



**Selected Papers of #AoIR2023:
The 24th Annual Conference of the
Association of Internet Researchers**
Philadelphia, PA, USA / 18-21 Oct 2023

THE IMPACT OF TIKTOK POLICIES ON INFORMATION FLOWS DURING TIMES OF WAR: EVIDENCE OF ‘SPLINTERNET’ AND ‘SHADOW-PROMOTION’ IN RUSSIA

Salvatore Romano
AI Forensics

Natalie Kerby
AI Forensics

Miazia Schüler
AI Forensics

Davide Beraldo
University of Amsterdam

Illir Rama
University of Milano

Introduction

This paper measures the effects of platform policies in times of war, and represents different entry-points into understanding how platforms impact information flows (Ciuriak 2022). We present the outcomes of three tests that assessed content accessibility and prioritization on TikTok in Russia in the aftermath of the invasion of Ukraine. The tests adopted a combination of scraping and sock-puppet algorithmic audits (Sandvig 2014) conducted through Russian-based IP addresses.

This study sheds light on how TikTok adapted to the Russian war censorship laws with new policies (TikTok 2022), highlighting the need to closely monitor the platform's policy decisions during times of conflict. Using Digital Methods (Rogers 2019) and a cross-national approach (Gray 2021), we reveal the existence of opaque content removal strategies, unevenly implemented policies, and their consequences for Russian users, including the spread of pro-war propaganda and their isolation from the global TikTok community. Our findings contribute to the understanding of how TikTok policies can

contribute to the creation of a 'Splinternet' - a scenario in which countries have vastly different information online and separate internet infrastructures (Lemley 2020).

We quantify the changes in the information regarding the war recommended by TikTok to Russian users and demonstrate how these changes may have contributed to modifying the information diet in the country. Our study also offers insights into the methodological challenges of quantifying the effects of Russian techniques that international companies navigate when operating in the country (Vendil Pallin 2016).

Test 1: Content Restriction

Our first probe demonstrates how, after the invasion, TikTok restricted access to non-Russian content in Russia. We started monitoring TikTok's response to the Russian invasion with the initial aim of comparing the content that was prioritized to Ukrainian versus Russian TikTok users. However, while submitting a query for 'ukraine war' with a Russian IP address, we encountered issues with our scrapers. We further investigated these issues by conducting tests with a set of automated browsers, querying for a number of war-related keywords that were accessible under a Russian residential IP address network.

The results of this audit varied over time. At first, the search function on the platform was unusable for non-Cyrillic alphabet queries, returning no results. After a few hours, the search function was restored for non-Cyrillic alphabet queries. However, we observed that the platform had restricted content to Russian-only in both the search results and the For You Page (FYP), resulting in a 95% reduction of available content in Russia. As a result, accounts created with non-Russian IPs returned an empty page when accessed from a Russian IP. We then compared popular accounts accessed through a Russian IP address to those created through European IP addresses to understand the extent of the content limitations in Russia.

Overall, we found that TikTok was not complying with their own publicly announced policies regarding bans on new content in Russia (TikTok 2022). Instead, they made all foreign content unavailable in Russia, thereby isolating Russian users from the rest of the world.

Test 2: Pro-War Content

Further investigations showed how TikTok unevenly applied their content policies, allowing pro-war content to proliferate in Russia despite their claim that they had enforced a ban on new content uploads in Russia (Gilbert 2022). To assess whether and how new content could be uploaded and made available from Russia, we attempted to upload content from a number of accounts and IP addresses—with some profiles created in Russia and others not. We repeated these experiments dozens of times over twenty days, which allowed us to observe changes and inconsistencies in content upload permissions.

Based on the sample of videos we collected, and after an hashtag snowball process, we identified the 12 most prevalent hashtags that were commonly used by both sides to express their opinions on the war.

To analyze the content, we automatically accessed around 1,000 of the most viewed videos for each selected hashtag and determined which were accessible from a Russian IP address. Through our analysis, we were able to identify temporal trends in the prominence of pro-war and anti-war content (Figure 1).

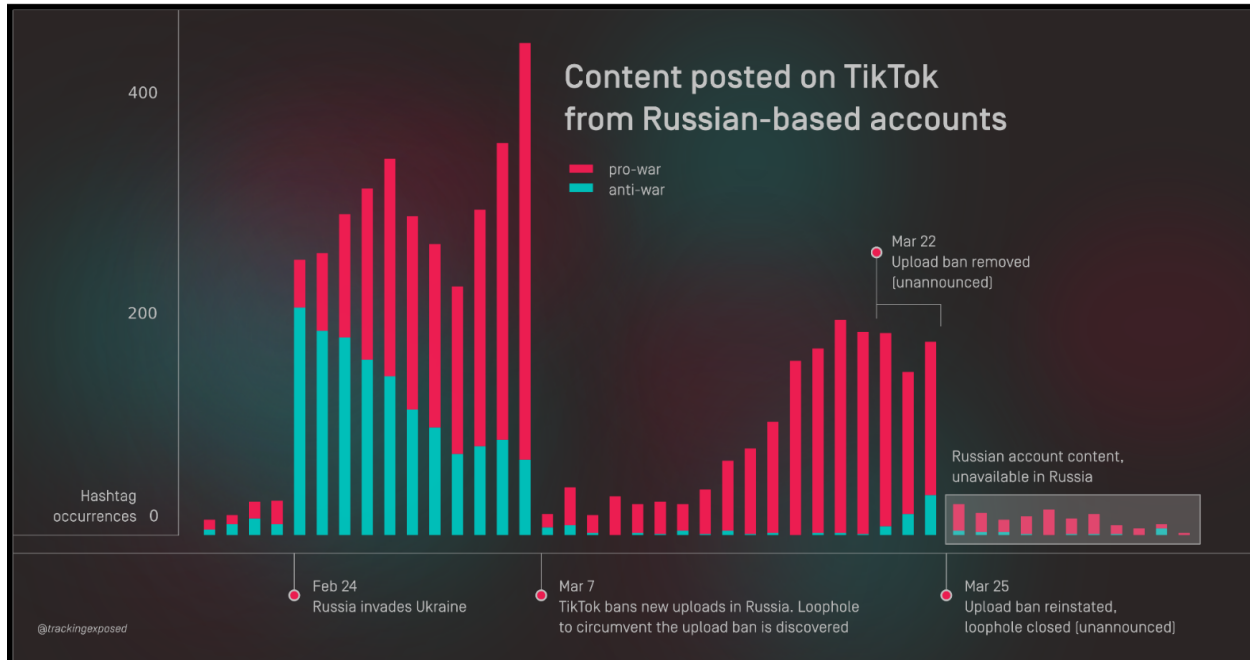


Figure 1: Content posted on TikTok from Russian-based accounts

Despite TikTok’s ban announcement on 6 March, we found that the company inconsistently applied the ban until 25 March. During the period in which TikTok did not comprehensively implement the ban, new content uploads related to the war were overwhelmingly pro-war (93.5%), while before the ban was announced, the balance of pro-war and anti-war content was roughly equal. Shortly thereafter, new uploads in Russia became impossible, and content from outside Russia was banned. Therefore, the only information about the Ukraine-Russia war available on TikTok in Russia was the content that was uploaded before 25 March, which was dominated by pro-war content and propaganda.

Test 3: Shadow-promotion

The third test brought to light a case of “shadow-promotion,” a term we coined to describe the algorithmic promotion of content that is supposed to be banned. According to TikTok’s own policy updated on 30 June, new content and international content was not available to domestic users inside Russia. However, we found that state controlled media, political, and non-political content were being promoted to Russian-based users via the algorithmically-compiled FYP, even though the videos did not appear on the profiles of the users who posted them.

We created a Russian TikTok profile and tried to upload a video while using a Russian IP and non-Russian IP. We then tried to access the video from different locations (Russia and Italy) using different access points, for example, accessing the uploader's profile through a direct link, or querying for keywords and/or hashtags.

Subsequently, we explored the availability of content supposedly banned for Russian users. We followed accounts with supposedly banned content, and checked the FYP afterwards to see if content from the account we had just followed was algorithmically recommended. We tested three account types: 1. A Russian media account, e.g.. Sputnikvideo, which was labeled by TikTok as 'Russian state-controlled media'; 2. A non-Russian media account, e.g. the official TikTok profile of BBC; 3. A non-political account, e.g. the official TikTok account of FC Dynamo, the Ukrainian professional football club based in Kyiv. For all the accounts, we followed the same protocol. We first checked their accessibility in Russia, in an EU-country (Italy), and in a non-EU country (U.S.) through non-logged browsers. We found that all the supposedly banned content, despite not accessible on the users' profile, did appear in the FYP for Russian users.

Conclusion

Ultimately, these three tests underscore the importance of monitoring TikTok's policy decisions in times of war, which we showcase through geographic-aware algorithmic audit methodologies. We were able to demonstrate the existence of unevenly implemented policies, and how it risked proliferating propaganda and perpetuating a Splinternet.

The Digital Service Act requires big online platforms to publish transparent and detailed reports on content moderation at least once a year (art. 13). Although TikTok's operations in Russia are outside the scope of EU legislation, as standards emerge for platform transparency, its lack of transparency in a high-profile region raises questions.

Our research emphasizes the importance of data transparency regarding TikTok's recommendation system and providing civil society access to data to hold the platform accountable for its algorithms and content moderation strategies. We have shown how this data can help identify biases, discriminatory practices, and manipulative algorithms. Further, our investigation stresses the need for social media platforms to have more transparent and consistent policies, particularly during times of conflict, to prevent the spread of harmful content and the fragmentation of the internet.

References

Ciuriak, Dan. 2022. "The Role of Social Media in Russia's War on Ukraine." Available at SSRN: <https://ssrn.com/abstract=4078863>.

Gilbert, David. 2022. "Russian TikTok Influencers Are Being Paid to Spread Kremlin Propaganda." Vice News (03/11/2022). <https://www.vice.com/en/article/epxken/russian-tiktok-influencers-paid-propaganda>

Gray, J. E. (2021). The geopolitics of "platforms": The TikTok challenge. *Internet policy review*, 10(2), 1-26.

Lemley, Mark A. 2020. "The Splinternet." Stanford Law and Economics Olin Working Paper #555. <http://dx.doi.org/10.2139/ssrn.3664027>.

Rogers, Richard. 2019. *Doing Digital Methods*. 1st edition. Thousand Oaks, CA: SAGE Publications.

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). Auditing algorithms: Research methods for detecting discrimination on internet platforms. *Data and discrimination: converting critical concerns into productive inquiry*, 22(2014), 4349-4357.

TikTok. 2022. "Bringing More Context to Content on TikTok." TikTok Newsroom. https://newsroom.tiktok.com/en-us/bringing-more-context-to-content-on-tiktok?utm_source=COMMSTWITTER&utm_medium=social&utm_campaign=030622 (accessed 23.02.2022).

Vendil Pallin, Carolina. "Internet Control through Ownership: The Case of Russia." *Post-Soviet Affairs*, vol. 33, no. 1, Jan. 2017, pp. 16–33. Taylor and Francis+NEJM, <https://doi.org/10.1080/1060586X.2015.1121712>.