Dynamic Topic Modelling: A Case Study on Russian Twitter Troll Data

Chunxia Cao, Bo Chai, Nicholas Chao, Johnny Chen, Liangkai Hu,
Shaobo Liang, Stephen Newns

Introduction

We have developed a visual tool to analyze Twitter troll dataset released by FiveThirtyEight implicating Russia of interfering with the 2016 US Presidential election, allowing users to uncover evolutions in topics and sentiment in tweets over time and to explore how discussion varies amongst different groups.

Data

We used only English tweets from Jan. 1, 2016 through Jan. 1, 2017, and after removing spammed tweets, tweets that were only links, we ending up with about 182,327 tweets for modeling.

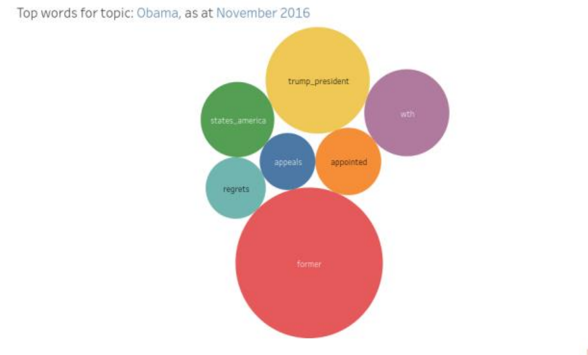
Approach

Dynamic topic modeling was used to better understand how trolls strategize misinformation campaigns on Twitter, as topic modeling analysis has often been static in the past.

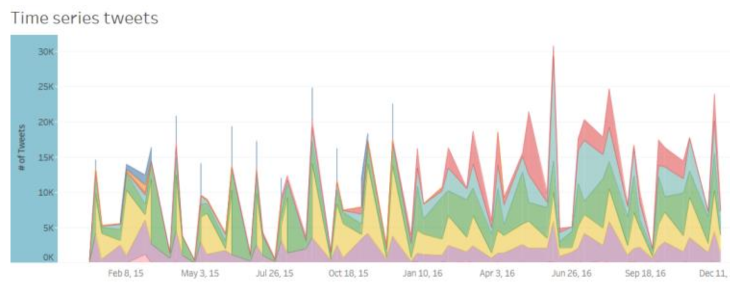
A Structural Topic Model (STM) was used to provide a temporal view of topic modeling, incorporating covariate effects from document-level metadata on a document's dirichlet prior, in terms of topic-content.

Having many documents makes determining the right number of topics non-trivial, as a bottleneck is created through the iterative process of topic modeling. This was managed with a t-distributed neighbor embedding (t-SNE) of a word co-occurrence matrix to find anchor words as starting points in expectation-maximization optimization of a structural topic model.

After tokenizing documents and converting and filtering words, we had 12,942 tokens and a 182,327x12,942 document-feature matrix for our model's dataset.



The word 'former' has a strong covariate effect on the Obama topic, corresponding to Trump's victory signaling the era



Dates with large spikes in Tweet volume can be further investigated to discover topics

Results

While dynamic effects were often difficult to comprehend, some meaningful effects are able to be found, and word-cloud representations of the most probable words drawn from a topic given a selected time period is an effective way of exploring dynamic topic models.

Further analysis should explore the possibility of incorporating time into topic and word distributions in a way that such distributions are dependent on those of the previous time frame.

Our approach simply considered publish month as a categorical factor, so adjacencies between time slices are not taken into account in topic prevalence and content, which might be important to capture if topic prevalence or content tends to grow in a sequentially dynamic way.