Network analysis of Russian Anti-War Discourse on Twitter during Russia's invasion of Ukraine

Iuliia Alieva[1[0000-0001-6270-8985]] and Kathleen M Carley[1[0000-0002-6356-0238]]

[1] Carnegie Mellon University, Pittsburgh PA 15213, USA
ialieva@cmu.edu

Abstract. This paper examines the dissemination and impact of Russian anti-war discourse on Twitter following Russia's invasion of Ukraine in 2022. This study aims to investigate the evolution of Russian anti-war discourse on Twitter, identify influential actors and communities, analyze the presence of bot accounts, and discern the predominant narratives propagated. The data collection and analysis employ a mixed-method pipeline, including data collection through Twitter API, network analysis using ORA software, bot detection using Bot-Hunter, and community detection using Leiden clustering. The findings provide insights into the dynamics of the Russian anti-war discourse on Twitter, shedding light on influential actors, narratives, and coordinated activities.

Keywords: Social Network Analysis, Disinformation, Russia, Ukraine, Computational Propaganda

1 Introduction

Prior to the onset of Russia's invasion of Ukraine in 2022, Russia was considered a hybrid regime, possessing both democratic and autocratic elements [1]. At that time, the relationship between democratic and authoritarian traits in the Russian political system was in a state of flux, with ruling elites adapting their approach based on the situation at hand. The media sector was particularly closely scrutinized, but in response to a demand for opposition voices, the Russian government maintained a mixed media system comprised of both state-controlled and independent outlets [2]. Nevertheless, following the anti-government protests in 2011-2012, Putin's regime encountered the so-called "dictator's dilemma" when endeavoring to implement censorship and regulate the Internet. The concept in question pertains to a predicament brought about by the emergence of novel forms of media that enhance public access to information, foster debate and mobilization, and pose challenges for the state [3]. Russia's invasion of Ukraine revealed the Russian government's belief that the presence of liberating media technologies represents a threat to the established regime, compelling a response from the ruling elites.

After the anti-government protests of 2011-2012, which utilized digital technology for mobilization and organization, the state responded by implementing harsh measures of oppression against the media and severely restricting freedom of speech. This situation escalated further after pro-Navalny protests in 2021, when opposition leader Alexei Navalny was promptly detained upon returning to Russia from Germany where

he had been receiving treatment following his poisoning the previous year. The recent war in Ukraine has exposed the purely authoritarian nature of Vladimir Putin's previously considered hybrid regime, which now appears to have zero tolerance for political criticism and dissenting views. As a result, independent media and non-profits have faced even stricter restrictions, with laws such as the 'foreign agents law,' 'undesirable organizations law,' and a ban on referring to a special military operation as a war being introduced. The government has also expanded its internal propaganda capacities and blocked social media access for Russian citizens. Since any form of protest, whether online or offline, could result in severe punishment such as imprisonment, many people are turning to online platforms to express their anti-war sentiments anonymously. Understanding the dissemination and impact of Russian anti-war discourse on social media is crucial for gaining insights into people's attitudes and their willingness to protest in the face of government oppression. To investigate the impact of the Russian anti-war discourse on Twitter, the study proposes several research questions:

RQ1: How has the spread of Russian anti-war discourse on Twitter evolved over time?

RQ2: Who are the most influential actors in the Russian anti-war discourse on Twitter?

RQ3: To what extent are bot accounts present in the Russian anti-war Twitter discourse?

RQ4: Which communities are the most influential in the Russian anti-war Twitter discourse?

RQ5: What are the predominant narratives that these actors and communities propagate?

RQ6: Which coordinated communities have the most influence in the Russian anti-war Twitter discourse?

RQ7: What narratives can be discerned from coordinated activities related to the Russian anti-war Twitter discourse?

2 Data and Method

The study's data collection and methodologies adhere closely to a pipeline established by social cyber-security studies [4-6]. To address our research inquiries, we collected a dataset of tweets using the Python package twarc through an archive search with the most recent version of the Twitter API, version 2.

For this study, the focus is on examining tweets in the Russian-language segment of Twitter to analyze the anti-war discourse related to the Russia's invasion of Ukraine. We collected tweets about anti-war discourse starting from February, 2022 till November, 2022. The keyword search includes the phrase "нет войне" and the hashtag #нетвойне ("no to war") in order to identify the anti-war tweets for our dataset. We converted the raw Twitter data into a meta-network consisting of user-to-user communication networks, user-to-tweet, and user to various tweet artifacts (hashtags, URLs) networks with ORA software to conduct network analysis [7].

A mixed-method pipeline was employed for data analysis, consisting of data collection, as well as bot detection using Bot-Hunter [8]. Bot-Hunter is a tool to identify

bot activities, a tiered supervised machine learning approach for bot detection and characterization. Our pipeline also involves Twitter data network analysis to detect key actors and influencers, including super spreaders and super friends, along with Leiden clustering to identify communities within the network. We then analyze the coordination between users and coordinated communities, followed by a qualitative analysis of the most influential agents in the network, their corresponding tweets, and narratives surrounding communities of agents (see Figure 1 for the overview).

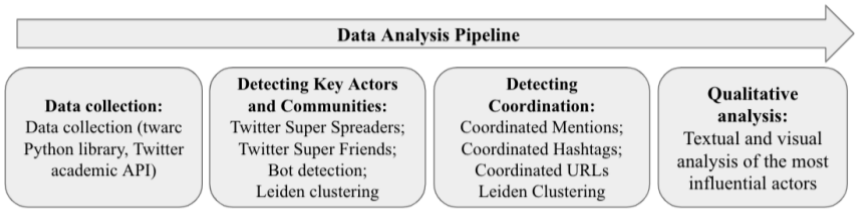

Fig. 1. The mixed-method pipeline used in the study.

ORA software is used to identify Twitter influencers and produces several metrics for Twitter data, such as the list of super spreaders (users that generate often shared content and hence spread information effectively) and super friends (users that exhibit frequent two-way communication, facilitating large or strong communication networks). The software uses several scores for computing super spreaders, including out-degree centrality, page rank centrality, and the membership of large k-core groups. ORA also computes scores for the list of super friends using total degree centrality and membership of large k-core groups.

To identify network communities participating in the conversations on Twitter, we use Leiden clustering method. Leiden clustering algorithm involves network partitioning and node movement that guarantees well- connected communities. Leiden algorithm was proved to be more efficient than others, such as Louvain; it is also faster and uncovers better partitions [9]. After identifying the communities, qualitative methods were used to compare content and user characteristics between groups.

Our research utilized a network-based approach to uncover narratives, communities, and coordinated activities. To detect these coordinated actions, we employed network analysis to identify connections between users who engaged in similar actions around the same time. To define coordination operationally, we employed the concept of synchronized action - a series of similar activities carried out within the same timeframe. These actions included tweeting with the same user mentions, hashtags, and URLs within a five-minute interval. The brief five-minute timeframe was particularly effective in identifying highly coordinated behavior since users who engaged in the same action repeatedly within this window were more likely to be deliberately coordinated rather than spontaneously so [10].

3 Results

As a result of data collection, our dataset contains 657,548 tweets with 497,431 retweets and 163,205 users (agents) (see Figures 2-4 for details on tweets over time and bot/not bot users).

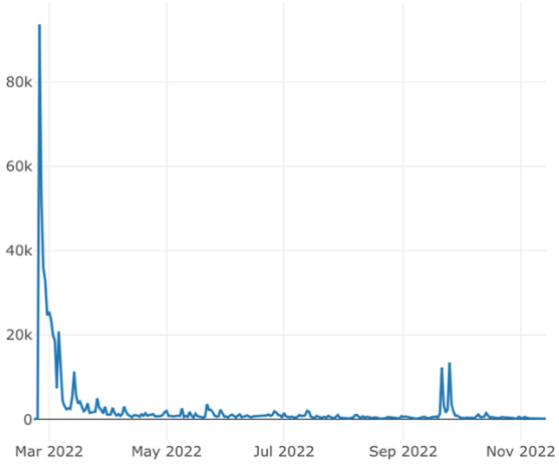

Fig. 2. Number of tweets over time.

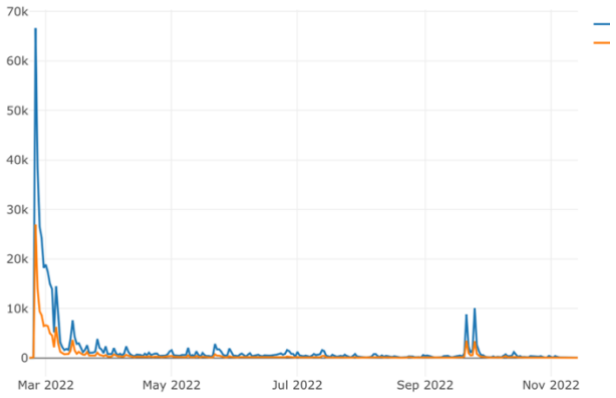

Fig. 3. Number of tweets from bots and not bots over time.

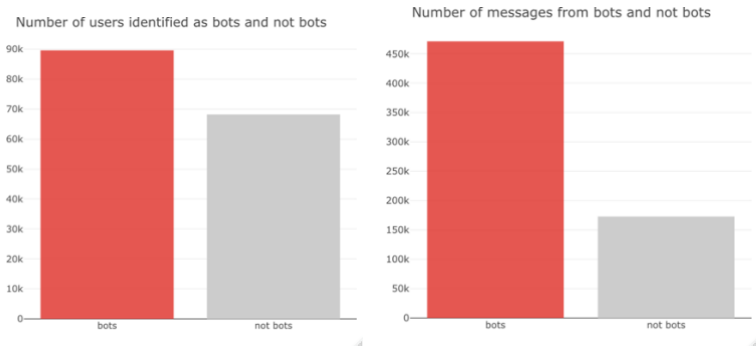

Fig. 4. Number of users identified as bots/not bots (left) and the number of tweets posted from bots/not bots (right).

3.1 **Twitter influencers**

In our dataset, various accounts and individuals could be identified as top super spreaders, including Lubov Sobol, the account of the French Ministry for Europe and Foreign Affairs in Russian, French politician Jean-Yves Le Drian, the human rights defense group OVD Info, news and opinion accounts such as @prof_preobr and @yoshkinkrot, the independent newsroom Media Zona, Russian opposition politician Alexey Navalny and his spokesperson Kira Yarmysh, their organization FBK Info, as well as the anti-war activist group Antivoenny bolnichny (@stranabolna).

Furthermore, among the top super friends accounts mentioning the "No war" slogan (нет войне) was ZavtraRu, which positions itself as a conservative and extreme-right newspaper in Russia with ultranationalist and anti-capitalist views. Additionally, another similar account, @DenTvRu, was identified as a super spreader. Den Tv is an extreme-right television media with a YouTube channel and website promoting Russian state propaganda and conspiracy theories with similar ultranationalist and anti-capitalist views.

Super friends list includes anti-war accounts such as @NoWar_Cats ("Котики против войны"), pro-Ukraine users, and pro-Navalny accounts.

After analyzing the hashtags and tweets spread by various accounts, differences in narratives were identified between the anti-war channels and Zavtra and Den TV. While anti-war channels used the hashtag #нетвойне (No war) in combination with other similar hashtags such as Stop Putin and No war with Ukraine, Zavtra and Den TV used the No war hashtag in combination with #Z and similar propaganda hashtags such as Yes to the victory (#ДаПобеде), Glory to Russia (#СлаваРоссии), and other state propaganda narratives. After further investigation, extensive hashtag hijacking was identified in both pro-war propaganda and anti-war groups. Hashtag hijacking is a form of cyber content attack in which a hashtag is used for a purpose other than its original intent such as labeling messages with undesirable content and promoting this content to a target audience [11]. While both groups were using hashtags from the opposite group for persuasion in order to promote their message.

3.2 **Leiden clusters for all communication and coordinated communication**

After applying Leiden clustering to the entire communication network, the first ten groups were examined. The four largest groups demonstrated authentic anti-war sentiments and comprised of opposition leaders and independent media such as Alexey Navalny, Lubov Sobol, Ilya Yashyn, MediaZona, and OVD Info, among others. However, groups #5 and #6 were predominantly composed of Russian state propaganda media and influencers, including mfa_russia, rian_ru, and dumagovru. Group #7 featured FBK and other opposition activists as top influencers, while group #8 consisted of pro-Ukraine users and Ukrainian politicians. Group #9 included Russian anti-war singers and performers as the main influencers, including Oxxxymiron, Anacondaz, and Danila Poperechny. Lastly, group #10 also had anti-war accounts as the primary influencers.

In the communication network analysis, Russian state propaganda and government accounts were found in both groups #5 and #6, but they were specifically identified as being in the first and second (largest) groups for coordinated communication. The third

and fourth groups, on the other hand, were made up of anti-war pro-Ukraine accounts. Groups #5 and #6 have accounts with anti-Putin and pro-Navalny narratives posting anti-war messages. Groups #7 to #9 were posting anti-war, pro-Ukraine messages, but cluster #10 has many accounts promoting Russian propaganda.

References

1. Petrov, N., Lipman, M., Hale, H. E.: Three dilemmas of hybrid regime governance: Russia from Putin to Putin. *Post-Soviet Affairs*, *30*(1), 1-26. (2014).
2. Denisova, A.: Democracy, protest and public sphere in Russia after the 2011–2012 anti-government protests: Digital media at stake. *Media, Culture & Society*, *39*(7), 976-994. (2017).
3. Shirky, C.: The political power of social media: Technology, the public sphere, and political change. *Foreign affairs*, 28-41. (2011).
4. Alieva, I., Moffitt, J. D., & Carley, K. M.: How disinformation operations against Russian opposition leader Alexei Navalny influence the international audience on Twitter. *Social network analysis and mining*, *12*(1), 80. (2022).
5. Alieva, I., Ng, L. H. X., & Carley, K. M.: Investigating the Spread of Russian Disinformation about Biolabs in Ukraine on Twitter Using Social Network Analysis. In *2022 IEEE International Conference on Big Data (Big Data)* (pp. 1770-1775). IEEE. (2022).
6. Uyheng, J., & Carley, K. M.: Characterizing bot networks on Twitter: An empirical analysis of contentious issues in the Asia-Pacific. *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, Springer, pp. 153-162. (2019).
7. Carley, K. M.: ORA: A Toolkit for Dynamic Network Analysis and Visualization. In Reda Alhajj and Jon Rokne (Eds.) Encyclopedia of Social Network Analysis and Mining, Springer. (2014).
8. Beskow, D. M., & Carley, K. M. (2018). Bot-hunter: a tiered approach to detecting & characterizing automated activity on Twitter. *SBP-BRiMS: International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, vol. 3.
9. Traag, V. A., Waltman, L., and Van Eck, N. J.: From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports*, 9(1):1–12. (2019).
10. Magelinski, T., Ng, L., & Carley, K.: A synchronized action framework for detection of coordination on social media. *Journal of Online Trust and Safety*, 1(2). (2022).
11. Xanthopoulos, P., Panagopoulos, O. P., Bakamitsos, G. A., & Freudmann, E.: Hashtag hijacking: What it is, why it happens and how to avoid it. *Journal of Digital & Social Media Marketing*, *3*(4), 353-362. (2016).