

Reviewing War: Unconventional User Reviews as a Side Channel to Circumvent Information Controls

José Miguel Moreno¹, Sergio Pastrana¹, Jens Helge Reelfs², Pelayo Vallina^{1,3}, Savvas Zannettou⁴, Andriy Panchenko², Georgios Smaragdakis⁴, Oliver Hohlfeld^{2,5}, Narseo Vallina-Rodriguez³, Juan Tapiador¹

¹ Universidad Carlos III de Madrid, ² Brandenburg University of Technology (BTU),
³ IMDEA Networks Institute, ⁴ Delft University of Technology, ⁵ University of Kassel

Introduction



At the start of Russia's full-scale invasion of Ukraine on February 24, 2022, Roskomnadzor blocked social media platforms such as Facebook and Twitter in an attempt to control the flow of information within Russia. In response, users turned to **unconventional channels** to bypass this censorship. An account presumably affiliated with the Anonymous movement suggested using Google Maps reviews to disseminate war-related information to Russian citizens [1].



Go to Google Maps. Go to Russia. Find a restaurant or business and write a review. When you write the review explain what is happening in Ukraine. Idea via @Konrad03249040
5:41 PM · Feb 28, 2022

Anonymous @YourAnonNews · Feb 28, 2022
Leave 5 star reviews (unless its Russian State owned, then feel free to leave 1 star ratings). The point is to push information to the Russian civilian population being lied to by Putin.

Figure 1. Thread asking to review bomb Google Maps.

Motivated by the mention of this phenomenon in the media, we analyze how two popular travel platforms (**Google Maps** and **Tripadvisor**) were used to disseminate war-related content.

Methodology



We collected **7,330 posts** from 1.3k threads in Tripadvisor travel forums over 7 months of 2022. Using a custom Chromium-based crawler, we harvested **2.2 million** Google Maps reviews from 122,826 places in Russia between March 4, 2022, and June 30, 2022.

We manually annotated these messages as **war** or **non-war** and extracted their most frequent keywords. We classified these keywords into **war-like**, **violence-like**, and **other** keywords appearing in war-related posts, giving each a score. We automatically labeled the rest of the dataset using a keyword-matching approach.

References



[1] @YourAnonNews. 2022. <https://twitter.com/YourAnonNews/status/1498337491056836610>

This research was partially supported by the Spanish AEI grants CYCAD (PID2022-140126OB-I00) and CIARRES (TED2021-132170A-I00), the EU's Horizon 2020 R&I Programme grant TRUST aWARE (101021377), and by the European Research Council (ERC) under Starting Grant ResolutioNet (ERC-StG-679158). José Miguel Moreno is supported by the Spanish Ministry of Science and Innovation with a FPI Predoctoral Grant (PRE2020-094224). Narseo Vallina-Rodriguez has been appointed as a 2020 Ramon y Cajal Fellow (RYC2020-030316-I).

RQ1: Prevalence of War-related Content

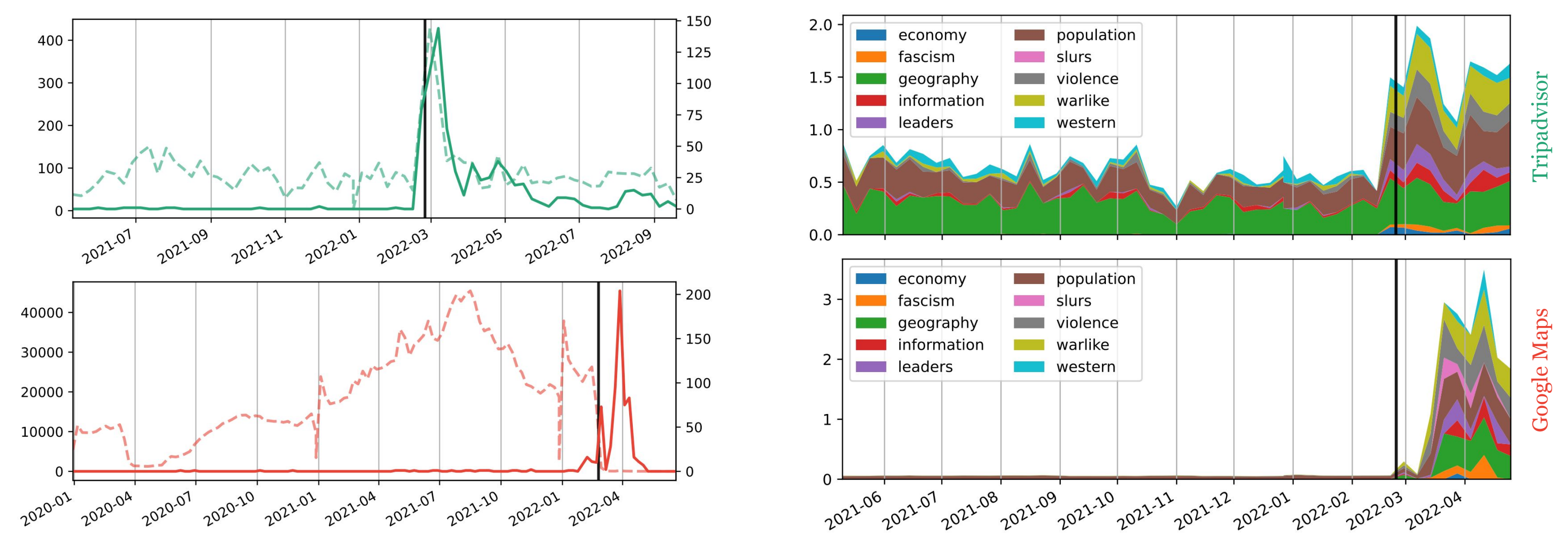


Figure 2. Weekly published posts and reviews. Dashed lines are non-war-related.

Figure 3. Ratio of topics found in posts and reviews per week.

We observe substantial changes in the number of daily posts and reviews since the beginning of the war that correlate with the blocking of major social platforms and news sites in Russia, and with a call by activist groups. Our automatic content labeling reveals that the traffic increase is **mostly due to war-related content**, and topic analysis confirms a noticeable change in user discourse. These findings ratify that Tripadvisor and Google Maps were used as side channels to disseminate war-related information to Russian citizens.

RQ2: Intent and Purpose of Content



Tripadvisor and Google Maps were used to disseminate messages about **(dis)information and censorship bypass, humanitarian help, hate speech, and travel advice**.

Here's a reminder of how to get around restrictions on BBC services in Russia: Download the Psiphon app [...]

A friend is involved in getting orphaned kids out of Dnipro [...] The charity, Edinburgh based Dnipro Kids' is hoping to get children to safety

*Russians are fascists
Moscow already knows that #!@% will strike, they are already digging trenches in Shebekino!!!!*

Tomorrow will be a bus (for free) in Lviv. Bus will be waiting for children and women next to main train station in Lwiw.

RQ3: Platform Moderation



There is evidence of both platform operators **actively removing messages that contain war-related content**. The most common reasons for content removal include off-topic message, harassment and hate speech, or more generally "content that violates ToS." The lifespan of moderated content ranges from a few hours to several weeks, with some messages escaping moderation.

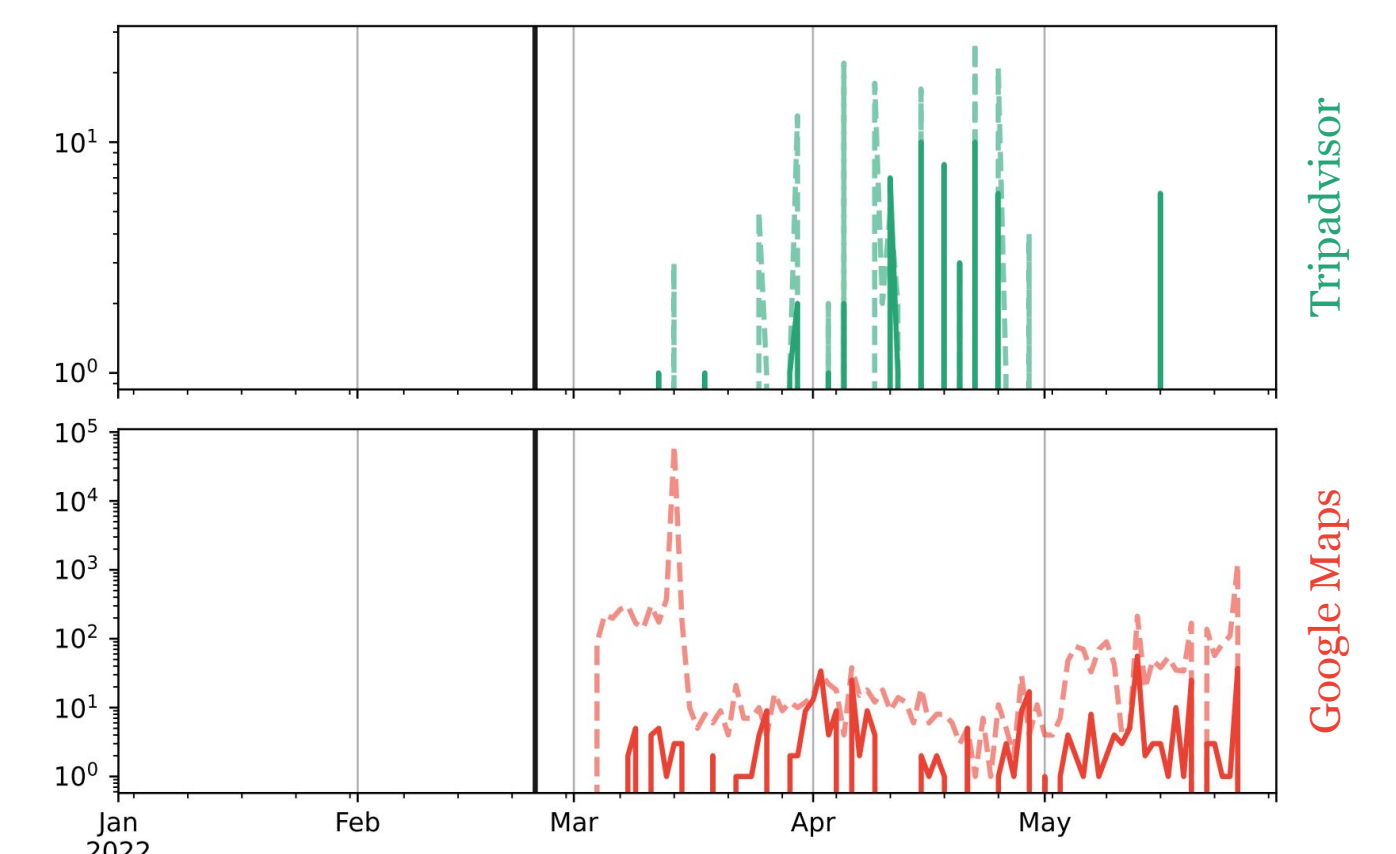


Figure 4. Removed posts labeled as war-related (solid line) and non-war-related (dashed).

Discussion and Conclusion



This study examines how Tripadvisor and Google Maps were used to circumvent Russian censorship during the 2022 Ukraine war. The platforms responded with content moderation, highlighting the need for **effective policies** and **specialized moderators** to manage unexpected social phenomena. The findings provide insights for researchers, policymakers, and service operators into harms that can arise from the use of seemingly innocuous services.

