

ROBOTROLLING  
2018. ISSUE 2



# ROBOTROLLING

PREPARED AND PUBLISHED BY THE  
**NATO STRATEGIC COMMUNICATIONS  
CENTRE OF EXCELLENCE**



# Executive Summary

Anonymous users stole the show this quarter. Never before have we observed such high levels of activity from anonymous accounts. At the same time, bot activity in Russian-language conversations about NATO activity in the Baltics and Poland has emerged from its winter slumber.

In the wake of the Skripal poisonings in the UK in March, Russian-language bot and anonymous activity about NATO more than doubled. Mentions of NATO on VK, in contrast, have been stable and declining during the whole period.

Social media companies are working to end platform misuse. But malicious activity is evolving. Today, anonymous accounts are dominating the conversation. These accounts are either operated manually, or they have become advanced

enough to fool human observers. The responses from open and free societies to the problem of online malicious activity have neither been strong enough, nor consistent enough.

Figures presented in this issue reveal a disparity between the conversation quality in English and Russian-language spaces. Currently, the Russian-language conversation about NATO in the Baltics and Poland has six times the proportion of content from bot and anonymous accounts. As Twitter has taken steps to remove bots, the disparity has only widened. We assess that 93% of Russian-language accounts in our dataset are operated anonymously or automatically. In no way does this conversation mirror opinions of citizens. Journalists, policy makers, and advertisers take note! ■

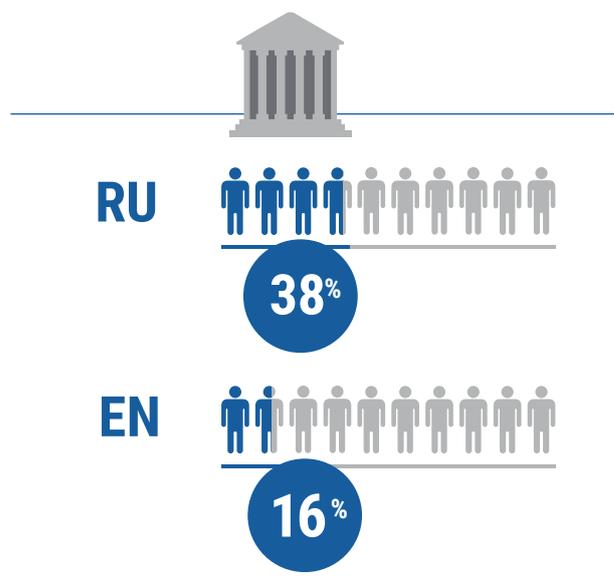
## The Big Picture

The Russian-language environment on Twitter is as filthy as ever. There was a sharp drop in bot activity for English-language content this quarter, when the posts from bots declined by 10 percentage points. The slight decline in bot-proportion can be explained by a surge in activity from anonymous accounts. Such accounts, especially those highly active and expressing strong political and ideological views, have a disproportionate likelihood of being fake.

Only 7% of active users who post in Russian are recognisable as humans or institutions; the remaining 93% are news accounts, bots, hybrid, or anonymous. The equivalent percentages for the English space are 45 and 55. By this indicator, the Russian-language space has more than six times the density of possibly inauthentic content.

In the period 1 February – 30 April 2018, Russian-language bots created 54% of all Russian-language messages about NATO in the Baltic States and Poland. Overall, 38% of accounts posting in Russian were predominantly automated. In comparison, bots make up 16% of accounts tweeting in English. These bots created 21% of all English-language messages this quarter.

In this issue we rely on a nuanced account classification scheme to go beyond a bot/human dichotomy. We analyse Twitter-mentions of NATO together with one or more of the host countries Estonia, Latvia, Lithuania, and Poland. The total number of posts considered is 6 400, of which 30% are in Russian. The number of active users is 3 900. Both the number of posts and the number of users increased by 15% compared to the period November 2017 – January 2018. ■



# Country Overview

In April, Russian-language bot activity about NATO increased sharply. Volumes were low during the winter of 2017–18, but the Skripal poisonings together with tensions in the Middle East prompted a new wave of automated and anonymous activity. More than half the Russian-language tweets about the NATO presence in Latvia and Estonia came from bot accounts.

The main news stories promoted by bots were about disturbances involving British officers in Latvia in late April, and Ukrainian (Russian-language) news reports on statements by Polish and Estonian officials about the country's hopes of joining NATO. English-language bot activity follows a similar pattern, albeit at greatly reduced volumes.

Bots promoting content in support of Ukraine joining NATO is a reminder that social media spaces are contested and interfered with by actors representing a diversity of interests.

## Estonia

This quarter Russian-language bots accounted for 55% of the 510 posts about NATO and Estonia. Compared to the previous quarters, these figures are modest. The main events coinciding with bot interest were the 14-year anniversary of Baltic States' NATO membership on 29 March 2018, and reports in mid-February and early March quoting Minister of Defence Jüri Luik that Estonia supported Ukrainian membership in NATO.

## Latvia

Compared to the previous quarter, the number of Russian language posts about NATO and Latvia has doubled, while the proportion of bot-messages has increased to 53%. The main incident emphasised by bots was an altercation in Riga between three British officers and local police on 24 April. Earlier that month reports about Russian missile tests forcing the closure of Latvian airspace were also disseminated by bots.

## Lithuania

Yet again, Lithuania attracted the lowest levels of unwanted social media attention. Activity about Lithuania was modest, even in early February when NATO Deputy Secretary General Gottemoeller visited Lithuania to mark the anniversary of NATO deployments to the Baltics. At the same time, President Dalia Grybauskaitė warned about the threat posed by the Russian nuclear-capable Iskander missiles located in the Kaliningrad enclave. In April, English-language bots shared reports that Lithuania wanted a stronger NATO anti-aircraft capacity in the region.

## Poland

This quarter a series of news reports about Poland's purchase of the American Patriot air defence system drew extensive Russian- and English-language comment from anonymous users. On 28 April, bots promoted Ukrainian news reports that Polish Foreign Minister Jacek Czaputowicz had advocated less formal talks in order for the Ukraine-NATO Commission to succeed.

Messages about NATO and Poland consistently attract elevated volumes of anonymous commentary, both in Russian and English. ■

Country Comparison of Russian-Language Bot Activity

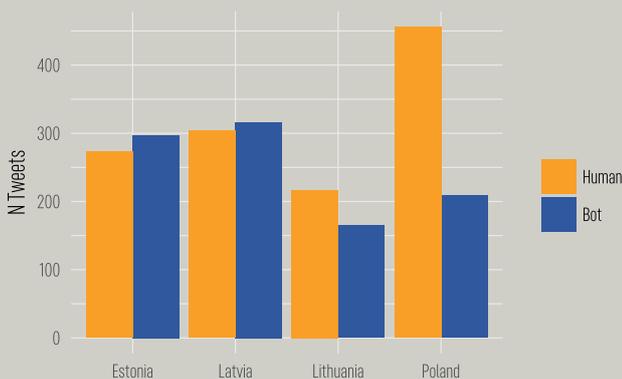


Figure 2: Distribution of Russian-language posts mentioning NATO and Estonia, Latvia, Lithuania, or Poland.



Figure 3: Messages from bots and anonymous users relative to content from recognisable human and institutional accounts. Values relative to May 2017 levels.

## Themes

This month's theme is how anonymous users are taking over the Twitter conversation. Over the last twelve months Twitter has been effective primarily at reducing English language bots. The modest reduction in Russian-language bot activity has been more than compensated for by a steady growth in human-controlled anonymous accounts.

The case of Poland illustrates this. In March 2018, Activity from anonymous accounts focused on Poland's purchase of the American Patriot air defence system. The proportion of anonymous human users posting about NATO and Poland is ~ 40% higher in both English and Russian than the equivalent figures for the Baltic States. The high density of such accounts suggests this conversation is being manipulated by actors posing as citizens.

Anonymity can be beneficial in many situations: citizens can speak about controversial topics without fear of judgement and prejudice. However, the ever-increasing proportion of anonymous accounts in specific conversation and language areas indicate that anonymity is being abused to cloak manipulation on social networks. Many anonymous accounts are fake identities created to promote sponsored messages.

Figure 3 shows how the posting activity from bot and anonymous accounts has changed year-on-year. Numbers are calculated relative to the proportion of tweets from recognisable human and institutional accounts for May 2017. Adjusting volumes in this way controls for changing activity levels due to the news agenda. A downward trend indicates declining levels of suspicious activity, while an upward trend indicates the opposite.

The dashed light blue line in Figure 3 shows that the reduction in bot activity is clear for English-language content. The ratio of tweets from anonymous accounts to those from recognisable human and institutional accounts has reduced to 40% of May 2017 levels. Anonymous human activity has also declined for English-language content, resulting in a space where authentic voices make up a larger share of the conversation.

For Russian-language content the picture is gloomier: in April 2018, bot activity spiked sharply, and although it is on an overall downward trajectory, it remains at about 90% of May 2017 levels. The relative difference in quality between the Russian and English online spaces has increased, because the volume of automated English-language tweets has more than halved in a year, while automated Russian-language messages declined only slightly.

The reduction of Russian-language bots over the past year has been matched by a rise in content from anonymous, human operated accounts. Compared to the volume-levels from recognisable humans, institutional accounts, and news outlets, the Russian-language bot level has dropped to 90% of May 2017 levels, but anonymous activity has increased by 20% over the same period. ■

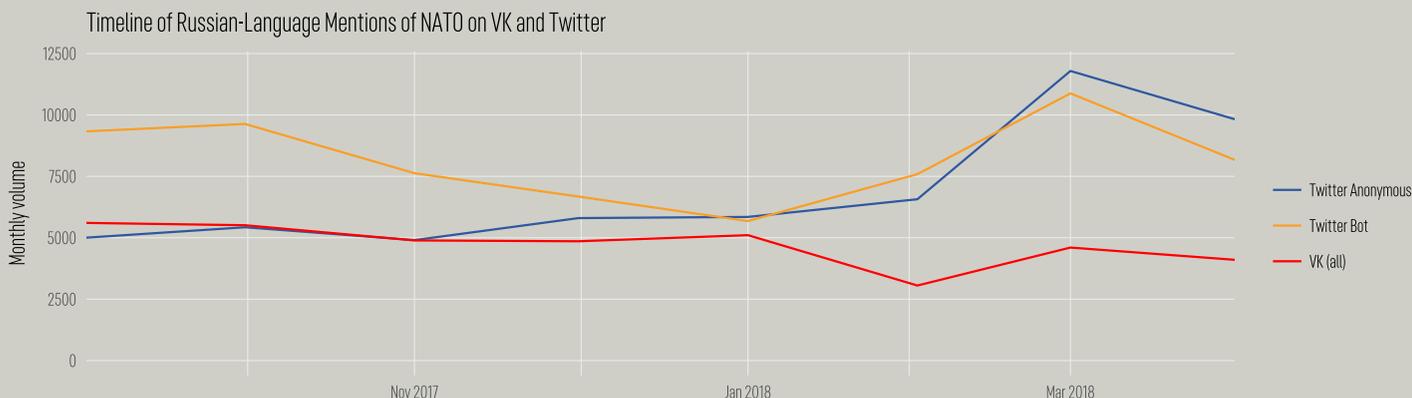


Figure 4: Timeline of all Russian-language mentions of NATO on Twitter and VKontakte for the period 1 September 2017 – 30 April 2018

# Robo-topics

This quarter saw strong reactions to three Russia-related incidents, only tangentially concerning NATO: the poisoning on 4 March of Sergei and Yulia Skripal and the subsequent expulsion of Russian diplomats; the re-election of Vladimir Putin as Russian President on 18 March, and on 7 April the chemical attack on the Syrian city of Douma. These events offer a backdrop against which to evaluate the efficacy of Twitter’s attempts to crack down on virtual manipulation.

Social media responses to these events spilled over into the Russian conversation about NATO. Figure 4 (previous page) shows the dynamic for all Russian posts mentioning NATO on Twitter and VK. Posts from Russian-language bots and anonymous accounts both increased by more than 250%. The trajectory change whereby modest reductions in bot content is outstripped by increased anonymous activity is visible also for the Russian-language conversation about NATO as a whole: anonymous activity has more than doubled in the wake of the Skripal poisonings.

Figure 4 shows that in March 2018, Russian-language bot activity about NATO surged past 11 000 messages per month and is now at the highest level seen to date. Mentions of NATO on VK, in contrast, have been stable and declining during the whole period.

Compared to the reference point of VK, it appears that inorganic messaging about NATO has been rampant on Twitter.

These trends are strong evidence that online manipulation remains rife in the Russian-language Twitter space. They demonstrate that Twitter’s attempts to reduce the presence of bots on the platform have incentivised other forms of virtual manipulation. Anonymous activity is not merely rising in relative significance because bots are removed from the platform; it is rising in absolute terms.

Changes introduced by Twitter to tackle the twin problem of bot and troll activity may be best suited to rooting out steady bot activity, rather than sudden deluges of content as during online disinformation campaigns. This indicates that there is much more work for Twitter to do in order to maintain a platform which protects its users against disinformation campaigns.

Robotrolling is based on a sample of Twitter-data about military activity in the Baltics and Poland and is not representative of Twitter or content on other social media platforms. See our online FAQ for details on methodology. ■

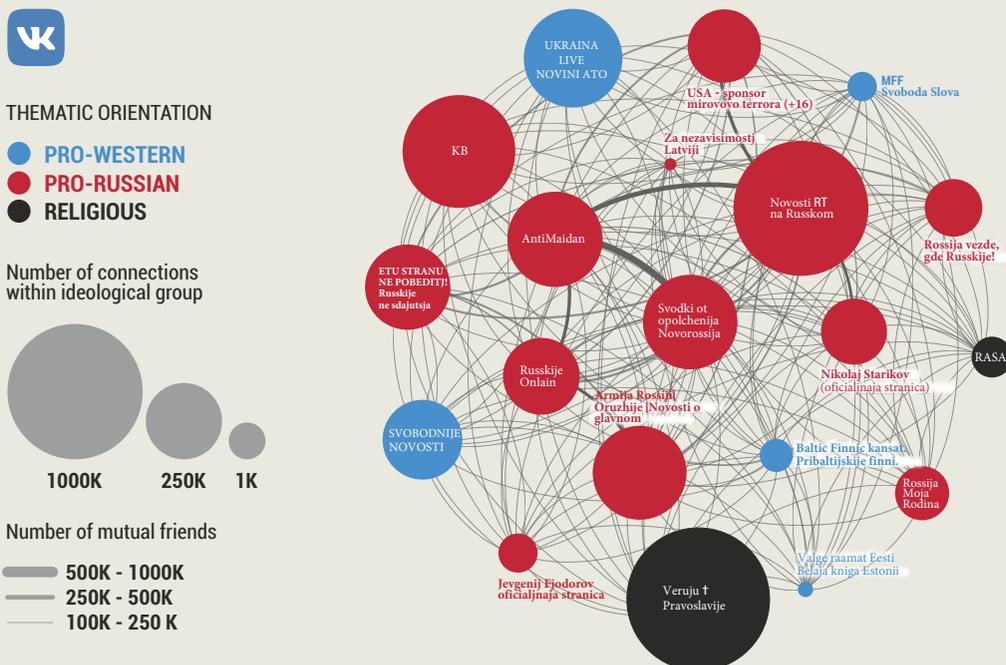


Figure 5: Orientation of ideological groups on VKontakte, as classified by the authors of 'Virtual Russian World in the Baltics' (see back page).

# In Depth: Virtual Russian World in the Baltics

The soon-to-be-released NCDS – COE report 'Virtual Russian World in the Baltics' draws on a massive anonymised collection of more than 2.8 million Russian-speaking users active on the social networks Facebook, Odnoklassniki, and VKontakte. The study gives an unprecedented insight into Russian-language online communities. The study also details how user groups are inundated by a toxic mix of disinformation and propaganda.

The authors reach a number of exciting conclusions of relevance to readers of Robotrolling:

- 1) The number and size of Russian-language Baltic communities is growing. The groups are overwhelmingly pro-Kremlin and anti-Western in orientation. Figure 5 (previous page) shows the thematic orientation of ideological Russian-language VK groups.
- 2) The Russian-language groups are dominated by a core of hyperactive users. As much as 70% of the ideological content stems from the top 10% of the most active accounts.
- 3) These accounts are identified as primarily human-operated, rather than robotic.
- 4) They are heavily engaged in promoting ideological content. Yet they are not ideologically motivated: rather than passionately promoting one or more related causes, they rapidly cycle through the issues of the day. This, the authors hint, suggests operators of such accounts may be financially compensated for their activity.
- 5) Offline events serve as pretexts to push certain topics into public discussion. Particular attention is paid to topics related to the Second World War, the USSR, present-day Russia, and anti-Western sentiment.
- 6) There is a small but significant proportion of active ideological users in every network in each of the Baltic states.
- 7) They tend to specialise in specific functions, namely creation (Writers), distribution (Distributors), or consumption (Readers) of ideological information.

'Virtual Russian World in the Baltics' is due to be published on the stratcomcoe.org website in May 2018. This cooperation between Estonia's National Centre of Defence & Security Awareness and the NATO StratCom COE, led by Dmitri Teperik, presents a psycholinguistic analysis of online behaviour and ideological content among Russian-speaking social media users.

The study reveals the complex and fluctuating trends in Russian-language online communities on VKontakte, Odnoklassniki, and Facebook in the Baltics. The researchers employ such diverse methods as applied linguistics tools, structural network analysis, elements of expert systems theory, trend and opinion analysis, neural networks, machine learning, and data visualisation. They identified profiles and public Russian-language posts or comments from areas with large Russian-speaking populations: the cities of Tallinn, Maardu, and Paldiski, as well as Ida-Viru county in Estonia; the cities of Rīga, Jūrmala, Jelgava, and Liepāja, as well as the Latgale region in Latvia; the cities of Vilnius, Visaginas, Klaipėda, Kaunas, and Šiauliai, as well as Šalčininkai and Vilnius districts in Lithuania. ■

Prepared by Dr. Rolf Fredheim 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](http://www.stratcomcoe.org) | [@stratcomcoe](https://twitter.com/stratcomcoe) | [info@stratcomcoe.org](mailto:info@stratcomcoe.org)