ROBOTROLLING
2019. ISSUE 4

ROBOTROLLING

PREPARED AND PUBLISHED BY THE

NATO STRATEGIC COMMUNICATIONS CENTRE OF EXCELLENCE

Executive Summary

In this edition of Robotrolling, we expose a coordinated network of bots on Facebook, Twitter, and VK connected to a militaristic YouTube channel. Through our analysis, we discovered that the group of automated accounts is used to spread anti-NATO videos in the pro-Donbass information space. Our findings demonstrate that the video-sharing platform is a ripe target for robotic exploitation.

During this period, the level of Russian-language bot activity decreased on Twitter. Meanwhile, English-language bot activity remained unchanged. On VK, the volume of messaging increased by 8%. Nearly a quarter of unique users engaging with NATO-related topics were identified as bot accounts.

Russian- and English-language conversations about the NATO presence in the Baltic States and Poland peaked on 13 August on both Twitter and VK. On Twitter, English-language bot and anonymous accounts targeted Poland, while the Baltic States received the majority of Russian-language bot attention.

Events commemorating the 80th anniversary of the Second World War attracted significant levels of fake engagement throughout the monitoring period. We observed two recurring anti-NATO narratives circulating in this context: (1) NATO is occupying the Baltic States and Poland, and (2) NATO supports fascism.

Finally, this instalment of Robotrolling provides a glimpse into the flourishing world of commercial social media manipulation or, put simply, bots for hire. In a forthcoming report, we measure the inability of Facebook, Instagram, YouTube, and Twitter to counter online manipulation. ■

The Big Picture

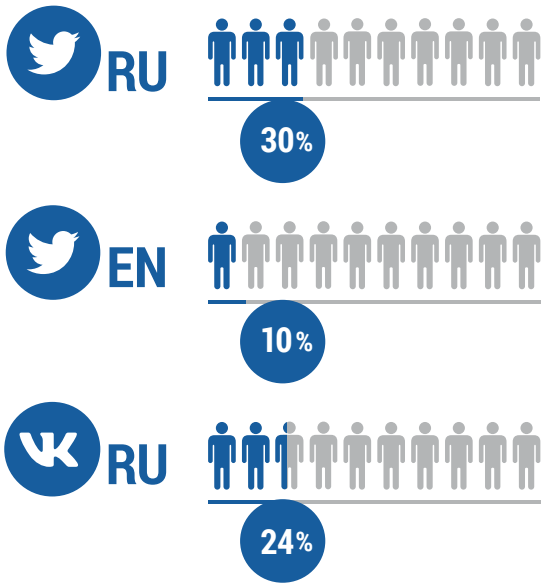
Robotrolling analyses the manipulation of information regarding the NATO presence in Estonia, Latvia, Lithuania, and Poland on the social media platforms Twitter and VK. Online manipulation is driven primarily by automated accounts (bots) and coordinated, anonymous human accounts (trolls). This issue of Robotrolling discusses the key findings and trends that emerged throughout the period 1 August – 31 October 2019.

During this quarter, we observed a decrease in the overall volume of conversations referencing NATO's presence in the Baltics and Poland on Twitter compared to the previous monitoring period, 1 May – 31 July. In contrast, on VK we saw an increase in both the total number of posts, and in the percentage coming from robotic accounts.

On Twitter, the volume of English-language messaging remained largely unchanged since the third quarter; the overall reduction was driven by a 25% drop in Russian-language activity. We also observed a decrease in Russian-language bot activity on Twitter, from 57% the previous quarter—the highest percentage of bot activity recorded this year—to 44%, the lowest percentage of bot activity recorded since data collection began in March 2017. Though sharp, this drop is largely explained by unusually high levels of bot activity in the previous quarter. We attribute this spike to the increase in NATO military exercises during the summer months.

The volume of automated English-language output on Twitter decreased from 17% to an estimated 12% of content produced by bots.

In the same period, VK users generated a total of 24 200 posts about NATO in the Baltics and Poland. Compared to the previous period, discussions regarding NATO's presence in Eastern Europe increased by 8%. Our algorithm identified 48% of VK timeline posts as automated, constituting an increase in bot activity of roughly 6% from the previous quarter. Nearly 25% of unique users engaging with NATO-related narratives on VK were identified as bot accounts. ■



Country Overview

Russian- and English-language messages about the NATO presence in the Baltics and Poland peaked on 13 August on both Twitter and VK. Though the spikes coincided, Russian- and English-language audiences discussed different events. While English-language users commented on the US threat to move NATO troops from Germany to Poland, Russian-language users focused on an incident over the Baltic Sea involving Russian and NATO warplanes.

English-language bot activity decreased significantly following 13 August, remaining low for the duration of the monitoring period. In contrast, Russian-language bot activity spiked several times throughout this quarter. These spikes coincided with NATO naval exercises in the Baltic Sea, events commemorating the start of World War II, and the publication of a strategic report entitled 'How to Defend the Baltic States' by a DC-based think tank.

On Twitter, Poland was targeted disproportionately by English-language bot and anonymous accounts, whereas Lithuania, Latvia, and Estonia received the bulk of Russian-langauge bot attention. On VK, bots disseminated messages about all four countries relatively equally. The largest spike on VK was also on 13 August, with many fake users framing the incident over the Baltic as a NATO provocation.

Estonia

The number of robotic posts regarding Estonia on Russian-language Twitter dropped to half the level of the previous quarter. Throughout September, bots circulated a satirical article about NATO building a small replica of Moscow on a training ground in Estonia for bombing in order to 'raise the morale' of Estonian soldiers. Several Twitter bots shared the link on the same day at the same time, indicating a coordinated effort to disseminate the article.

Latvia

Of the four countries, Latvia was mentioned most frequently by Russian-language bots. This conversation was driven by the controversies about Nazi and Soviet histories that re-emerged against the backdrop of events commemorating the Second World War. Russian-language bots amplified the narratives that (1) Latvia and NATO are Nazi sympathisers and that (2) the Baltics are being occupied by NATO forces.

Lithuania

This quarter, Russian-language bot activity on Twitter and VK responded to a six-month rotation of NATO troops in Lithuania. The deployment featured a battalion of roughly 500 troops and dozens of tanks to Lithuania near the border with Belarus. This followed a statement by the President of Lithuania, in which he characterised NATO's current presence in the country as 'insufficient' to deter Russia. The increased deployment was criticised as a demonstration of force by Alexander Lukashenko, Prime Minister of Belarus, whose outrage was shared widely by bots on both platforms.

Poland

Poland was disproportionately targeted by English-language bots this quarter. The English-language space was dominated by positive responses to US Ambassador to Germany Richard Grenell's statement that Germany's failure to increase defence spending will force the US to transfer their troops to Poland. Russian-language users on both platforms focused on Polish Prime Minister Mateusz Morawiecski's comments that Germany should increase its NATO contributions. On both platforms, Russian-language bots engaged with accusations that NATO dictates Polish policy. ■

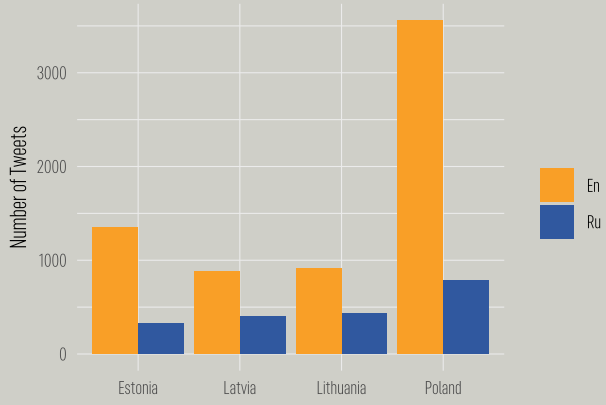


Figure 2: Total number of tweets mentioning NATO and Estonia, Latvia, Lithuania, or Poland, by language.

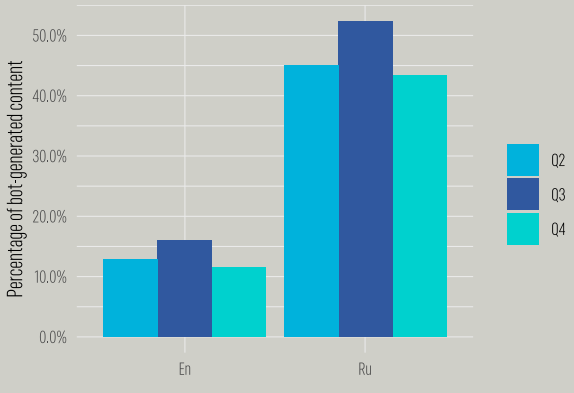


Figure 3: Changing percentage of bot posts on Twitter, by language.

Themes

This quarter, we observed the resurgence of two narratives that fuel negative perceptions of the NATO presence in the Baltics and Poland: (1) NATO is occupying the Baltic States and Poland, and (2) NATO supports fascism. These narratives have been identified in media coverage of the Enhanced Forward Presence (eFP) in Estonia, Latvia, Lithuania, and Poland since the deployments began in 2016. Their renewed appearance correlated with events commemorating the Second World War in Europe, especially with regard to Poland and Latvia. While Russian-language bot activity spiked several times during this quarter, posts propagating these narratives were observed throughout.

The narrative that Poland is a vassal state of NATO emerged in early September, as various world leaders gathered in Warsaw to commemorate the 80th anniversary of the outbreak of the Second World War. Missing from the crowd was Russian president Vladimir Putin; Russian representatives had not been invited. In response, Russia's Foreign Ministry spokesperson Maria Zakharova alleged that NATO was behind Poland's decision to exclude Russia. On 5 September, we observed a spike in bot activity sharing the story on Russian-language Twitter, with our algorithm identifying 71% of tweets referencing Zakharova as posted by bots.

In a similar vein, bots on both Twitter and VK promoted content accusing NATO of directing Polish officials to demolish a monument to Soviet soldiers in Trzcianka. Although the monument was destroyed in 2017, in late September bots circulated a recent article asserting NATO's complicity in the demolition as part of a deliberate anti-Russian campaign. On VK, 100% of posts linking to this article were shared by automated users.

This narrative persisted throughout October as Moscow celebrated the 75th anniversary of what Russia considers to be the 'Liberation of Riga by the Soviet Army'. The Latvian Foreign Ministry objected to the commemoration on grounds that the Red Army's entry into the Baltics signified the resumption of Soviet occupation. In response, Zakharova claimed that Latvian defence spending is in fact 'payment for occupation' by NATO. Users on both Twitter and VK shared a TASS report of Zakharova's statement. On Twitter, 70% of shares were generated by either anonymous or bot accounts, while 75% were classified as bots on VK.

The second narrative—NATO supports fascism—gained traction in early October when the Kremlin-backed media outlet RT criticised Latvian Defence Minister Artis Pabriks for delivering a speech in honour of members of the Latvian Legion. According to RT, Pabriks glorified Nazism, while NATO 'looked blindly' at the alleged rise of neo-Nazi sentiment in Latvia. The surrounding conversation centred on historical revisionism and Russophobia in the Baltics.

The high levels of fake engagement surrounding these controversies indicate how historical themes are a popular target for online mobilisation. ■

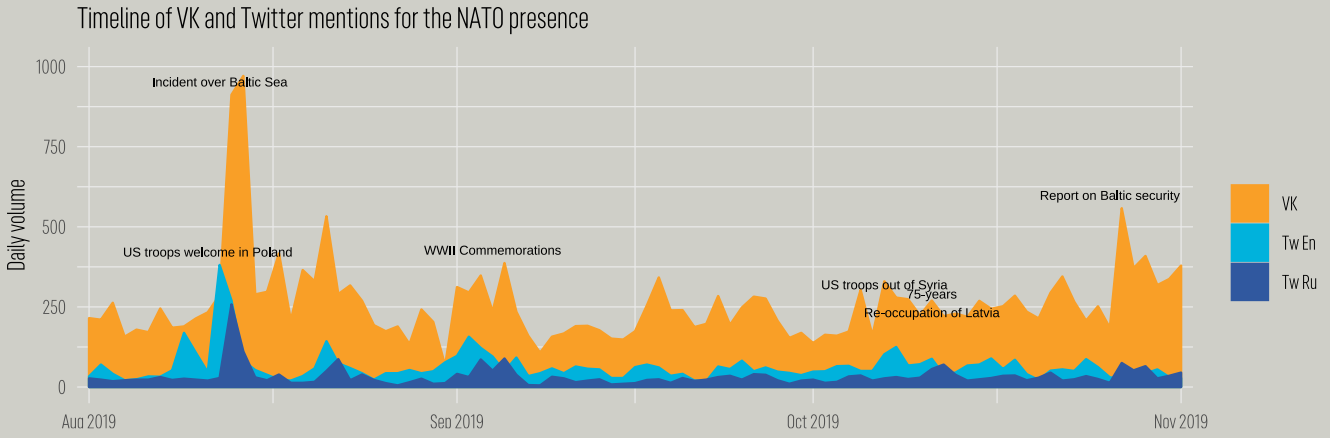


Figure 4: Timeline of VK and Twitter mentions.

Robo-topics

This quarter, we observed an intriguing pattern in the Russian-language Twitter space: multiple videos from the same YouTube channel—named *Andrey Andreev* but bearing a logo reading *BystroNovosti* [Fast news]—were shared en masse by bot users. Through further investigation, we uncovered a network of bot accounts spanning Twitter, VK, Facebook, and multiple pro-Donbass news resources that systematically share content from this Russian YouTube channel.

The channel was created in 2011. It attracted little attention until 2019, when the channel suddenly devoted its output to military-themed content. The total subscriber count more than tripled from 2 800 in December 2018 to 9 800 in January 2019. The monthly subscriber count has continued to grow exponentially since then. Today it boasts 52 500 subscribers, and has received over 26 million total views on its 981 videos.

The videos share a simple format: a voice-over reads a news report published by one of a number of pro-Donbass resources. The videos consist entirely of still images of military hardware, while their messages support recurring anti-NATO, anti-Ukraine, and pro-Russian themes.

We take the example of one video, *NATO is ready to sacrifice Poland and the Baltic States in the conflict with Russia*, to illustrate how the network disseminates its content. Figure 5 visualises the level of bot coordination involved in sharing the video. The video was uploaded to YouTube at 5AM on 22 August. Its content was based on an article from pro-Donbass website topwar.ru, published a few days earlier. Fifteen minutes after the video was uploaded, an automated cascade triggered crossposting by bots circulating the video on Facebook, Twitter, and VK. These automated users shared the video at exactly the same time, down to the second.

Following this burst in automated activity, the video was shared on several fringe pro-Donbass web resources. Each video is typically embedded directly in the webpage with little text accompanying it, mimicking its composition on YouTube. The speed and regularity with which material from the channel is hosted indicates that crossposting is in some cases automated.

Currently the channel produces multiple military-related videos every day. Often copied verbatim from a simple text source, audio and images are combined to create video content which is then automatically promoted by the network of bot accounts in the hope of making the content go viral. In some cases this tactic is effective: the most-viewed videos on this channel received up to 900 000 views. Regardless, every video spawns a dozen or more news reports within the pro-Donbass information space, each of which also has the potential for catching public interest. ■

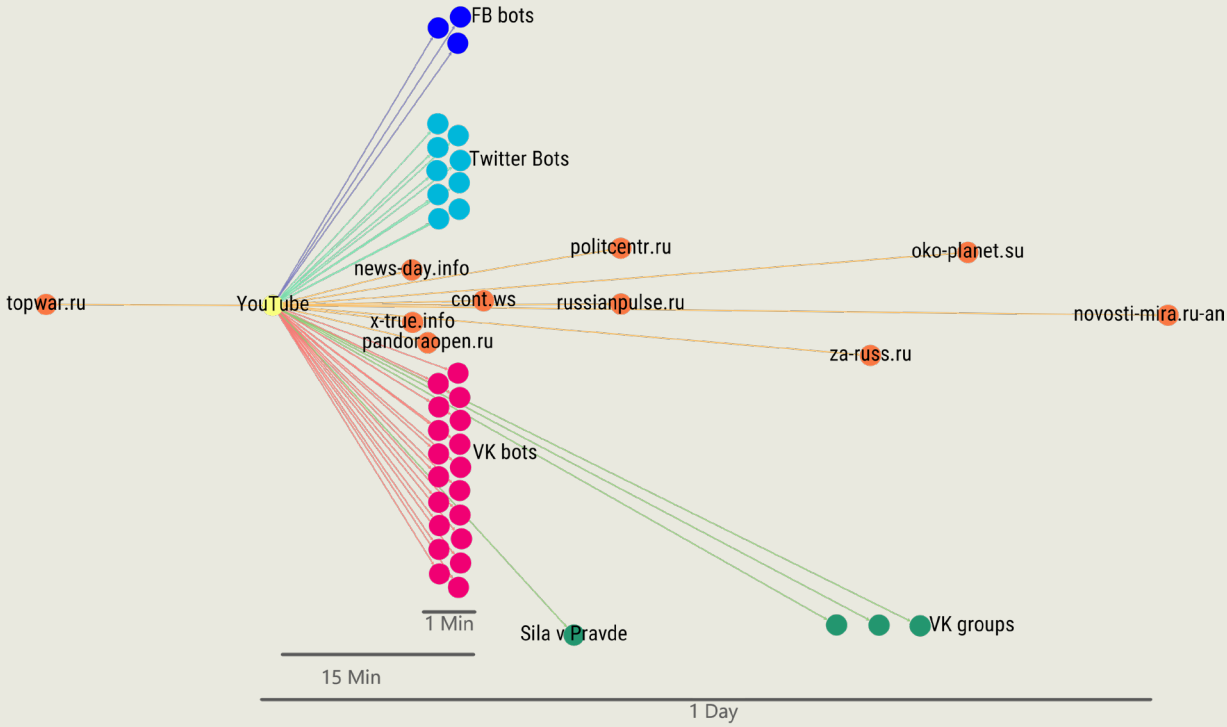


Figure 5: Network automatically disseminating BystroNovosti.

In Depth: Bots for Hire

Since the 2014 invasion of Ukraine, antagonists seeking to undermine democratic institutions have increasingly turned to social media manipulation to spread disinformation and interfere in democratic elections. In 2018, major social media companies agreed to the self-regulatory EU Code of Practice on Disinformation, which includes several commitments, ranging from transparency in political advertising to the closure of fake accounts. This year, the NATO StratCom COE conducted an experiment to evaluate the ability of social media companies to identify and remove inauthentic behaviour from their platforms.

The forthcoming COE report Falling Behind: How Social Media Companies are Failing to Combat Inauthentic Behaviour Online presents the results of this experiment and provides a glimpse into the world of commercial social media manipulation. The cross-platform analysis exposes the flourishing market of social media manipulation and the ease with which a range of actors can purchase social media engagement in the form of comments, clicks, likes, and shares. Ultimately, the study finds that self-regulation is losing the battle against the manipulation industry.

In order to test the ability of social media companies to counter online manipulation, the study evaluates how effectively Facebook, Instagram, Twitter, and YouTube responded to manipulation of 105 social media posts. At a total cost of just 300 EUR, 16 separate Russian and European manipulation service providers delivered a total of 54 000 fake interactions (comments, likes, shares, and video views). The authors were able to identify nearly 20 000 separate accounts being used for social media manipulation. Within this pool of identified inauthentic accounts there was at least one known pro-Kremlin troll-account.

Six weeks after purchase, 4 in 5 of the fake engagements were still online, revealing how ineffective the platforms are at independently detecting and removing fake content. Most alarmingly, even after reporting a sample of the fake accounts to their respective platforms, more than 95% remained active.

The authors assess and compare the performances of Facebook, Instagram, Twitter, and YouTube's ability to counter abuse of their platforms. Overall, Twitter is the social network most effective at tackling the gamut of social media manipulation. From the perspective of Robotrolling, this finding is surprising, given the persistently high levels of manipulation reported. It can be inferred that the platforms Facebook, Instagram, and YouTube are also being manipulated to at least a degree comparable to the numbers reported for Twitter.

Based on their experiment, the authors recommend:

- Setting new standards for social media companies and require reporting based on more meaningful criteria
- Establishing independent and well-resourced oversight of the platforms' ability to counter inauthentic activity
- Increasing transparency of the platforms to understand the scope and effect of manipulation
- Regulating the market for social media manipulation

Although the fight against online disinformation and coordinated inauthentic behaviour is far from over, an important finding of this experiment is that the different platforms aren't performing equally poorly—in fact, some are significantly better at identifying and removing manipulative accounts and activities than others. Investment, resources, and determination make a difference.

The full report, Falling Behind: How Social Media Companies are Failing to Combat Inauthentic Behaviour Online is due to be published on the stratcomcoe.org website in December 2019. ■

Prepared by Dr Rolf Fredheim and Kristina Zina-Joy Van Sant published by
NATO STRATEGIC COMMUNICATIONS
CENTRE OF EXCELLENCE

The NATO StratCom Centre of Excellence, based in Latvia, is a Multinational, Cross-sector Organization which provides Comprehensive analyses, Advice and Practical Support to the Alliance and Allied Nations.

www.stratcomcoe.org | @stratcomcoe | info@stratcomcoe.org