Who Do You Think You're Fooling?

The Internal Russian Disinformation Campaign

Sean Norton

University of North Carolina at Chapel Hill

Motivation

The well-known 2016 Russian disinformation campaign targeted at the US presidential elections, was proceeded by a much larger (in terms of tweets and users) Russian-language Twitter campaign, which peaked during the Euromaidan Revolution and subsequent annexation of Crimea in spring 2014. While previous work on authoritarianism has established the utility of social media for distracting citizens and increasing the costs of finding undesirable information, this seems unlikely in the context of post-Crimea Russia. The annexation of Crimea brought approval of Vladimir Putin and the Russian regime to all-time highs as the regime purposefully activated feelings of Russian national pride and turned them against the new Ukrainian government, the EU, and the US. Distraction seems counter-productive at best - so what was the Russian regime doing?

The Russian Language Campaign
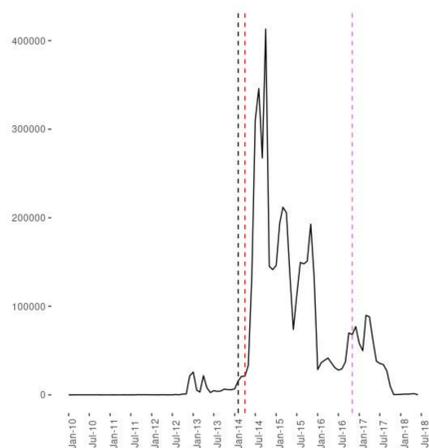
Dashed lines (left to right): annexation of Crimea (Feb. 2014), imposition of sanctions (Mar. 2014), US presidential election (Nov. 2016)

What Were They Talking About?

Based on coherence scores and human verification, I fit a 15 topic model to the corpus of $\sim$ 4.9 million Russian language tweets from the Twitter Election Integrity Dataset. I present the 5 most substantively interesting topics here.

(a) **Pushing the news**: "Thoughtfully written. The Russian people are sincerely worried about the situation in Ukraine."; 12% of total tweets

(b) **Putin and Patriotism**: "Don't get it confused: the government is the government, but the state is each of us!"; 8% of total tweets

(c) **"Personal" Opinions**: "Lie, lie, and lie again! That's the motto of (Ukrainian Minister of Internal Affairs) Avakov!"; 6% of total tweets

(d) **America**: "Why is Russia feeding the US?" and "Trump instead of Clinton! What a week!"; 9% of total tweets

(e) **Russia vs. the EU**: "Russia hasn't even felt a thing yet, but the EU itself can't withstand the effect of its own sanctions!"; 7% of total tweets
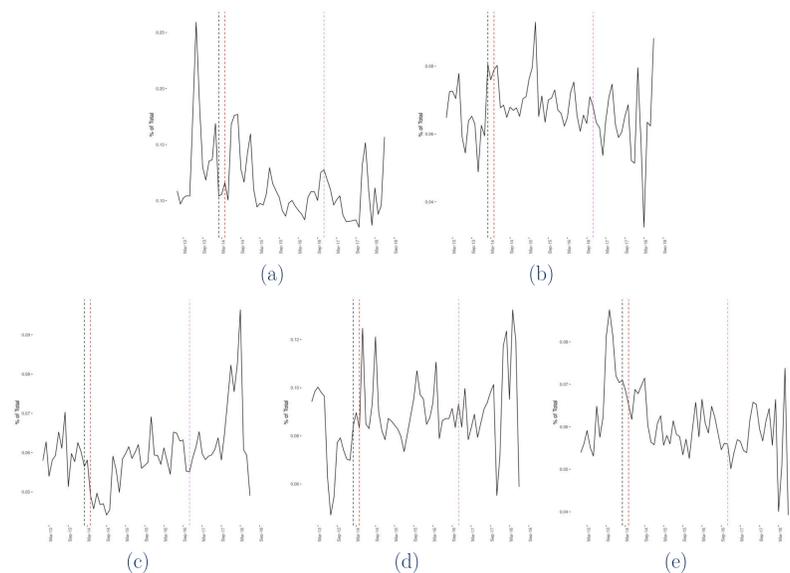
How Were They Talking About It?

Sentiment analysis using CNN pre-trained on large corpus of Russian twitter replies, coded as positive or negative.
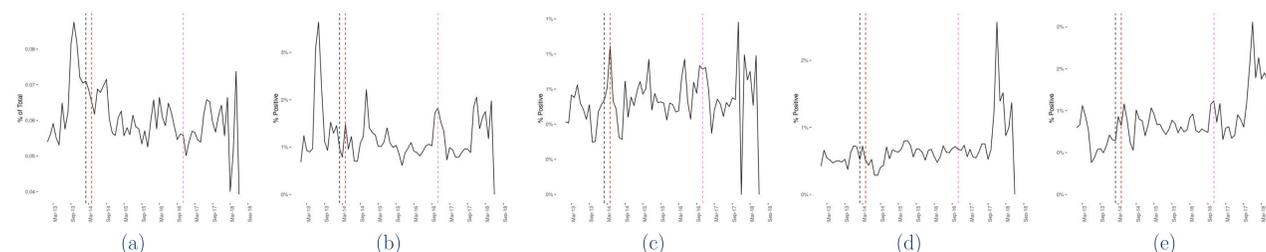
Distracting or Participating?

- Russian propaganda campaign reacting to events on the ground, boosting preferred narrative
- Acting out a "rally around the flag" effect
- Riding the wave of the Russian Spring, not leading it; amplifying existing trends over creating their own
- Open question: is there a substantive difference between participation and distraction?

Did It Work?

The Russian Spring was an event driven by *emotion* - could social media amplify these emotions? Planned experiment:

- Better sentiment classification; train RuBERT model, include neutral category
- Identify high-interaction tweets on a single subject: sanctions
- Expose sample to tweets; measure feelings towards sanctions, Putin, Russia, etc. before and after
- Identify features - across both participants and texts - associated with change in feeling

The Whole Picture

Defining feature of modern social media is how *visual* it is - $\sim$ 500 GB of images in dataset

- Humans are hard-wired to engage emotionally with images - will test tweets with/without included images in experiment
- What *type* of images are most effective? Memes? Journalistic? Amateur?
- Can we recover image features that increase emotional engagement?

Contact Details

Email: stnorton@live.unc.edu

Personal Website: https://www.seantnorton.net