Information Operations Analysis of NATO Trident Juncture Exercise 2018

Joshua Uyheng[1], Thomas Magelinski[1], Ramon Villa Cox[1], Christine Sowa[1], and Kathleen M. Carley[1]

[1] CASOS, Institute for Software Research
Carnegie Mellon University, Pittsburgh, PA 15213
{juyheng, tmagelin, rvillaco, csowa, carley}@andrew.cmu.edu

Abstract. Information operations have the potential to shift public opinion through the strategic deployment of automated bot accounts driving targeted messages. This paper examines the Twitter conversation surrounding NATO and the 2018 Trident Juncture Exercise to determine the extent to which such operations took place. Using a pipeline of textual mining, network analysis, and machine learning tools, this report assesses and characterizes the activity of automated accounts which may have driven public discourse about the exercise. Our results show that tweets surrounding the exercise featured a combination of promotional NATO messages, coverage of a frigate collision, discussions of world politics, and opportunistic marketing schemes. A noticeable number of bots were detected across all topics, with estimates ranging between 12-30% of users in our dataset. While some bot activity appears to be driven by Russian media accounts, their influence on the network is apparently minimal. Furthermore, influencer analysis indicated that official NATO accounts dominated the social network, except for a few bot-like users with high volumes of pro-Russian content. We discuss the implications of our findings in view of utilizing and developing frameworks for detecting and characterizing online information operations at various granularities.

Keywords: Information Operations, Social Network Analysis, Bot Detection.

1 The NATO Trident Juncture Exercise 2018

The Trident Juncture Exercise (TRJE) is an international military event symbolizing the joint commitment of NATO member-nations to uphold their shared mandate of protecting security in the region. From late October through November 2018, the TRJE brought together 50,000 military and civilian personnel from 31 NATO and partner countries in Norway, making it NATO's largest exercise in two decades.

ACKNOWLEDGMENTS. This material is based upon work supported by the Office of Naval Research Multidisciplinary University Research Initiative (MURI) under award number N00014-17-1-2675. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Office of Naval Research. Additionally, Thomas Magelinski was supported by an ARCS foundation scholarship.

With its significant show of military force, the TRJE attracted widespread discussion on mainstream and social media platforms. However, the same online attention may enable the spread of manipulative messaging surrounding NATO and the TRJE. One long-standing issue encompasses NATO's tense relationship with Russia, especially in view of the exercise's geographic proximity to the superpower's borders. Another possible trigger for negative messaging was the unforeseen crash of a Norwegian frigate, the Helge Ingstad, while returning from the exercise [1]. Both represent sensitive concerns vulnerable to adversarial information operations.

Information operations are an urgent concern in the context of modern society's deep embeddedness in online social networks. Prior work has shown, through various methods, how adversarial actors infiltrate online communities to manipulate public opinion toward coordinated objectives [2, 3]. This paper empirically analyzes information operations in the Twitter conversation surrounding NATO TRJE 2018, focusing on the use of automated bot accounts to drive targeted messages. Using a combination of text mining, machine learning, and network analysis tools, we characterize how bots attempted to influence the public narrative surrounding the exercise and assess their overall influence over the online conversation.

2 Data and Methods

To assess the Twitter conversation surrounding NATO TRJE 2018, we analyzed 236,809 tweets collected from October 22, 2018 to November 13, 2018. Data collection was primarily conducted using the Twitter APIs with hashtags #tridentjuncture, #nato, and their non-English variants. Each tweet came with metadata on its respective user account and related interactions (e.g., retweets, mentions).

A total of 81,555 unique users were represented in the dataset. For certain time-sensitive aspects of the analysis, tweets were divided into four periods: **phase 1**, which included all tweets before the official start of TRJE on October 25; **phase 2**, which spanned the main days of the exercise from October 25 to November 7; **phase 3**, which encompassed the crash of the Helge Ingstad on November 8 and the two days afterward; and **phase 4**, which accounted for all tweets past November 10.

As visualized in Figure 1, a series of interoperable tools was used to leverage textual, user, and interaction information for topic analysis, bot detection, role identification, location prediction, and influencer analysis. Latent Dirichlet allocation (LDA) [4] was performed to provide an initial characterization of the main topics of conversation and track their prevalence across the four time periods. Extensive text preprocessing was conducted to standardize contractions, remove URLs and stop-words, etc. The final number of topics was evaluated using the coherence score [5].

To assess bot activity, Bothunter was used to examine Twitter accounts in the dataset and output a probability of the user being a bot based on their profile features [6]. Role identification employed a neural network model trained on a large dataset of user descriptions and tweets to classify accounts as belonging to special actor classes such as news agencies, reporters, government offices, and celebrities. Location prediction was also conducted using a neural network model trained on a dataset of user

descriptions and known locations [7]. Finally, ORA enabled the analysis of multi-modal networks to identify influential users, as well as characterize the overall structure of the Twitter conversation [8].

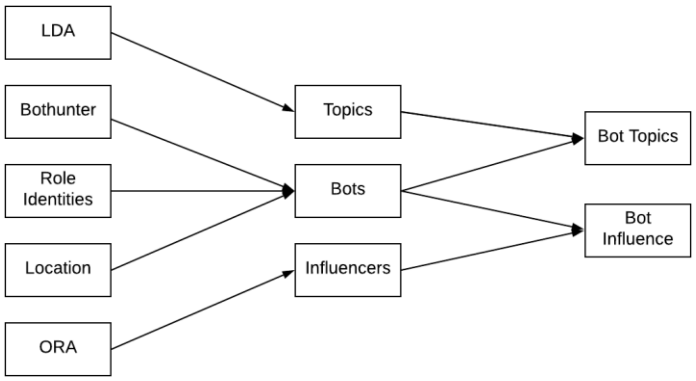

Fig. 1. Pipeline of tools to analyze topics, sentiment, bot activity, and influencers.

3 Results

Our findings characterize the Twitter conversation surrounding NATO TRJE 2018 as primarily dominated by official NATO accounts with potential information operations taking place using bot and bot-like accounts. In the succeeding sections, we present the topics of the online discussion, estimate the number of automated accounts involved, and describe their activities and impact on the social network.

3.1 Topic Analysis

Six topics were chosen for LDA based on coherence scores and manual assessment of the derived topics [5]. Multidimensional scaling suggested that the last three topics are relatively similar so they are interpreted concurrently in the succeeding analysis for convenience. Every tweet was assigned to a topic based on LDA topic probabilities. Figure 2 depicts the scaled distance between final topics used as well as their diffusion over time. Table 1 provides representative words and a sample tweet for each topic selected based on LDA results. We note that while the first two topics have coherent word lists, the latter two are relatively difficult to decipher on their own. This may be attributed to the fact that the topics themselves remain internally mixed.

Topic 1: NATO Trident Juncture. Accounting for about 70% of all tweets (including retweets), the first topic appears to concern general updates about the NATO TRJE. Messages in this topic are characteristically descriptive of exercise activities. As demonstrated by the sample tweet, the use of an official exercise hashtag #WeAreNATO signals solidarity with the efforts of NATO.

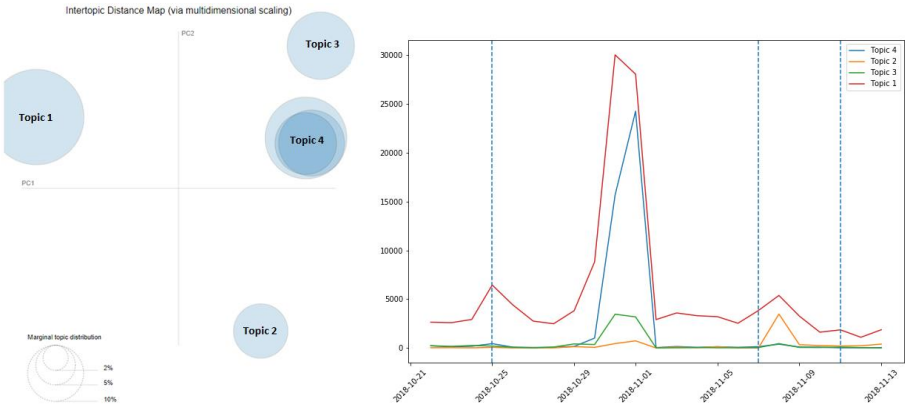

Fig. 2. Distance metrics and temporal diffusion of LDA topics. Left shows inter-topic distance map using multidimensional scaling. Right shows tweets per topic over time.

Topic 2: Collision of Helge Ingstad. Accounting for about 2-3% of all tweets (including retweets), the second topic aggregates tweets that discussed naval vessels. Importantly, it captures the tweets mentioning the Helge Ingstad collision. As shown in the diffusion plot, most of the content in this topic was produced during the period directly following the crash. In the sample tweet presented, the crash of the Helge Ingstad is used to discredit the strength of NATO.

Topic 3: World Politics. Comprised of tweets discussing general events around world politics, the third topic is not specific to the TRJE. It accounts for 6% of all tweets. Like the first topic, most of the conversation around this topic took place around the end of October and start of November, although there is also a spike at the start of the exercise. In the given example, NATO is mentioned as one of many entities linked to global conspiracies.

Topic 4: Opportunistic Marketing. The final topic identified is an aggregation of three different topics with a high degree of overlap in terms. It accounts for 37% of original tweets and 20% of all tweets (when considering retweets). Again not primarily related to the NATO exercise, most of the conversation took place around the end of October and start of November (similar to the other topics), however it does not seem to increase activity during the start nor the accident. Other tweets found in this topic appear to insert the #nato hashtag without actually talking about NATO. Hashtags are used opportunistically just to boost the visibility of the commodities being marketed.

Table 1. Sample tweets per topic.

Topics	Key Words	Sample Tweets
NATO Trident Juncture	nato, exercise, trident juncture, norway, Russia, military, maneuver, soldier, large, participate	JORSTADMOEN, Norway – A U.S. Army AH-64 Apache assigned to the 1st Battalion, 3rd Aviation Regiment. 12th Combat Aviation Brigade departs Rena Leir Airfield, Norway, during Exercise Trident Juncture, Nov. 5, 2018. #TridentJuncture2018 #WeAreNATO
Collision of Helge Ingstad	Russian, ship, photo, October, navy, frigate, sink, tanker, great, take	#NATO has to learn from his #TridentPUNCTURE. A fire on #Canadian frigate HMCS #Halifax. The storm nearly sank ship of #USNavy #GunstonHall. #CanadianNavy ship #Toronto lost its turn. A collision with tanker, Navy frigate #Norwegian KNM 'Helge Ingstad' Via @Syrian_Uruk.
World Politics	go, prepare, need, medium, thank, want, leave, ukraine, man, injure	There might be more truth here than you can handle…Just sayin' #USA #Syria #Palestine #MidEast #Yemen #Iraq #Libya #Afghanistan #Kushner #Zionism #Nazi #Wahhabism #NATO #NewWorldOrder #UN #AIPAC #GenieEnergy #Trump #Obama #Hillary #GeorgeBush #Obama #BillClinton
Opportunistic Marketing	Say, member, test, do, amp, defense, war, time, big	New on ebay: Arc Touch Wireless USB Receiver Mouse Slim Optical Flat Microsoft Touch Mouse KZ [URL]

3.2 Bot Analysis

Using a 60% threshold on Bothunter, we detected 24,868 bots in the dataset. This represents 30.49% of the unique users captured in our dataset. Cross-referenced against the role identity algorithm to remove special actors, only 10,072 bots remained, accounting for 12.35% of the unique users. Table 2 cross-tabulates the Bothunter results with role identity predictions. Percentages are relative to total users.

Table 2. Detected bots by role identity classifications.

Role Identity	Classified Users	Detected Bots
Normal users	38086 (46.70%)	**10072 (12.35%)**
Government	25325 (31.05%)	9411 (11.54%)
News agencies and reporters	15035 (18.44%)	4456 (5.46%)
Companies	1213 (1.49%)	308 (0.38%)
Celebrities	1139 (1.40%)	233 (0.29%)
Sports	757 (0.92%)	206 (0.25%)

The large reduction in final predictions is due to the significant proportion of reporters and government accounts classified as bot-like. Such Twitter users included the official NATO accounts which generated a high volume of content during the exercises. Figure 2 depicts the ORA visualization of bots on the social network of users connected by communication, with and without the role identity filtering. Filtered results are used as final predictions for conservative estimates of bot activity.

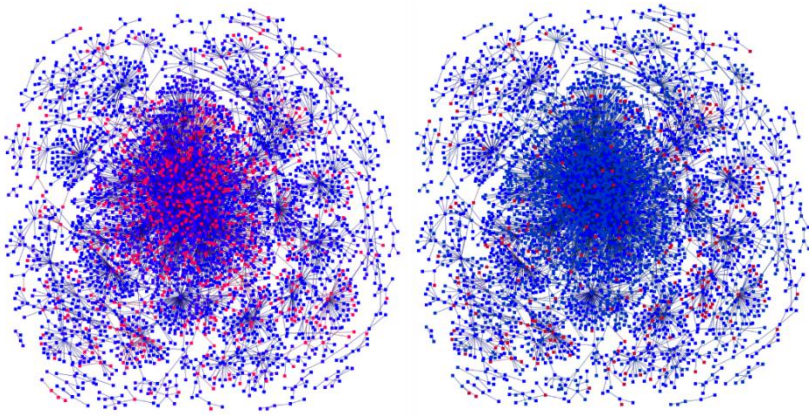

Fig. 3. Detected bots (red) in communication networks for NATO TRJE 2018. Left depicts bot predictions at 60% probability threshold. Right depicts predictions filtered by role identities.

Bot Locations. Cross-tabulated against our location prediction results, the top five countries with the highest numbers of detected bots included key NATO nations like the United States and Great Britain, with Russia also featuring a sizeable amount. Using filtered bot predictions, we report location results in Table 3.

Table 3. Detected bots by predicted geographic locations.

Location	Classified Users	Detected Bots
United States	20853	7956
Great Britain	7410	2716
Norway	5270	1865
Russia	6666	1293
Spain	7373	1203

Bot Conversations. Analyzed in conjunction with the previously identified topics, we found that the largest number of bot tweets talked about the exercises in general. As summarized in Table 4, bot messages comprised 25.63% of tweets in this topic, including both pro- and anti-NATO tweets. World politics tweets also featured content that was 20.30% from detected bot accounts. However, the largest proportion of bot

tweets per topic was found for the topic concerning the frigate collision. Thus, it appears that bots may have capitalized on the crash of the Helge Ingstad.

Table 4. Topics by proportion of tweets from detected bot accounts.

Topic	Bot Tweets
Collision of Helge Ingstad	2385 (31.97%)
NATO Trident Juncture	42512 (25.63%)
World Politics	3018 (20.30%)
Opportunistic Marketing	3799 (7.82%)

Sputnik Analysis. Following the suspicious activity geographically linked with Russia, we performed specialized analysis on the sub-networks associated with Sputnik accounts. Sputnik is a state-supported Russian media organization that has been previously linked to online propaganda [9]. The succeeding analysis drill down on Sputnik activity in the following manner. First, all Sputnik Twitter accounts in the dataset are enumerated. Second, all users with any form of communication (retweets, mentions) with these accounts are identified as first-order connections. Finally, users communicating with first-order connections are included. The social network composed of this subset of users (N = 6905) is then subjected to analysis.

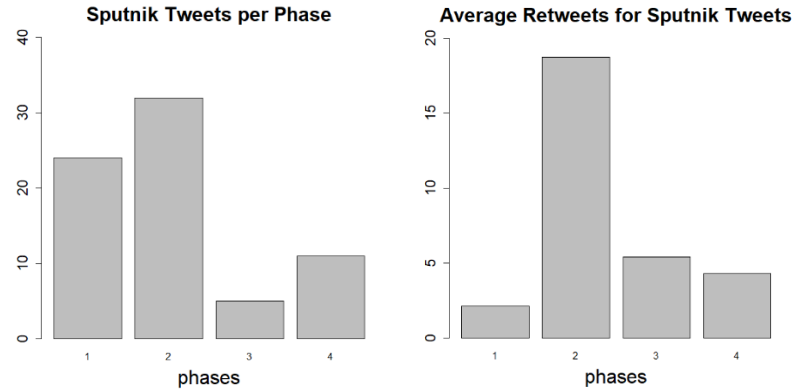

Fig. 3. Sputnik activity related to NATO TRJE 2018.

Our analysis showed that while Sputnik accounts tweeted about the exercises, they did not receive much traction in terms of retweets, as seen in Figure 3. In terms of topics, about 82.21% of the tweets in the Sputnik sub-network discussed the main NATO hashtags. Drilling down to the level of individual tweets, Sputnik stories included diverse types of anti-NATO messages and stories. Such stories revolved around themes of how TRJE aggravated local conflicts despite its ostensibly defensive objectives, triggering protests by peace activists around several Norwegian cities

against the “NATO war”. In one story, NATO troops had allegedly bruised and arrested an autistic man during drills. Stories had less than 10 likes and retweets each.

Detected bot activity, however, was significant in the Sputnik sub-network. Despite their relatively low influence on the Twitter conversation, about 41.00% of users in the Sputnik sub-network were classified as bots, noticeably higher than the detected 12.35% throughout the entire dataset. ORA visualization is given in Figure 4.

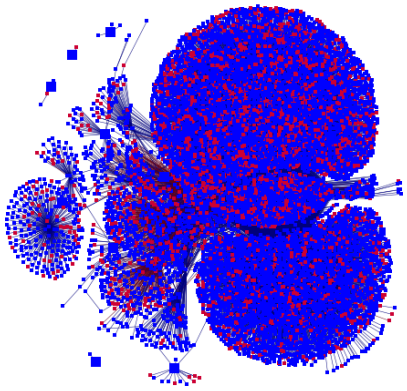

Fig. 4. Detected bots (red) in the Sputnik sub-network.

3.3 Influencer Analysis

Finally, we assess the impact of various agents on the overall Twitter conversation surrounding NATO TRJE 2018. One approach for impact assessment of information operations is to search for influencers and identify the messages they promote. Influencer analysis asks: Given a large-scale conversation on Twitter, which users impact the conversation the most? Under a network framework, measures of centrality are used to determine important users. For example, a user can be labeled an influencer if they have a high in-degree in their following network, since their tweets reach many people. Users can also reach many people if they are retweeted by someone with a large following. Additional measures include but are not limited to the total number of retweets, favorites, or mentions a user receives.

Combining meta-network measures of centrality, ORA finds three types of influencers: Super Spreaders, Super Friends, and Other Influencers. Super Spreaders are users which generate content that is shared often, and hence spread information effectively. Super Friends are users that exhibit frequent two-way communication, facilitating large or strong communication networks. Other influencers are users which have an active network presence, by tweeting often or mentioning users often, and operate in central parts of the conversation, such as by using important hashtags, or mentioning important users. In the succeeding analysis, usernames are withheld for anonymity due to the sensitivity of content; however, activities are characterized in detail to illustrate the type of agents detected by our tools.

Super Spreaders. Super Spreaders for each of the four periods were government-run or otherwise official NATO accounts, with one exception. While this is intuitive it points to something important about the conversation surrounding the exercise: NATO effectively controlled the conversation by pushing content about the exercise.

One exceptional user was classified as a Super Spreader after the crash of the Helge Ingstad. The account received a bot probability of 0.78. This score was likely obtained due to their extremely high Twitter activity (averaging 32 tweets per day for 5 years). Upon examination of their public profile, they appear to be a Russian patriot. In our dataset, the user calls the United States weak, citing that the NATO exercise could not be held without loss. Due to the coherence and awareness this tweet shows, it is unlikely to have been made by a bot. Further, the account exhibits intelligible direct replies to other users, in both English and Russian. From this analysis it appears that they may be a cyborg account, a politically engaged citizen, or a paid user promoting Pro-Russian content.

Super Friends. Super Friends again mostly found official accounts, showing that official NATO accounts also effectively pushed their message through two-way communication with other NATO accounts. Additionally, some Super Friends were normal users, as indicated both by Bothunter and the role identity tool. Again, there was one exception, this time during the exercise itself.

Like the Super Spreader described above, the suspicious Super Friend is incredibly active on Twitter (averaging more than 55 tweets a day for 10 years, with increasing frequency). They claim to live in Russia, tweet anti-USA content, and respond coherently to other users. In our dataset, this user claims that Trump is Putin's puppet, and that he ordered Trump to surround Russia with troops. With so many of the user's tweets being replies or shared articles with detailed commentary, this user is also likely a paid tweeter or cyborg account.

Other Influencers. Lastly, the Other Influencer category traditionally catches the most bot-like users, since accounts automatically tweeting often will achieve reasonably high scores. For instance, one user appears in this category across all four time periods in our dataset. This user is the most active user in the data set, with over 115 tweets per day on average for the last 10 years, with increasing frequency. Unlike other suspicious users, they claim to live in Sweden, with corroborating profiles on other websites linked in their Twitter account. Most of the tweets from the account are retweets, conveying pro-NATO/Anti-Russian content. Even though most contents are retweets, the links to other social media profiles indicate that this account is likely not a bot. Hence, it seems most likely that this is an active, politically charged user.

4 Discussion

Official NATO messages dominated the Twitter conversation surrounding the Trident Juncture Exercises in terms of quantity and influence. However, we also discovered

considerable bot activity, geographically concentrated around key NATO member nations and Russia. While our work does not in itself design new models or algorithms, we demonstrate the value of interoperable pipelines for engaging different facets of online information operations: analyzing actors, messages, and their influence over the public conversation.

Several challenges were encountered in the analysis we conducted. The multilingual dataset necessitated painstaking pre-processing procedures for textual analysis. Furthermore, the high-profile nature of NATO attracts opportunistic marketing schemes that use the relevant hashtags but do not actually discuss the event in question. While this is also an interesting finding, such factors must be taken into consideration as they may obscure more pertinent online information operations. Finally, while we examine social media traces for evidence of information operations, it is currently beyond our methodology to examine actual belief changes among the public outside the Twittersphere. Both our analytic insights as well as the limitations encountered highlight the importance of developing effective and adaptable frameworks for analyzing online information operations in context such as international security.

References

1. Forsberg, T., Herd, G.: Russia and NATO: From Windows of Opportunities to Closed Doors. Journal of Contemporary European Studies. 23, 41–57 (2015). https://doi.org/10.1080/14782804.2014.1001824.
2. Arif, A., Stewart, L.G., Starbird, K.: Acting the Part: Examining Information Operations Within #BlackLivesMatter Discourse. Proc. ACM Hum.-Comput. Interact. 2, 20:1–20:27 (2018). https://doi.org/10.1145/3274289.
3. Wilson, T., Zhou, K., Starbird, K.: Assembling Strategic Narratives: Information Operations As Collaborative Work Within an Online Community. Proc. ACM Hum.-Comput. Interact. 2, 183:1–183:26 (2018). https://doi.org/10.1145/3274452.
4. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet Allocation. Journal of Machine Learning Research. 3, 30 (2003).
5. Omar, M., On, B.-W., Lee, I., Choi, G.S.: LDA topics: Representation and evaluation. Journal of Information Science. 41, 662–675 (2015). https://doi.org/10.1177/0165551515587839.
6. Beskow, D.M., Carley, K.M.: Its all in a name: detecting and labeling bots by their name. Comput Math Organ Theory. (2018). https://doi.org/10.1007/s10588-018-09290-1.
7. Qian, Y., Tang, J., Yang, Z., Huang, B., Wei, W., Carley, K.M.: A Probabilistic Framework for Location Inference from Social Media. arXiv:1702.07281 [cs]. (2017).
8. Carley, K.M.: ORA: A Toolkit for Dynamic Network Analysis and Visualization. In: Alhajj, R. and Rokne, J. (eds.) Encyclopedia of Social Network Analysis and Mining. pp. 1219–1228. Springer New York, New York, NY (2014). https://doi.org/10.1007/978-1-4614-6170-8_309.
9. Aro, J.: The Cyberspace War: Propaganda and Trolling as Warfare Tools. European View. 15, 121–132 (2016). https://doi.org/10.1007/s12290-016-0395-5.