ROBOTROLLING
2020. **ISSUE 1**

ROBOTROLLING

PREPARED AND PUBLISHED BY THE

**NATO STRATEGIC COMMUNICATIONS
CENTRE OF EXCELLENCE**

Executive Summary

Inauthentic English- and Russian-language conversations on Twitter about the NATO presence in Poland and the Baltic States peaked on 4 and 5 December, respectively, coinciding with the 2019 NATO Leaders' Meeting in London. Robotic accounts focused heavily on the meeting this quarter, particularly on English-language Twitter, which saw roughly 3 times the usual level of bot activity. On VK, an anomalous increase in activity from anonymous human-controlled accounts coincided with the meeting.

Due to the contentious atmosphere surrounding the meeting in London, a considerable increase in the proportion of posts generated by bots was observed on English-language Twitter this quarter. At the same time, Russian-language bot activity on Twitter decreased to the lowest level observed thus far.

In this issue of Robotrolling, we dig deeply into a sample of political pages amassed by a COE report on commercial social media manipulation in order to identify patterns in inauthentic activity on Facebook. We demonstrate that the 2019 elections in Ukraine were the primary focus of actors willing to pay for inflated social media engagement. Our analysis also reveals several shared traits among political manipulators on Facebook and provides a network visualisation that shows the connections between them.

As a new year of Robotrolling begins, we review trends observed in VK groups over the past 18 months. A steady reduction in the proportion of content shared in communities dedicated to the so-called Novorossia region and the Donbass coincides with inauthentic content increasingly being posted in community spaces such as private groups or pages.

The Big Picture

This edition of Robotrolling continues to monitor the online manipulation of information regarding the NATO presence in Poland, Estonia, Latvia, and Lithuania on the social media platforms Twitter and VK. Our analysis focuses on the activities of automated accounts (bots) and coordinated, anonymous human accounts (trolls). This issue tracks the key trends that emerged in the Russian- and English-language information spaces during the period 1 November 2019 – 31 January 2020.

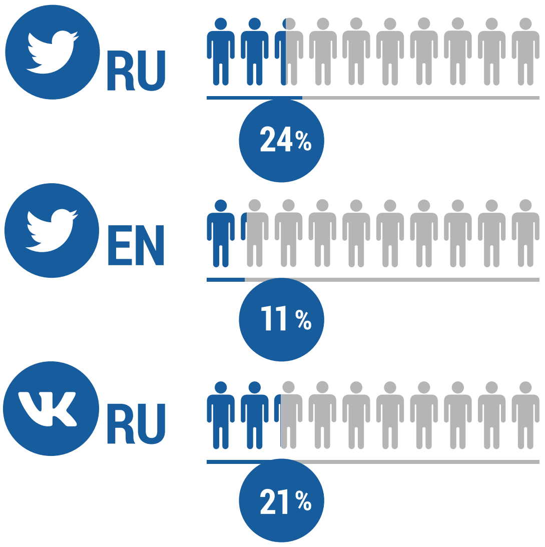
This quarter, we observed 11 500 messages on Twitter referencing the NATO presence in Poland and the Baltics, constituting a 30% increase in the overall volume of conversations compared to the previous period, 1 August – 31 October 2019. In addition to post volume, we observed an increase in the number of unique users engaging in these conversations. On VK, we noted a slight increase in post volume but a simultaneous decrease in total users active during this period. On VK, more than half of discussions about NATO in the region occurred in groups.

On Twitter, the volume of English-language bot activity increased considerably compared to the previous quarter. Overall, bot activity took up 15% of the English-language information space, a jump from 12% in the previous quarter.

In our last report, we observed the lowest percentage of Russian-language bot activity to date. This quarter, we witnessed an even more drastic reduction in messaging generated by Russian-language bot accounts, as relative bot activity fell from 43% to 38%. Despite commanding a smaller segment of the discussion space, the absolute number of bot posts increased this quarter.

Bots were less engaged on VK than on Russian-language Twitter this quarter. Despite witnessing an increase in active bot users, the average level of bot activity on the Russian platform fell to 35%. Consequently, more bot users contributed less to conversations about the NATO presence in Poland and the Baltics.

This quarter, Lithuania received the majority of Russian-language bot and anonymous activity on both Twitter and VK while Poland remained the primary target of English-language bots and anonymous users (see Figure 2). Many anonymous accounts are legitimate Twitter users, but spikes in messages from such accounts may also indicate manipulation by human-controlled accounts.



Country Overview

Discussions about the NATO presence in the Baltics and Poland peaked on both platforms in early December, coinciding with the NATO Leaders' Meeting in London. The bulk of Russian- and English-language activity this quarter focused on events in London and culminated on 4 and 5 December. While bot activity spiked during November, it decreased following 5 December and remained negligible for the duration of the quarter.

Otherwise, bot activity was driven by military exercises in the region, such as Estonia's Cyber Coalition cyber threat training, Lithuania's Iron Wolf II exercises, and Defender Europe 2020. Bots claimed that these exercises demonstrated that NATO was preparing for a large-scale conflict with Russia. On VK, bots circulated a Chinese article claiming that NATO's regional defence plan posed an imminent threat to Kaliningrad.

On 31 January, the final day of this quarter, we observed an anomalous spike in English-language user activity. On Twitter, adherents to the far-right fringe movement QAnon coordinated trolling around a tweet referencing NATO posted in 2014 by US representative Adam Schiff, lead prosecutor during the Senate impeachment trial of US president Donald Trump. QAnon began in 2017 as a conspiritorial pro-Trump movement against the so-called 'deep state', but has been recently creeping into the mainstream conservative political arena.

Estonia

Robotic activity directed at Estonia increased this quarter, corresponding with a cyber security exercise, a diplomatic visit to NATO troops, and an interview with Interior Minister Mart Helme about a separate Baltic defence plan. In late December, while visiting troops in Estonia, UK Prime Minister Boris Johnson condemned Russia's repeated violations of Estonian airspace. Bots circulated the response of the Russian embassy in London, which denounced his statement as propaganda to justify the military presence in the Baltics.

Latvia

In contrast with the previous quarter, Latvia was the country least mentioned by English- and Russian-language bot accounts on both platforms. In November, Latvian Minister of Defence Artis Pabriks called for the modernisation of Latvia's military arsenal and the establishment of a permanent NATO combat group in the country. Bots operating in the Russian-language information space framed Pabriks' proposal as an example of Russophobic paranoia.

Lithuania

Lithuania was the principal target of bot activity in the Russian-language information space. Inauthentic activity focused on Lithuanian president Gitanas Nauseda, who was notably vocal during and after the 2019 NATO Leaders' Meeting in London. Nauseda urged fellow NATO members to recognise Russia as a threat to the existing world order, stating that it is 'very important' to do so. In late December, Nauseda's criticism of Putin's assertion that Poland bears partial responsibility for the outbreak of the Second World War garnered bot attention.

Poland

This quarter, Poland once again was the subject of the highest number of English-language bot mentions on Twitter. This is likely due to Polish officials' public responses to criticism of NATO and to the controversy surrounding the 2019 NATO Leaders' Meeting. In November, Prime Minister Matteusz Morawiecki criticised President Emmanuel Macron's remarks that NATO is experiencing 'brain death'. Bot activity spiked again on 5 December when the Polish president stressed that Russia was not as an enemy but a neighbouring country whose unacceptable actions have resulted in many disagreements. ■

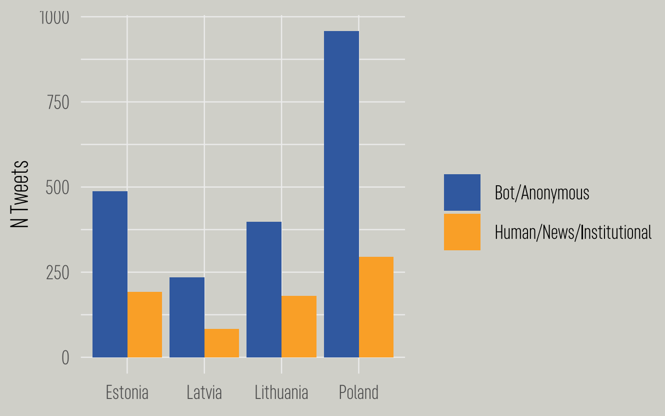


Figure 2: Total number of tweets mentioning NATO and Estonia, Latvia, Lithuania, or Poland, by account type.

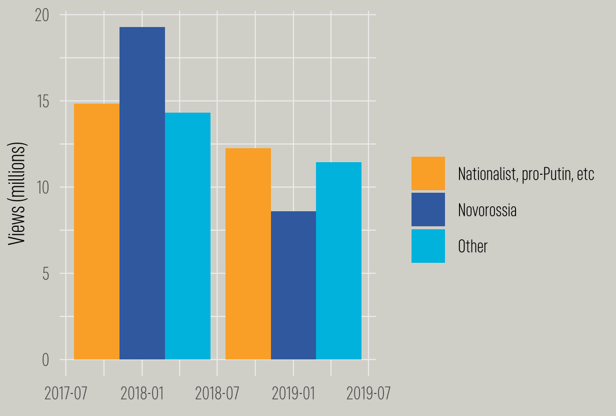


Figure 3: Views of content about NATO in the Baltics and Poland, posted to community spaces on VK and subdivided by community cluster.

Themes

This quarter, both authentic and robotic activity fixated on NATO's 70th anniversary Leaders' Meeting in London and the diplomatic disputes that preceded it. The meeting, held on 3 and 4 December, gathered NATO heads of state and government in London for their 30th formal meeting. The meeting attracted the highest volume of bot activity on Twitter this period, as one third of the total volume of English- and Russian-language bot posts this quarter were shared between 2 and 5 December. As a result, English-language bot activity was 3 times the usual level. Inauthentic users depicted the meeting as exposing deep divisions within the alliance, amplifying the theme of NATO in disarray throughout this quarter.

The dispute at the centre of the two day meeting emerged a week before it began, when Turkish president Erdogan announced that Turkey would oppose a NATO defence plan proposed for Poland and the Baltics if the military bloc refused to recognise the Kurdish YPG militia as terrorists. This resulted in a spike in Russian-language bot activity on 27 November, with several inauthentic accounts on Twitter framing Turkey as blocking NATO's 'anti-Russian' plans. On VK, bot users empathised with Turkey, emphasising that the NATO member is actively threatened by the YPG, while the Baltics and Poland face an 'imaginary threat' from Russia. The debate surrounding the Baltic and Polish defence plan continued as the NATO meeting commenced on 3 December.

As the events of the NATO meeting proceeded, we noted differences between English- and Russian-language bot activity. English-language bots disseminated periodic updates from the gathering, dramatically increasing their activity on 4 December when Turkey reversed its stance and supported the defence plan. Alternatively, Russian-language bots focused on the statements of Baltic and Polish leaders throughout the meeting. On 3 December, Russian-language bots circulated a report that the president of Lithuania called on fellow members to label Russia a threat and, immediately following the meeting, praised Poland for 'refusing' to call Russia an enemy of the alliance. Thus some pro-Kremlin outlets (notably Sputnik) took statements by the Lithuanian and Polish presidents out of context and presented them as contradictory, even though in reality they were not. This line was pushed by bots resulting in the increased volume.

Though bot engagement with the 2019 NATO Leaders' Meeting portrayed the alliance as institutionally weak and disorganised, additional bot activity indicated otherwise. Throughout this quarter, bots consistently disseminated the notion that NATO is an aggressive behemoth attempting to expand its presence along the Russian border. Russian-language bot activity on Twitter and VK intensified around military training exercises, with many users claiming that NATO is preparing for a large-scale confrontation with Russia. In late December, Russian-langauge bot users shared an RT article claiming that Poland was attempting to 'force' Belarus away from Russia and into NATO's sphere of influence. Similarly, inauthentic activity on both platforms reacted to Baltic and Polish leaders' support for Ukraine becoming a member of NATO and the EU. ■

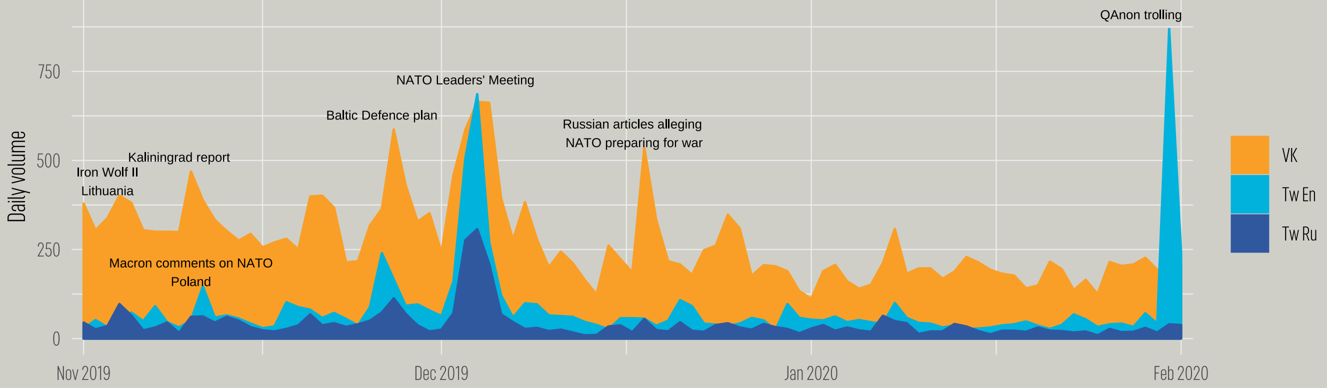


Figure 4: Timeline of VK and Twitter mentions.

Robo-topics

As the Robotrolling newsletter enters its fourth year of monitoring bot activity, we decided to review the patterns and dynamics we have observed thus far among groups on VK. In our second issue of 2019, we found that the vast majority of posts on VK originate from groups dedicated to the conflict in Ukraine and to Russian nationalism. There has been a sustained reduction in the proportion of content from communities dedicated to posting about the Donbass and the so-called Novorossia region, which encompasses the self-proclaimed republics of Donetsk and Luhansk. The volume of nationalist and pro-Putin content, on the other hand, has continued at roughly the same rate and has attracted similar levels of engagement. Whereas nationalist content continues to generate considerable interest, material posted in Novorossia communities gains ever less traction, insinuating a mismatch in the supply and demand of content as visualised in Figure 3.

Throughout this same period, the proportion of posts in communities less accessible for researchers has nearly doubled—from 26% to 45%. These communities are difficult to classify, because they either are closed groups, have hidden members lists, or are so peripheral within the normal network of groups that they fit poorly into our major categories of VK communities. While the proportion of these posts is very high, they have comparatively limited reach, accounting for ~3% of all views. However, at the start of November 2019 we observed that views for content within these communities had doubled.

A form of polarisation is taking place within VK: while on the one hand, the number of posts to inaccessible communities is on the rise, the relative significance of mega-groups and pages is increasing. 18 months ago, 50% of views for community content was disseminated in spaces with more than 150 000 members or subscribers; today that percentage has risen to 57%.

In the past, we have noted that most VK posts are published in groups or pages as opposed to users' timelines, and group content receives the vast majority of views. This trend has become even more dramatic in the past year. A year ago, only 5% of views were for timeline content (this includes human, anonymous, and automated accounts), but today that figure has dropped to 2.5%. This trend holds when we consider the metrics for number of shares and likes: the percentage of timeline shares has dropped from 16% to 11%, and from 18% to 13% for likes.

Finally, our analysis of VK this quarter revealed two insights. First, we observed that, for the second year in a row, VK users lost interest in NATO-related content during January. Just as in 2019, January 2020 saw approximately half the normal level of activity. Second, we observed a surge in the proportion of views for content from anonymous accounts during the NATO Leaders' Meeting, perhaps signalling a coordinated effort to dominate the information space during the event. ■

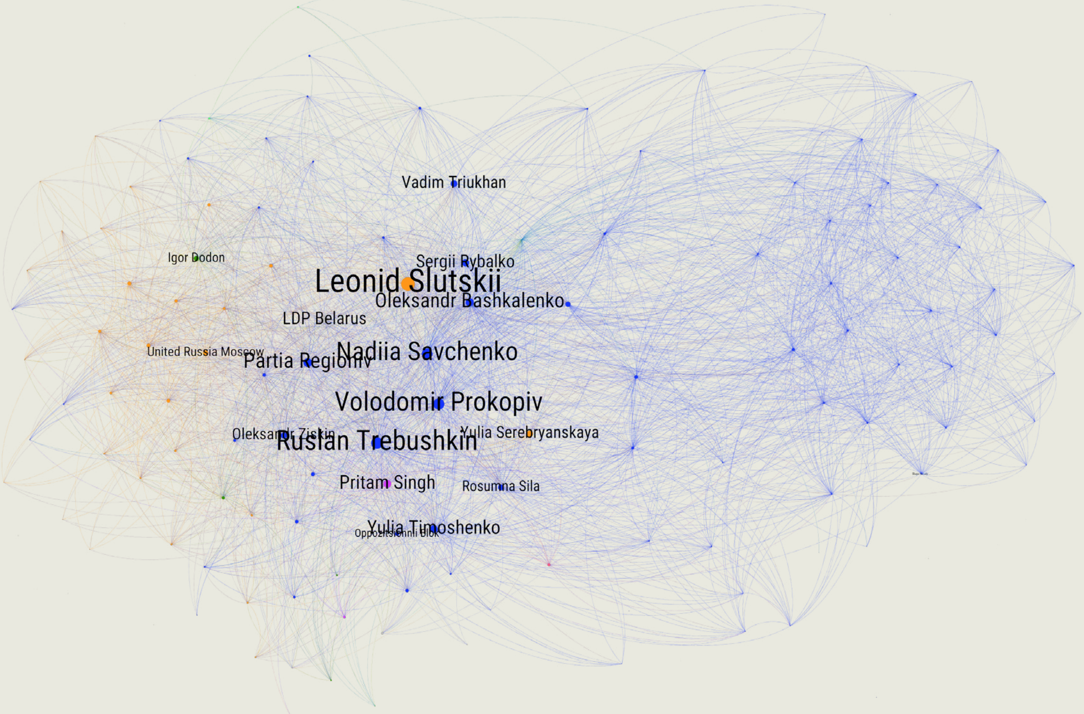


Figure 5: Political pages promoted by the same fake accounts. Nodes in orange are Russian-based politicians or parties; nodes in blue are Ukrainian. Larger nodes are liked by a larger number of fake accounts. See 'In Depth' section.

In Depth: Bots for Hire on Facebook

In this case study, we explore how political pages on Facebook have made use of commercial social media manipulation services. This data derives from the recently published NATO StratCom COE study, Falling Behind: How Social Media Companies are Failing to Combat Inauthentic Behaviour Online. The report demonstrates how the world's leading social media companies are struggling to defend their platforms against the growing social media manipulation industry. In this experiment, the authors purchased engagement on 105 posts on Facebook, Instagram, Twitter, and YouTube in order to test the ability of social media companies to identify and remove bought manipulation. The report's findings, which suggest that undetected inauthentic activity may interfere in democratic processes, have reverberated internationally and have been shared by major media outlets, such as the New York Times, the BBC, and Politico.

By purchasing thousands of fake engagements, researchers at the NATO StratCom COE were able to observe networks of inauthentic users that provide social media manipulation services on Facebook, Twitter, YouTube, and Instagram. While the vast majority of purchased engagements on social media were used for commercial purposes, the authors identified bought engagement on 721 political pages and 52 government pages, carried out by at least one known pro-Kremlin bot account. We compiled these political Facebook pages in a dataset and analysed those that received the highest levels of engagement, as well as the for-hire accounts that delivered it.

Our analysis resulted in three main takeaways. First, it is clear that the 2019 Ukrainian presidential and parliamentary elections were the main target of inauthentic activity; of the 20 most-engaged with pages, 13 related to elections in Ukraine. Among these pages were Ukrainian politicians, political parties, and government entities. Additionally, we found manipulation on several pages associated with the 2019 election of the Moscow City Duma, the regional parliament in Moscow. The remaining pages in our sample were connected to Singaporean, Belarusian, Moldovan, Polish, Georgian, Indian, and US politics.

The second finding relates to the accounts that provided politically-charged social media manipulation services. We observed that the same accounts were active on politically and ideologically diverse pages, often supporting opposing views or competing politicians simultaneously. The resulting tightly-woven network structure shown in Figure 5 is less of a network than a free-for-all where everyone is connected to everyone else via the activity of manipulation providers.

Finally, we observed that individual account activity was geographically varied. The pages that a single account engaged with were often tied to the politics of several countries, primarily Ukraine and Russia, but also Belarus, Poland, India, and others. For example, we observed that the same account interacted with the pages of an Italian politician, a Ukrainian politician, and the Liberal Democratic Party of Belarus.

These findings indicate that the Ukrainian and Russian information spaces are especially polluted by commercially-driven inauthentic activity. Robotrolling has consistently found this to be true around political discussions on Twitter, a platform that is far less popular among Russian speakers than Facebook. They also echo the conclusions of the Falling Behind report: Facebook may be adept at blocking fake account creation, but those accounts that bypass Facebook's security mechanisms are free to engage in inauthentic activity.

Our conclusions also have implications for social media regulation. This type of online behaviour—engagement with ideologically and geographically inconsistent targets—exhibits clear inauthentic properties. Identifying these accounts as used for commercial purposes should be low-hanging fruit for major social media companies; their failure to do so further demonstrates the lamentable insufficiency of current bot-detection methods. ■

Prepared by Dr. Rolf Fredheim, Kristina Van Sant, Marina Paramonova and 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