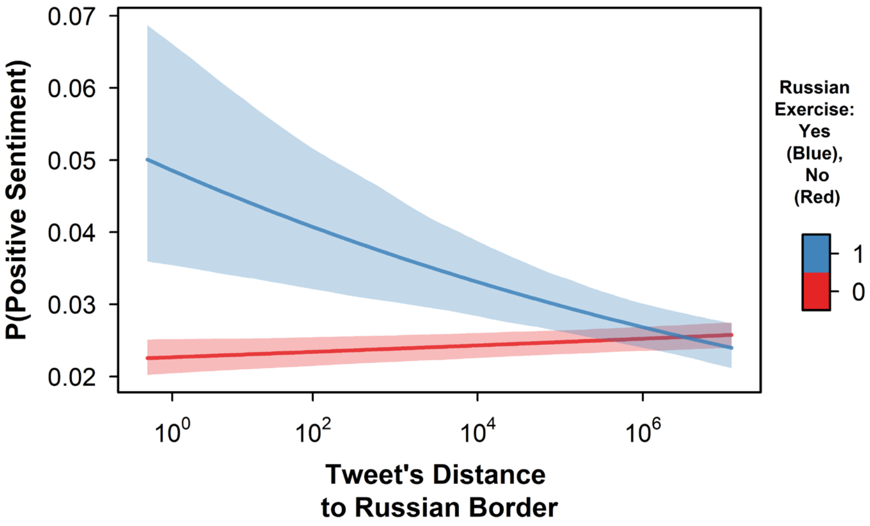

Figure 8. Russian Military Exercises and Distance to Russia vs. Probability of Positive Sentiment

The regression analysis of the daily networks of Twitter users also reveals statistically significant relations between the activity of Russia's overt and covert information agencies and network topology measures. This conclusion suggests that, during NATO exercises, these networks change their topology due to the actions of overt Russian media outlets and the covert operations of entities such as the IRA. In line with Hypothesis 4, the network becomes larger, and the clustering coefficient is higher when NATO exercises occur than in periods without military drills. Their values further increase if IRA or Russian media intensify their activity on Twitter. The finding on network enlargement can be explained by the expected appearance of new accounts engaging in the online conversation about NATO when that organization's exercises take place. The observed change in the clustering coefficient suggests that IRA and online media accounts also increase their second-order interactions. They start tweeting, mentioning, or quoting each other more frequently. This interconnectedness becomes even higher during NATO exercises. However, there is an important difference between the effect of the online activities of the IRA and the overt media outlets, as Figure 6 shows. During NATO exercises, IRA activities increase the network interconnectedness at the same rate as they do during periods without exercises. Howard et al. explain that "IRA accounts typically operated in ‘teams’ of co-mentioners …[that]… tended to mention teammates far more often than non-teammates; thus, forming a number of coherent communities of interaction.”[115] Similarly, Russian media organizations push the conversation to become more interconnected, but with a rate that increases substantially during NATO exercises. In other words, participants in the online conversation about NATO start retweeting, mentioning, or quoting each other more during NATO exercises.

Although the regression results for degree centralization partially support Hypothesis 4, they also reveal more specific patterns. Contrary to Hypothesis 4, the increase in IRA tweets results in a decrease in centralization. On the other hand, Russian media activity increases the centralization of the daily networks. These findings suggest that the increase in the Russian media’s online activity probably makes some nodes in the network more central. This conclusion corresponds to the qualitative findings in Chapter III that several central and larger online news agencies such as RT, Sputnik, and Ruptly appear to dominate the Russian media landscape. In contrast, covert entities such as the IRA seem to represent as a more “egalitarian” network composed of what Howard et al. referred to as small “communities of interaction” with a more similar number of ties.

[115] Howard et al., “The IRA, Social Media and Political Polarization in the United States, 2012–2018,” 26.

VI. CONCLUSIONS

This thesis has sought to examine the conversation about NATO on Twitter and to explore the relationships between the average sentiment in that conversation and the online activities of Russia's overt and covert online media agencies. In addition to the sentiments expressed in the tweets, the thesis analyzed how the basic topology measures of the social networks hosting this conversation relate to the online activity of overt Russian media and so-called troll factories. The thesis's research period was one year, from July 1, 2013, through August 31, 2014. The focus was on the conversation about NATO that appeared on Twitter and the activity on Twitter caried out by the major Russian media outlets and the Internet Research Agency during the NATO exercises that occurred in this period. The regression analysis of the online conversation about NATO in the specified period provided evidence of a statistically significant relationship between the sentiment of the tweets and the online actions of the Russia's covert and overt agencies. Several findings follow from the tested models' results.

First, the overall sentiment of the online conversation about NATO generally becomes more positive if there is an increase in online activity by the IRA and overt Russian media. Further, the IRA's tweets have a more substantial impact on the probability for sentiment change than do tweets by Russia's traditional media outlets. Both the overt and the covert Russian actors more effectively influence positive than negative sentiments.

Second, the prevailing sentiment of the online conversation about NATO during NATO exercises becomes more negative as the online activity of covert Russian media organizations increases. When a NATO exercise takes place, an increase in the IRA's online activity results in a decreased probability for positive sentiment. By contrast, when there is no NATO exercise, the probability for positive sentiment increases when the number of IRA messages increases.

Third, the sentiment expressed in tweets during NATO exercises is generally more negative if the location of the tweet's originator is closer to the borders of the Russian Federation.

Fourth, IRA and Russian media online activity during NATO exercises increases the size of the daily network of users participating in the conversation on NATO, and that network becomes less centralized and more interconnected. During NATO exercises, Russian media activities online tend to increase the centralization of the daily network. In comparison, the IRA seems to act as a more "egalitarian" network composed of small teams with a similar number of ties.

These conclusions suggest that close tracking and examination of the online activities of Russia's covert and overt media agencies can provide the necessary base for detecting ongoing information operations. However, their proper identification faces specific difficulties. First, today's social media platforms are rapidly increasing their number, making monitoring Russian online activities difficult and resource-consuming. Second, each social media platform has its own specific and often restrictive rules for sharing information about its users' behavior and how other individuals can extract and collect such data. As in the example of the Twitter API,[116] there is a limit on the number of tweets collected as well as on access to historical data.

Nonetheless, further refining of the analytical methods can overcome some of these limitations and deliver a more comprehensive outcome. These improvements should focus on several directions. First, the measurement of sentiment in tweets could include machine learning algorithms that can increase the precision of the final result. Second, the online activities of covert agencies similar to the IRA can be tracked across different online platforms by natural language processing algorithms. They have great potential to discover patterns in online communication that will enable the timely identification of trolls' accounts.

Finally, it is hoped that the results of this thesis contribute to our better understanding of Russian information operations. These results indicate that the activities of Russia's covert and overt media agencies can effectively influence online conversation and shape public perception. Although the thesis focuses on the narrow topic of NATO

[116] Twitter, "Twitter API Documentation," February 27, 2022, https://developer.twitter.com/en/docs/twitter-api.

exercises, its results could have implications for other significant events such as elections, referendums, or conflicts such as the February 2022 war between Russia and Ukraine. The results demonstrate that the online activities of companies such as the IRA need close monitoring. The thesis's findings show that the IRA's actions more substantially affect online conversation than do channels operated by Russia's overt media.

Moreover, the events during the current war between Ukraine and Russia suggest that the role of covertly operating troll networks can be even more significant in a time of crisis or conflict. The European Union imposed restrictions on RT and Sputnik media platforms and publicly labeled them as sources of disinformation in 2022.[117] As a result, major online platforms including Twitter, Facebook, and Instagram declared that they will comply with the sanctions and will prevent these Russian media networks from posting content.[118] Such restrictions will reduce the ability of these covert influencers to interact with targeted audiences and diminish their role in the information operations. In this situation, only networks covertly established by trolls can continue to operate and affect the information environment. Thus, preventive measures against information operations must include constant monitoring and timely counteraction of such networks as well.

[117] Foo Yun Chee, "EU Bans RT, Sputnik over Ukraine Disinformation," Reuters, March 2, 2022, sec. Europe, https://www.reuters.com/world/europe/eu-bans-rt-sputnik-banned-over-ukraine-disinformation-2022-03-02/.

[118] Elizabeth Culliford, "Twitter to Comply with EU Sanctions on Russian State Media," Reuters, March 2, 2022, sec. Technology, https://www.reuters.com/technology/twitter-comply-with-eu-sanctions-russian-state-media-2022-03-02/.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Arif, Ahmer, Leo Graiden Stewart, and Kate Starbird. "Acting the Part: Examining Information Operations Within #BlackLivesMatter Discourse." *Proceedings of the ACM on Human-Computer Interaction* 2, no. CSCW (November 2018): 1–27. https://doi.org/10.1145/3274289.

Awan, Imran. "Cyber-Extremism: Isis and the Power of Social Media." *Society* 54, no. 2 (April 2017): 138–49. https://doi.org/10.1007/s12115-017-0114-0.

Beauchamp-Mustafaga, Nathan, and Michael S. Chase. *Borrowing a Boat Out to Sea: The Chinese Military's Use of Social Media for Influence Operations*. Policy Papers. Washington, DC: Johns Hopkins University School of Advanced International Studies, 2019.

Bennett, W. Lance, and Steven Livingston. "The Disinformation Order: Disruptive Communication and the Decline of Democratic Institutions." *European Journal of Communication* 33, no. 2 (April 2018): 122–39. https://doi.org/10.1177/0267323118760317.

Bernatis, Vincent. "The Taliban and Twitter: Tactical Reporting and Strategic Messaging." *Perspectives on Terrorism* 8, no. 6 (2014): 25–35. http://www.jstor.org/stable/26297291.

Beskow, David M., and Kathleen M. Carley. *Bot-Hunter: A Tiered Approach to Detecting & Characterizing Automated Activity on Twitter*. Pittsburgh, PA: Carnegie Mellon University, 2018. https://www.researchgate.net/publication/326606376_Bot-hunter_A_Tiered_Approach_to_Detecting_Characterizing_Automated_Activity_on_Twitter.

———. "Its All in a Name: Detecting and Labeling Bots by Their Name." *Computational and Mathematical Organization Theory* 25, no. 1 (March 2019): 24–35. https://doi.org/10.1007/s10588-018-09290-1.

———. "Social Cybersecurity: An Emerging National Security Requirement." *Military Review*, March-April (2019). https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/MA-2019/Beskow-Carley-Social-Cyber.pdf.

Blei, David M., Andrew Ng, and Michael Jordan. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3 (January 2003): 993–1022.

Brzezinski, Jan J., and Nicholas Varangis. "The NATO-Russia Exercise Gap." *Atlantic Council* (blog), February 23, 2015. https://www.atlanticcouncil.org/blogs/natosource/the-nato-russia-exercise-gap/.

Carley, Kathleen M., Guido Cervone, Nitin Agarwal, and Huan Liu. "Social Cyber-Security." In *Social, Cultural, and Behavioral Modeling*, edited by Robert Thomson, Christopher Dancy, Ayaz Hyder, and Halil Bisgin, Vol. 10899. Lecture Notes in Computer Science. Cham, Switzerland: Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-93372-6_42.

Carley, Kathleen M., Jana Diesner, Jeffrey Reminga, and Maksim Tsvetovat. "Toward an Interoperable Dynamic Network Analysis Toolkit." *Decision Support Systems* 43, no. 4 (August 2007): 1324–47. https://doi.org/10.1016/j.dss.2006.04.003.

Center for Systemic Peace. "Polity5 Project, Political Regime Characteristics and Transitions, 1800–2018," April 23, 2020. https://www.systemicpeace.org/inscrdata.html.

Chan, Elvis M. "Fighting Bears and Trolls: An Analysis of Social Media Companies and U.S. Government Efforts to Combat Russian Influence Campaigns during the 2020 U.S. Elections." Master's thesis, Naval Postgraduate School, 2021. http://hdl.handle.net/10945/68309.

Charon, Paul, and Jean-Baptist Jeangène Vilmer. *Chinese Influence Operations: A Machiavellian Moment*. Paris, France: Ministry for the Armed Forces, Institute for Strategic Research, 2021. https://www.irsem.fr/report.html.

Chee, Foo Yun. "EU Bans RT, Sputnik over Ukraine Disinformation." Reuters, March 2, 2022, sec. Europe. https://www.reuters.com/world/europe/eu-bans-rt-sputnik-banned-over-ukraine-disinformation-2022-03-02/.

Csardi, Gabor, and Tamas Nepusz. "The Igraph Software Package for Complex Network Research." *InterJournal Complex Systems* (2006): 1695. https://igraph.org.

Culliford, Elizabeth. "Twitter to Comply with EU Sanctions on Russian State Media." Reuters, March 2, 2022, sec. Technology. https://www.reuters.com/technology/twitter-comply-with-eu-sanctions-russian-state-media-2022-03-02/.

Cunningham, Daniel. Understanding Dark Networks: A Strategic Framework for the Use of Social Network Analysis. Lanham, MD: Rowman & Littlefield, 2016.

Ferrara, Emilio. *Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election*. Los Angeles, CA: University of Southern California, Information Sciences Institute, 2017. https://ssrn.com/abstract=2995809.

Garrett, R. Kelly. "Echo Chambers Online?: Politically Motivated Selective Exposure among Internet News Users." *Journal of Computer-Mediated Communication* 14, no. 2 (January 2009): 265–85. https://doi.org/10.1111/j.1083-6101.2009.01440.x.

Gerasimov, Valery "The Science's Value is in the Prediction." *Military Industrial Courier* no. 8(476), 02/27-03/05/2013, https://vpk-news.ru/sites/default/files/pdf/VPK_08_476.pdf, accessed February 1, 2022

Golovchenko, Yevgeniy, Mareike Hartmann, and Rebecca Adler-Nissen. "State, Media and Civil Society in the Information Warfare over Ukraine: Citizen Curators of Digital Disinformation." *International Affairs* 94, no. 5 (September 1, 2018): 975–94. https://doi.org/10.1093/ia/iiy148.

Hanley, Monica, and Andrey Kuzichkin. *Russian Media Landscape:Structures, Mechanisms, and Technologies of Information Operations*. Riga, Latvia: NATO Strategic Communications Centre of Excellence, 2021. https://stratcomcoe.org/publications/russian-media-landscape-structures-mechanisms-and-technologies-of-information-operations/215.

Helmus, Todd C. *Russian Social Media Influence: Understanding Russian Propaganda in Eastern Europe*. Research Report (Rand Corporation), RR-2237-OSD. Santa Monica, CA: RAND Corporation, 2018. https://www.rand.org/pubs/research_reports/RR2237.html.

Howard, Philip N., and Samantha Bradshaw. "The Global Organization of Social Media Disinformation Campaigns." *The Journal of International Affairs, SIPA* 71 (September 2018). https://jia.sipa.columbia.edu/global-organization-social-media-disinformation-campaigns.

Howard, Philip N., Bharath Ganesh, Dimitra Liotsiou, John Kelly, and Camille François. "The IRA, Social Media and Political Polarization in the United States, 2012–2018." *Oxford, UK: University of Oxford, Computational Propaganda Research Project, 2019*, October 2019, 48. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1004&context=senatedocs.

Hu, Minqing, and Bing Liu. *Mining and Summarizing Customer Reviews*. Research Track Paper. Chicago, IL: University of Illinois at Chicago, 2004.

Kania, Elsa. "The PLA's Latest Strategic Thinking on the Three Warfares." *China Brief* 16, no. 13 (August 2016). https://jamestown.org/program/the-plas-latest-strategic-thinking-on-the-three-warfares/.

Lee, Kyumin, Brian David Eoff, and James Caverlee. "Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter." In *Fifth International AAAI Conference on Weblogs and Social Media*, 185–92. Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media. College Station, TX: Texas A&M University, 2011.

Linvill, Darren L., and Patrick L. Warren. "Troll Factories: Manufacturing Specialized Disinformation on Twitter." *Political Communication* 37, no. 4 (February 2020): 447–67. https://doi.org/10.1080/10584609.2020.1718257.

———. *Troll Factories: The Internet Research Agency and State-Sponsored Agenda Building*. Clemson, SC: Clemson University, 2018. http://pwarren.people.clemson.edu/Linvill_Warren_TrollFactory.pdf.

Mejias, Ulises A., and Nikolai E. Vokuev. "Disinformation and the Media: The Case of Russia and Ukraine." *Media, Culture & Society* 39, no. 7 (October 2017): 1027–42. https://doi.org/10.1177/0163443716686672.

Mihaylov, Todor, Georgi Georgiev, and Preslav Nakov. "Finding Opinion Manipulation Trolls in News Community Forums." In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*, 310–14. Beijing, China: Association for Computational Linguistics, 2015. https://doi.org/10.18653/v1/K15-1032.

Morales, James. "Assessing Anti-American Sentiment through Social Media Analysis." Master's thesis, Naval Postgraduate School, 2016. http://hdl.handle.net/10945/51587.

Morstatter, Fred, Jürgen Pfeffer, Huan Liu, and Kathleen M. Carley. "Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose." *ICWSM 2013*, June 2013. http://arxiv.org/abs/1306.5204.

Nekmat, Elmie. "Prosocial vs. Trolling Community on Facebook: A Comparative Study of Individual Group Communicative Behaviors." *International Journal of Communication* 12 (2018): 1–22. http://ijoc.org.

Nimmo, Ben. "Anatomy of an Info-War: How Russia's Propaganda Machine Works, and How to Counter It." *Central European Policy Institute* 15 (2015). https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/.

Panarin, I. N. *Informatsionnaia Voina i Geopolitika*. [Information Warfare and Geopolitics] Velikii Put′. Moskva: Pokolenie, 2006.

Pomerantsev, Peter, and Michael Weiss. *The Menace of Unreality:How the Kremlin Weaponizes Information, Culture and Money*. The Interpreter. New York, NY: The Institute of Modern Russia, 2014. https://www.almendron.com/tribuna/wp-content/uploads/2015/08/The_Menace_of_Unreality_Final.pdf.

Qi, SiHua, Lulwah AlKulaib, and David A. Broniatowski. "Detecting and Characterizing Bot-Like Behavior on Twitter." In *Social, Cultural, and Behavioral Modeling: 11th International Conference*, SBP-BRiMS 2018. Washington, DC, USA, July 10–13, 2018.

Qiao, Liang, and Xiangsui Wang. *Unrestricted Warfare*. 2nd ed. Wuhan, China: Chongwen, 2011.

Reynolds, Nathaniel. *Putin's Not-So-Secret Mercenaries: Patronage, Geopolitics, and the Wagner Group*. The Return of the Global Russia. Washington, DC: Carnegie Endowment for International Peace, 2019. https://carnegieendowment.org/files/GlobalRussia_NateReynolds_Vagner.pdf.

Rotaru, Vasile. "Forced Attraction?: How Russia Is Instrumentalizing Its Soft Power Sources in the 'Near Abroad.'" *Problems of Post-Communism* 65, no. 1 (January 2, 2018): 37–48. https://doi.org/10.1080/10758216.2016.1276400.

Selph, Gregory R., Michael H. Crain, and Andrew Anderson. "Measuring Sentiment Response to Collective Violence through Social Media." Master's thesis, Naval Postgraduate School, 2018. http://hdl.handle.net/10945/66275.

Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. *Fake News Detection on Social Media: A Data Mining Perspective*. Vol. 19, 2017. https://arxiv.org/pdf/1708.01967.pdf.

The Joint Chiefs of Staff. *Joint Publication 3-13: Information Operations*. JP 3-13. Washington, DC: Joint Chief of Staff, 2012. https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_13.pdf.

U.S. Congress. House of Representatives. *The Internet Research Agency's Handles as of June 2018*. Washington, DC: U.S. House of Representatives Permanent Select Committee on Intelligence, 2018. https://intelligence.house.gov/uploadedfiles/ira_handles_june_2018.pdf.

———. *Russia Investigative Task Force: Testimony before Permanent Select Committee on Intelligence*, 115th Cong. (2017) (statement of Sean J. Edgett Acting General Counsel, Twitter, Inc). Accessed February 4, 2022. https://docs.house.gov/meetings/IG/IG00/20171101/106558/HHRG-115-IG00-Wstate-EdgettS-20171101.pdf.

Twitter. "Twitter API Documentation," February 27, 2022. https://developer.twitter.com/en/docs/twitter-api.

Uyheng, Joshua, Thomas Magelinski, Ramon Villa-Cox, Christine Sowa, and Kathleen M. Carley. "Interoperable Pipelines for Social Cyber-Security: Assessing Twitter Information Operations during NATO Trident Juncture 2018." *Computational and Mathematical Organization Theory* 26, no. 4 (December 2020): 465–83. https://doi.org/10.1007/s10588-019-09298-1.

Weisburd, Andrew, Clint Watts, and Jim Berger. "Trolling for Trump: How Russia Is Trying to Destroy Our Democracy." War on the Rocks, November 6, 2016. https://warontherocks.com/2016/11/trolling-for-trump-how-russia-is-trying-to-destroy-our-democracy/.

World Bank. “World Development Indicator.” Accessed February 27, 2022. https://datatopics.worldbank.org/world-development-indicators/.

Zhou, Xinyi, Reza Zafarani, Kai Shu, and Huan Liu. “Fake News: Fundamental Theories, Detection Strategies and Challenges.” In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, 836–37. Melbourne, ,Australia: ACM, 2019. https://doi.org/10.1145/3289600.3291382.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California