Textual and Statistical Analysis of Russian IRA Facebook Posts

*The paper is written in the scope of a student-faculty collaborative summer research with professor Richard K. Merritt.

David Oniani

Luther College

oniada01@luther.edu

Month Day, Year

Abstract

The 2016 United States Presidential Election was targeted by an unprecedented intelligence and influence campaign. Arising out of Russian so-called Internet Research Agency (IRA), it sought to sow discord, attack the fissures of the United States, and ultimately, sway the election results. [1] [2] Recently, some of the IRA-backed Facebook advertisements were released by The United States House Permanent Select Committee on Intelligence. All of the advertisements are in the PDF format. We have scraped the PDF files and present the results obtained by both textual and statistical analyses of the above-mentioned data. Authorship attribution and sentiment analysis were also performed. [1] [4] In addition, we have made the data publicly available for other researchers and/or interested people in a much nicer and easier-to-manipulate CSV format.

[1]Please note that this paper does not discuss political implications of the campaign, but attempts to explore the methods of persuasion that were employed in this influence campaign.

Table of Contents

Data and Preparation

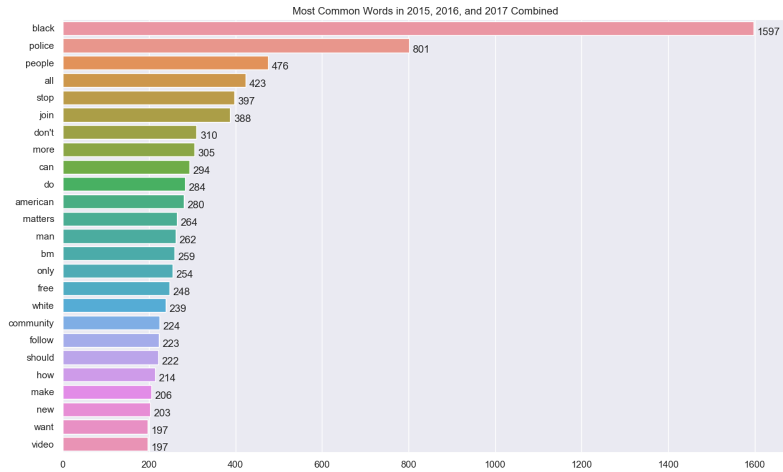
The dataset was scraped from [2] more than 3,500 Russian IRA Facebook posts made publicly available in the PDF format by the House Intelligence Committee. We used the free and open-source Python library [5] `pdftotext` to scrape the data. Many CSV files were formatted in a way that it was virtually impossible for `pdftotext` to parse them correctly. Due to this reason, we have manually reviewed most of the CSV files for validity. [4] All the CSV files have been made publicly available. It is important to note that the dataset is just a sample of a bigger dataset and, albeit LESS LIKELY, might not be a good representatation of the overall campaign.

Common Words

Among the strategies utilized by this influence campaign, one that stood out the most was the exploitation of internal issues of the United States by realizing social, political, and historical backgrounds of the country. This was achieved primarily by using words and phrases strongly linked to the above-mentioned issues.

[2] Below is the barchart showing top 25 most commonly used words after eliminating linking verbs, prepositions, pronouns, and some other (non-relevant) words. Three most commonly used words in the campaign were black, police, and people.

[2] The list of all eliminated words is provided at
`https://github.com/oniani/ira-analysis/blob/master/eliminated-words.txt`.



Most common words in 2015, 2016, and 2017 combined.

It is apparent that the vast majority of the Facebook posts featured words and phrases related to race and race-based discrimination. Words related to "taking action" such as "join", "stop", "do", "follow" etc. were used extensively for persuasive purposes. These words in combination with other words like "american", "matters", "community" etc. further increased the persuasive quality of the advertisement.

Sentiment Analysis

For the sentiment analysis purposes, we used `python`'s `TextBlob` library. Suprisingly, out of all Facebook posts with `Ad Text`, 1643 were positive, 900 negative, and 933 neutral leaving the overall tone of the posts positive. Yet, the statistical significance of this claim is rather questionable as the polarity levels were always near zero. Such low polarity level, however, demonstrates a highly intelligent design of posts.

As the `Ad Texts` did not give us any strong proofs, we looked at the negativity levels across all three years and found a consistent trend.

Subjectivity	Year 2015	Year 2016	Year 2017	2016 (%)
1.0	154	562	176	63.004
0.75	150	547	169	63.164
0.5	123	405	120	62.500
0.25	15	81	22	68.644
0.15	5	33	10	68.750
0.1	2	11	6	57.895
0.05	1	5	2	62.500

Analyzing Negativity

Notice that for any given year and at any subjectivity level, year 2016 is consistently comprising around 60% of all posts. The year of the Presidential Election was rather negative.

Authorship Attribution

Since all of these posts were issued by the same organization/entity, it was interesting to see if there are some common patterns between the Facebook ads of 2015, 2016, 2017. For this exact reason, we have performed authorship attribution tests by implementing [3] the paper by Koppel et. al.

We have performed 3 authorship attribution tests:

1. Assuming that the author of the Facebook posts of 2016 and 2017 was the same and checking accuracy for the author of 2015.
2. Assuming that the author of the Facebook posts of 2015 and 2017 was the same and checking accuracy for the author of 2016.
3. Assuming that the author of the Facebook posts of 2015 and 2016 was the same and checking accuracy for the author of 2017.

In all three cases, we had to merge some of the data to achieve the required minimum text length of 500 words. [3] This was done using randomization to avoid bias. [4] Results are available in the form of JSON files formatted in the manner shown below.

```
{
  "answers": [
    {
      "unknown_text": "2015-unknown1.txt",
      "author": "candidate2016",
      "score": 0.58
    },
    ...
    {
      "unknown_text": "2015-unknown56.txt",
      "author": "candidate2016",
      "score": 0.76
    }
  ]
}
```

JSON example

The first answer tells us that the unknown text was written in year 2015, and that there is 58% chance that it was written by the author of Facebook posts of 2016. The last answer shows that there is 0.76% chance that the given text was authored by the entity who wrote the posts in year 2016.

[3] Note that because of randomization, texts were somewhat scrambled and are not available in any format. That said, one can easily redo the authorship attribution with similar accuracy using the algorithm which already implemented and resides in the GitHub repository, https://github.com/oniani/ira-analysis/tree/master/koppel11.

[4] Results JSON: https://github.com/oniani/ira-analysis/tree/master/koppel11/results

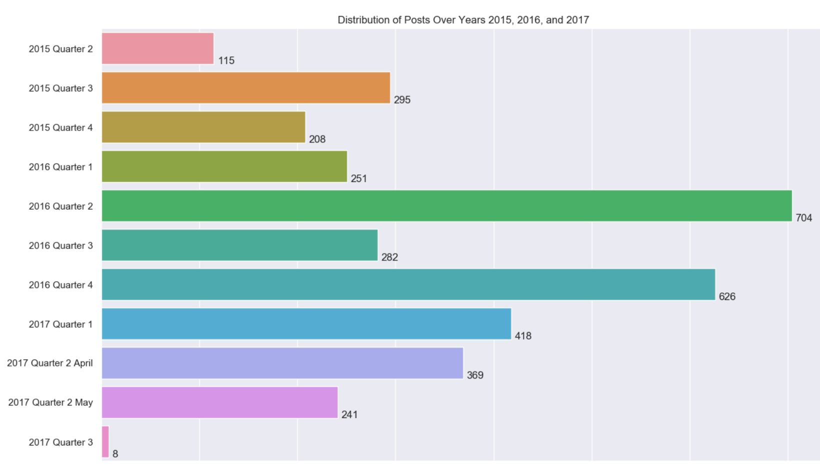
Below are the results obtained after performing all three above-mentioned authorship attribution tasks.

Years	Similarity (%)
2016 and 2017	72.509 (similarity to 2015)
2015 and 2017	68.516 (similarity to 2016)
2015 and 2016	62.392 (similarity to 2017)
Average (2015, 2016, and 2017)	67.806%

Average Similarity

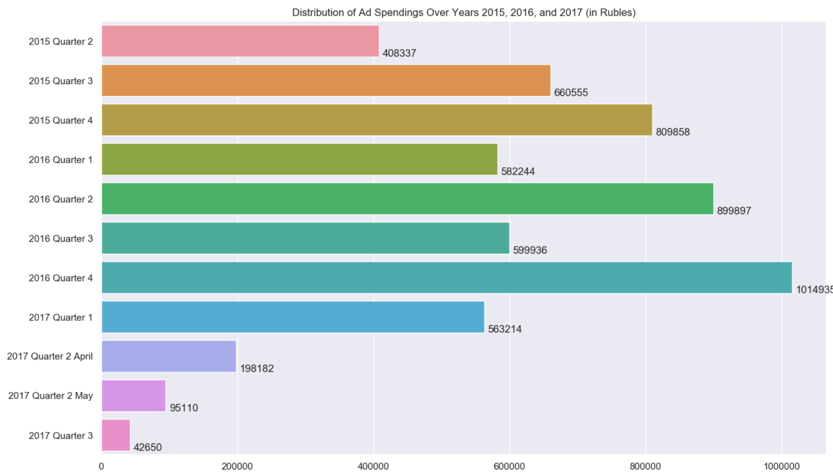
General Statistics

The distribution of posts over all three years shows a bimodal distribution with two peaks in quarters 2 and 4 of year 2016. Given the fact that the US Presidential Election was held in the fourth quarter of 2016, it is surprising that the number of posts in the second quarter exceeded that of the fourth quarter.



Distribution of Posts Over Years 2015, 2016, and 2017.

As for the ad spendings, the fourth quarter of 2016 exceeds that of any quarter, with second quarter coming next. Interestingly, most of ads were paid in the Russian currency (ruble) with two exceptions (4 posts in total) in 2016 quarter 3 and 2017 quarter 1 when IRA spent $74.000 and $35.330 respectively.



Distribution of Ad Spendings Over Years 2015, 2016, and 2017 (in rubles).

Furthermore, 99.8% of all paid ads across all years were paid in rubles. Below is the chart showing the number of posts based on a currency.

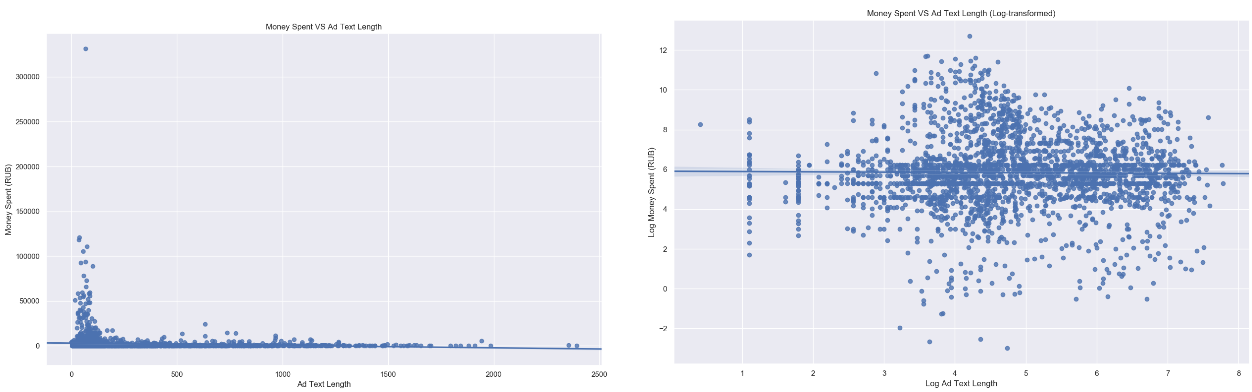
Currency	Total (All Years)
RUB	2549
USD	5
[5] None	787
0	176

As the information for the reader, the Russian ruble is used only in Russia, Belarus, and two regions of Georgia, which are considered by Russia as partially recognized states of Abkhazia and South Ossetia.

Regression Analysis

We would like to determine if there is a statistically significant relationship between two quantitative variables: the textual length of advertisement and the money paid for it. Determining the relationship between these variables will give us an idea if there was some kind of priority attached to posts, depending on its textual length.

For this task, we use a linear regression approach. Note that we perform the test only on the ads paid in rubles, the primary reason being not having enough data points for USD (only 5 values).



(a) Money Spent VS Ad Spend Length

(b) Log Money Spent VS Log Ad Spend Length

Simple VS Log-transformed

The first plot showed no strong relationship between the response and the predictor. That said, the shape of the plot hinted us to apply the log transformation. The log transformation, once again, verified our assumption of no relationship. The numerical data shown below further reinforces our observations.

RUB (Without Log Transform)	
Slope	-2.594
Intercept	2980.780
R-squared	0.007
P-value	$3.680 * 10^{-5}$
Standard Error	0.627

RUB (With Log Transform)	
Slope	-0.015
Intercept	5.904
R-squared	0.0001
P-value	$0.594 * 10^{-5}$
Standard Error	0.028

From the chart above, we see the R-squared value of $0.007, which tells us that only 0.7% of variation in the money paid for advertisements is explained by the number of words. Standard error is 0.627 which is very small compared to the intercept whose value is 2980.780 (RUB). This tells us that the error in our estimates is really small. The p-value equals $3.680 * 10^{-5}$ which is a lot smaller than 0.05 and makes the conclusion statistically significant. Therefore, we accept the null and reject the alternative conclude that there is no statistically relationship between the amount of money paid and the number of words in the advertisement.

Even after performing a log-transform, the results show no evidence of a strong or even a weak relationship between the variables. The p-value tells us that the results are statistically significant.

Although having no relationship is usually not a desired result, in our case, we can conclude that posts were designed in a rather intelligent manner, with no noticeable priorities depending on the number of words in the text section.

Conclusions

We have scraped PDF files made publicly available by The United States House Permanent Select Committee on Intelligence and performed a number of different analyses including textual, linguistic, and pure-statistical. The results obtained showed a clear attempt of influencing people.

References

[1] The United States House Permanent Select Committee on Intelligence. *Exposing Russia's Effort to Sow Discord Online: The Internet Research Agency and Advertisements*. 2018. URL: https://intelligence.house.gov/social-media-content/ (visited on 08/27/2019).

[2] The United States House Permanent Select Committee on Intelligence. *Social Media Advertisements*. 2018. URL: https://intelligence.house.gov/social-media-content/social-media-advertisements.htm (visited on 08/27/2019).

[3] Moshe Koppel, Jonathan Schler, and Shlomo Argamon. "Authorship attribution in the wild". In: *Language Resources and Evaluation* 45 (Mar. 2011), pp. 83–94. DOI: 10.1007/s10579-009-9111-2.

[4] David Oniani and Richard Merritt. *CSV data scraped from the PDF files of IRA Facebook posts*. 2019. URL: https://github.com/oniani/ira-analysis/tree/master/data/csv (visited on 08/27/2019).

[5] Jason Alan Palmer. *pdftotext*. 2018. URL: https://pypi.org/project/pdftotext/ (visited on 08/27/2019).