Investigating Russian Influence on the 2016 Presidential Election: A study of Internet Research Agency Tweets

Pedro Freitas
Francesco Gandolfi
I Vivek Kai-wen

Introduction

Foreign influence in the domestic political contexts of other countries is not a new phenomenon. Elections represent part of the command-and-control system for the decision-making of a nation and influencing an election can alter domestic decisions to align with a foreign nation interest.[1] States have been seeking to influence each other for centuries and what is new in the modern context are the methods. If in the 19th-century tactics to influence an election included physical force, nowadays it is used coercive influence and manipulation, including through disinformation[2] , where social media strategies may be used to achieve a political purpose.

The modernization of tactics to enforce political influence occurs from the advent of the internet and social media platforms such as Facebook and Twitter. They have changed the way people communicate about politics and access information in general, creating new opportunities for a wide range of political actors, including foreign actors. The social networks facilitate the spread of disinformation and content to be propagated among users with no significant third party filtering, fact-checking, or editorial judgment. A user with no track record or reputation can in some cases reach as many readers as Fox News, CNN, or the New York Times.[3]

Even though the use of social media to exert foreign influence gained public attention after the US 2016 elections, over the last years it has been documented the use of social media to influence politics in a range of countries by promoting propaganda and advocating controversial viewpoints[4]. There is evidence of cyber attacks and computer-driven propaganda in countries like Argentina, Brazil, France, Germany, Philippines, the United Kingdom (UK), and many others.[5]

The Internet Research Agency (IRA) is a Russia-based company focused on media and information propagation that has been accounted for executing influence campaigns over the 2016 US elections. According to the U.S. Department of Justice indictment, the campaigns initiated in 2014 with the goal of interfering with the U.S. political system, largely through social media.

They made use of social media accounts masqueraded as American citizens to divide voters over a range of issues. These accounts were accused of posting derogatory information about several

---

[1] Baines, P., & Jones, N. (2018). Influence and Interference in Foreign Elections: The Evolution of its Practice. The RUSI Journal, 163(1), 12-19.
[2] Baines, P., & Jones, N. (2018). Influence and Interference in Foreign Elections: The Evolution of its Practice. The RUSI Journal, 163(1), 12-19.
[3] Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. Journal of economic perspectives, 31(2), 211-36.
[4] Martin and Shapiro (2019) Trends in Online Foreign Influence Efforts. https://scholar.princeton.edu/jns/research-reports
[5] Tenove, C., Buffie, J., McKay, S., & Moscrop, D. (2018). Digital threats to democratic elections: How foreign actors use digital techniques to undermine democracy.

candidates, and by early to mid-2016, their campaigns included supporting the candidate Donald Trump and defaming Hillary Clinton.

To fight against these malicious activities, Twitter has launched a data archive of potential state-backed operations identified on the network. This disclosure has empowered academic research on the topic and helped to better understand how those efforts occur. The goal of this paper is to analyze Twitter activity for IRA associated accounts during the years of 2015 and 2018 and grasp a better understanding of how those attacks are coordinated.

Related Literature

Foreign Influence Efforts

The Digital Threats To Democratic Elections[6] report analyzes the influence of foreign actors' use of digital techniques to interfere in democratic elections. According to the report, foreign actors do so to undermine three critical elements of democratic elections: (i) fair opportunities for citizen participation (such as voting, running for office, or contributing to public debates); (ii) public deliberation that enables citizens to share and understand each other's insights and perspectives; and (iii) key institutional actions by electoral commissions, political parties, and other organizations, including the enforcement of electoral regulations.

The methods used include four main techniques: (i) hacking attacks target systems, accounts and databases, to access, change or leak private information; (ii) mass misinformation and propaganda campaigns promote false, biased and provocative messages; (iii) develop messages for micro-targeted manipulation; and (iv) online "trolling" operations to intimidate, discredit, and harass individuals or groups.

The element of public deliberation, for example, can be harmed through misinformation practices, weakening the possibility for people to have good discussions about what is true or false, acceptable or inappropriate. Foreign actors can also use polarizing and degrading messages to corrode norms of inclusivity and respect. Through these and other means, processes by which citizens come to understand their shared problems and to pursue legitimate collective solutions are threatened.

Martin and Shapiro have cataloged a few of foreign interference campaigns, which they call Foreign Influence Efforts (FIE). For a campaign to be considered an FIE they analyzed if there was evidence of (i) action by one state against another in the media; (ii) an identifiable political goal; and (iii) content that is meant to appear as being produced organically in the target state.

Those campaigns possess different strategies depending on the objective. The main strategies defined by the authors are (i) defamation, which is an attempt to harm the reputation of people or institutions; (ii) persuasion, which is the attempt to move the average citizen to one side of the issue; (iii) polarization, defined as trying to move opinion to the extremes on one or more issues; (iv) shift political agenda, which is when the efforts add something new in the political agenda; and (v) undermine institutions, which is an effort to reduce the credibility/reputation of one or more institutions in the target country.

[6] Tenove, C., Buffie, J., McKay, S., & Moscrop, D. (2018). Digital threats to democratic elections: How foreign actors use digital techniques to undermine democracy.

Defamation

Defamation, defined as a direct attack on a person with the objective of harming their reputation, was the most widely used strategy in Foreign Influence Efforts (FIEs) according to Martin and Shapiro.[7] In their study, defamation was used in 65% of FIEs, while persuasion was used in 55% of FIEs and polarisation was only used in 15% of FIEs. Therefore, defamation campaigns should be the strongest focus of governments and agencies involved in the fight against FIEs. Bradshaw and Howard argue that defamation is also used frequently in domestic political influence efforts to mount smear campaigns against the opposing political party during election campaigns or to discredit dissidents and critics of the government in normal times.[8]

Many defamation campaigns pursue their goals by creating fake content, much like other forms of political influence campaigns. They create videos, news articles and blogs with false information and spread this information through social media, either through their own posts or by interacting with posts from real accounts. Hundreds of new websites that publish fake content have been set up in recent years. In fact, during the 2016 US Presidential Election campaign, fake news stories about Hilary Clinton ordering the murder of an FBI agent and participating in a child abuse ring were "shared hundreds of thousands of times on social media".[9] Nonetheless, defamation campaigns usually do not exclusively use fake news to achieve their objectives; rather, they mix false information with real news that portrays their targets negatively. This not only allows trolls to gain some credibility, but also makes it even harder for social media operators to identify and stop their activities.

The distinguishing action of defamation campaigns is the use of "hate speech or various forms of online harassment".[10] In contrast to other techniques, hate speech and harassment do not primarily aim to convince people of a certain opinion. Instead, they attempt to tap into pre-existing dislike or doubts and fuel the growth of these negative thoughts and feelings.

The Data

We will analyze a set of tweets that were identified by Twitter as part of a FIE in the United States, including during the 2016 United States Presidential Elections. It was later posted publicly by FiveThirtyEight, a site that publishes data-oriented articles. The tweets had been posted on accounts linked to the Internet Research Agency (IRA), a Russian company that carries out online influence operations for public or private interests. According to intelligence agencies in the United States, the IRA was heavily involved in influence campaigns during the 2016 Presidential Elections, specifically with the aim of improving the election prospects for Donald Trump. With this in mind, we focused on the 2016 elections in our analysis.

News reports, academic papers and Congressional reports indeed make the argument that the tweets by the IRA on the whole supported the Trump campaign. On the defamation front, its tweets

---

[7] Martin and Shapiro (2019) Trends in Online Foreign Influence Efforts. https://scholar.princeton.edu/jns/research-reports

[8] Bradshaw, S. & Howard, P. N. (2018), 'Challenging truth and trust: A global inventory of organized social media manipulation', The Computational Propaganda Project.

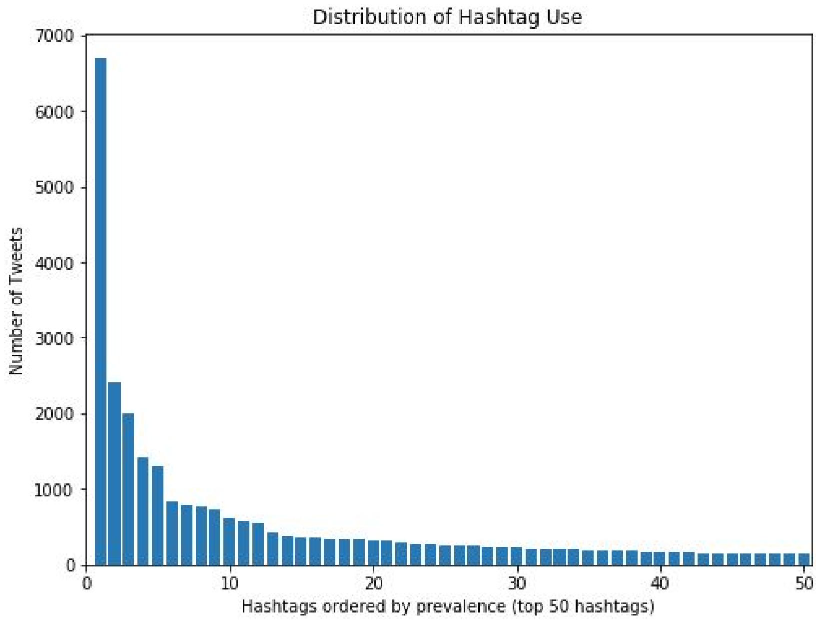
[9] Hindman, M. & Barash, V. (2018), 'Disinformation, and influence campaigns on twitter'.

[10] Bradshaw, S. & Howard, P. N. (2018), 'Challenging truth and trust: A global inventory of organized social media manipulation', The Computational Propaganda Project.

attempted to discredit the Democratic presidential candidate Hilary Clinton, former President Barack Obama and the Democratic Party as a whole. The knowledge of this context would drive our efforts to label part of the data, so that we could make use of supervised learning techniques to detect defamatory content.

Data Exploration and Summary Statistics

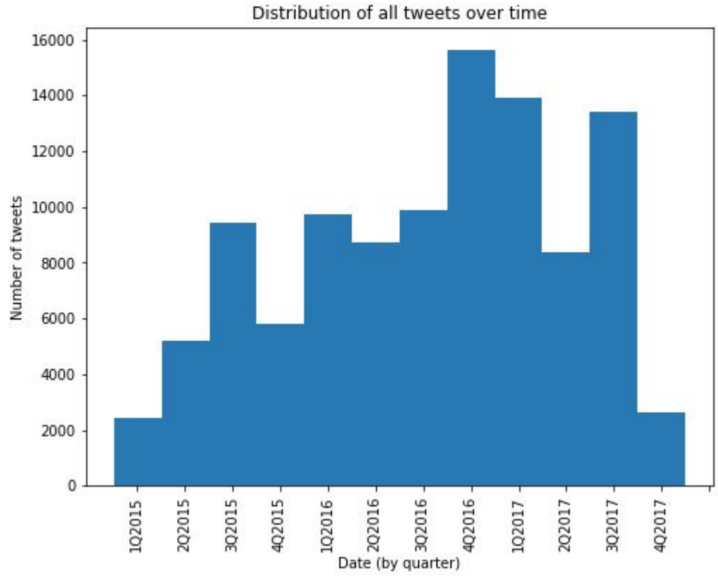
One of the major indicators of the content of a tweet is the hashtags used. This served as a starting point for our analysis of the dataset, so we first looked into the distribution of hashtags in the dataset and the top hashtags.



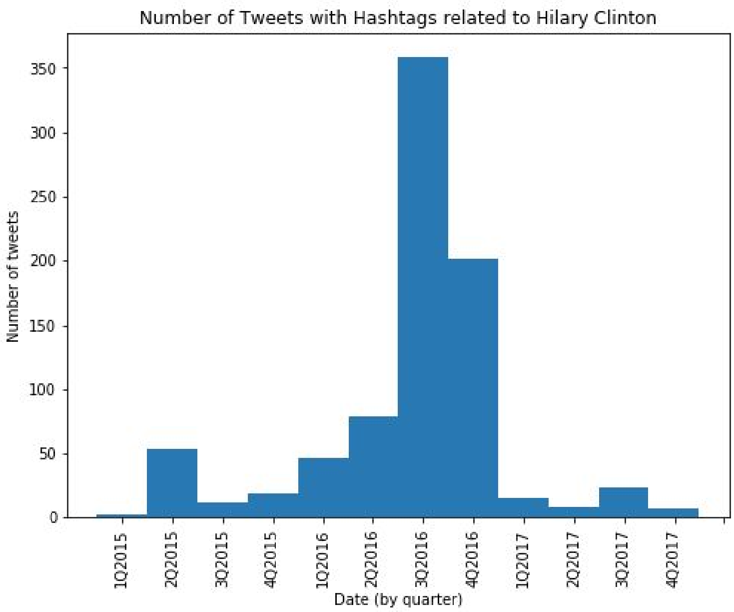
hashtag	density
#news	0.091467
#sports	0.032805
#politics	0.027431
#world	0.019431
#local	0.017954
#maga	0.011541
#tcot	0.010885
#topnews	0.010406
#blacklivesmatter	0.009941
#pjnet	0.008478

We noted from the histogram that the distribution of hashtag use followed the Power Law, meaning that the hashtag use decreased exponentially going from more to less commonly used hashtags. We can also see this in the table, which shows that top 8 hashtags have a density of between 1% and 10%, while all the remaining hashtags have a density lower than 1%. Furthermore, the top 5 hashtags actually were not politically charged, but instead involved neutral terms like news and sports that did not point to a specific political movement. Nonetheless, we did not assume that the tweets associated with these hashtags did not have a political slant, since they could still contain selected news or fake news that could be used to defame politicians.

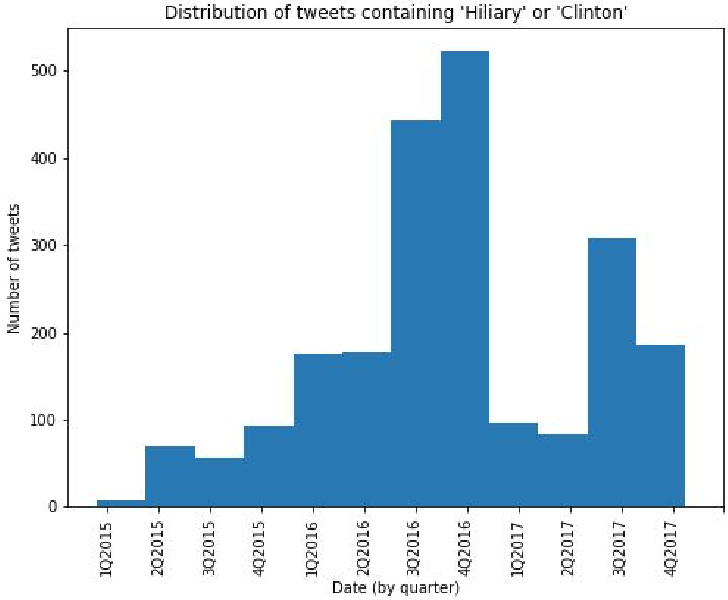
The next key dimension to look at was the distribution of tweets over time. Since we decided to study the IRA campaign during the 2016 Presidential Election, we looked at the distribution of tweets over time from 2015 to 2017. As one would expect, we see a general rise in the number of tweets by IRA-linked accounts from the start of 2015 to the end of 2016 when the election was held. However, the IRA accounts did not suddenly die off after that; rather, they continued to be active at certain times, indicating further campaigns after the election.



Next, we went on to look at tweets with hashtags related to Hilary Clinton and how the number of such tweets evolved over time. Again, we see a build-up in the number of tweets to the election in late 2016, indicating that the IRA was ramping up its efforts as it got closer to the election. After the election, this number of tweets fell dramatically, which we expect to happen, given that the IRA had already achieved its goals and no longer needed to defame Hillary Clinton.



Following this, we looked at the distribution of tweets containing either 'Hillary' or 'Clinton'. Once again, the number of tweets rose rapidly and peaked around the time of the 2016 election. After that, the number dropped significantly. However, in contrast to the previous plot, the number of tweets in this case rose again in the middle of 2017. This difference probably means that Hillary Clinton was mentioned, but was no longer the focus of the IRA campaign in 2017.



Our Methods

First Approach to the Problem

Our first approach involved sentiment analysis of the tweets, followed by a linear regression of sentiment on hashtags. First, we extracted the top 1000 hashtags by the number of tweets associated with the hashtags. We performed a one-hot encoding of these hashtags. For tweets associated with the top 1000 hashtags, we used the Python package *textblob* to conduct a basic sentiment analysis to learn if the tweets were positive, negative or neutral. Each tweet was evaluated with a value on a scale from -1 to 1, where 0 indicated a neutral sentiment, -1 indicated the highest possible level of negativity and 1 indicated the highest possible level of positivity. We then regressed the sentiment value on the one-hot encodings of the hashtags. Through this process, we hoped to get a sense of the most negative tweets and to check if these tweets were linked to defamation efforts of the IRA.

Surprisingly, hashtags with the smallest coefficients were not directly related to the 2016 Presidential Election. Instead, many of these hashtags were related to key issues in the general public discourse. For instance, two of the most negative hashtags were '#amber' and '#amberheard', which were references to the celebrity Amber Heard that allegedly suffered domestic abuse. As a result, this approach did not work well.

Nonetheless, this gave us some insight into the IRA tweets, especially the fact that these trolls were more sophisticated than we had expected. Rather than just focusing on the elections and posting content to defame politicians, these trolls were engaged in social discourse even in issues not directly linked to the election. This would make trolls harder to detect and more credible to other Twitter users that may read and share their content. This observation also hints at a possible polarisation effort by the IRA as well. By sharing views related to social issues and movements, these trolls might help to reinforce extreme views and persuade other users of those views. However, this was not our focus for this project and we did not pursue that line further.

Second Approach to the Problem

Our second approach involved labeling some of the tweets we had the confidence to relate to defamation campaigns and to train a supervised model on it to classify all documents in the corpus. We labeled the tweets based on hashtags. For the “True” label, the hashtags selected were the ones associated with defamation of candidate Hillary Clinton. Some neutral and positive hashtags were also picked to be part of the "False" label of our training dataset.

We tried two different representations for our corpus: bag-of-words and word2vec. For each representation, we experimented with different parameters. Some of the parameters we explored in the bag-of-words representation include bi-grams, inverse-document-frequency reweighting, and l2 normalization of the vectors. The parameters we tested for the word2vec included: cbow/skip-grams, size of the vectors, and negative sampling.

Later, we trained a standard Logistic Regression model and scored the whole corpus. We thought initially that the word embedding representation would yield better results because we could extract information also from the unlabelled tweets, however, the model with the bag-of-words representation seemed to get better results on the test set.

Even though the model scored well in our labeled test set, it did not pass the visual-inspection test when used to label the entire corpus. A lot of the tweets classified as defamation were not considered defamation tweets in our interpretation. Therefore, we ultimately chose to use a different model that would perform better.

Final Approach: Topic Modelling on Clinton-related Tweets

From our first attempts, it is evident that studying defamation on the entire IRA corpus is a challenging task, and would require a more complex approach or a significant effort in labeling the data. To streamline the analysis, we decided to focus on the well-documented defamation campaign by the IRA against Hillary Clinton. The first step was to select the tweets related to Hillary Clinton. To do so, we created a list of hashtags about Hillary Clinton and selected:

- All the tweets containing Hillary and/or Clinton in the text
- All the tweets containing at least one hashtag from the list

For a total of 54,148 from 791 distinct accounts. The complete list of hashtags is provided in Appendix A.1. We then removed stop-words from the tweets, as well as the words ‘Hillary’ and ‘Clinton’.

To study measures of defamation, we extracted topics from this corpus using LDA and NMF (Non-Negative Matrix Factorization). With NMF, for k topics we approximate the $d*n$ document-term matrix X as W*H, where W is a $d*k$ matrix whose rows can be interpreted as the amount of each topic in a document , and H is the $k*n$ matrix of words-in-topics weights. H and W are calculated minimizing the frobenius norm between X and W*H. Studying the weights for each word in H we can give an interpretation of the topics. To simplify the analysis, we do not use a multi-topic representation of the tweets, but instead we label each tweet as being of the topic with the highest weight in the corresponding row in W. Although less common than LDA, the use of NMF for topic modelling is well established, and can result in more interpretable topics[11].

[11] See for example O’callaghan, Derek, et al. "An analysis of the coherence of descriptors in topic modeling." *Expert Systems with Applications* 42.13 (2015): 5645-5657.

In both models the most important hyperparameter is *k,* i.e. the number of topics. We fit LDA and NMF for a wide range of *k*, focusing on the interpretability of the results, and especially on identifying anti-Clinton topics that could be used as a proxy for the defamation FIE against Hillary Clinton.

On the basis of interpretability, the model selected is an NMF with 12 topics. We labelled these topics on the basis of the top words and by looking at representative tweets. Two were clearly anti-Clinton, and thus potential expressions of the defamation campaign. We label these topics as *Clinton emails*[12] and *DOJ plea deal*[13]. It is worth noting that *DOJ plea deal* assigns a large weight to the word "crooked". We believe that this topic is related to the plea deal offered to past Clinton aides in the investigation into Hillary Clinton's private email server. In addition, one topic is about the Trump vs Clinton presidential race[14], but is mostly associated with tweets that are negative toward Clinton, and is thus considered one of the defamatory topics.

On the other hand, two topics seemed to be in support of Clinton's presidential race, *Pro Hillary*[15] and *Vote for Hillary*[16]. Lastly, one topic focuses on the Democratic Party presidential primaries, and especially on Clinton versus Sanders[17], and, trusting Twitter's classification of IRA twitter accounts, can be seen as evidence of the IRA effort to divide the Democratic Party. The complete list of topics and their corresponding top words is provided in Appendix A.2.

Results

When analyzing the tweets from the IRA accounts, we first considered tweets that were only either pro-Hillary or anti-Hillary, for accounts with at least 25 tweets. We then evaluated the pro-Hillary or anti-Hillary level of each account using the proportion of tweets that fell into either category. Most of the accounts were on average anti-Hillary and accounts that were strongly biased were all anti-Hillary accounts. For the most biased accounts, anti-Hillary tweets reached up to 100% of the tweets. In contrast, the strongest pro-Hillary accounts posted pro-Hillary tweets less than 75% of the time.

Overall, there was certainly more anti-Hillary content, but there was less of this defamatory content than we had initially expected from an IRA campaign with the aim of defeating Hillary Clinton in the election. Moreover, many of the accounts were not polarized in any particular way, but instead had a significant share of both pro-Hillary and anti-Hillary content. This reflects two characteristics of the data. First, there were many IRA accounts that masquerade as news sites and therefore post some positive tweets and neutral content about Hillary Clinton and second, the IRA trolls were more sophisticated than we had expected and maintained some degree of balance in their tweets.

---

[12] Most common words: *emails, state, foundation, fbi, email, department, via, Clintons, dept, release, campaign, deleted, investigation, wikileaks, benghazi*

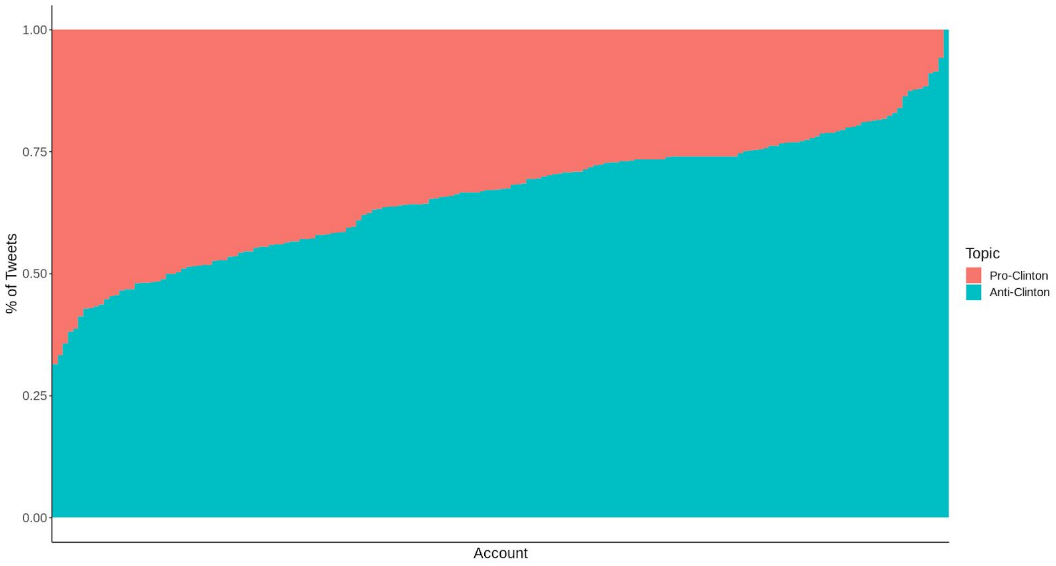
[13] Most common words: *breaking, doj, deal, plea, offers, expanding, probe, rumored, take, report, video, crooked, grounds, ample, prosecuting*

[14] *Trump v Clinton,* most common words: *trump, donald, poll, supporters, debate, russia, vs, says, voters, leads, lead, win, rally, wins, like*

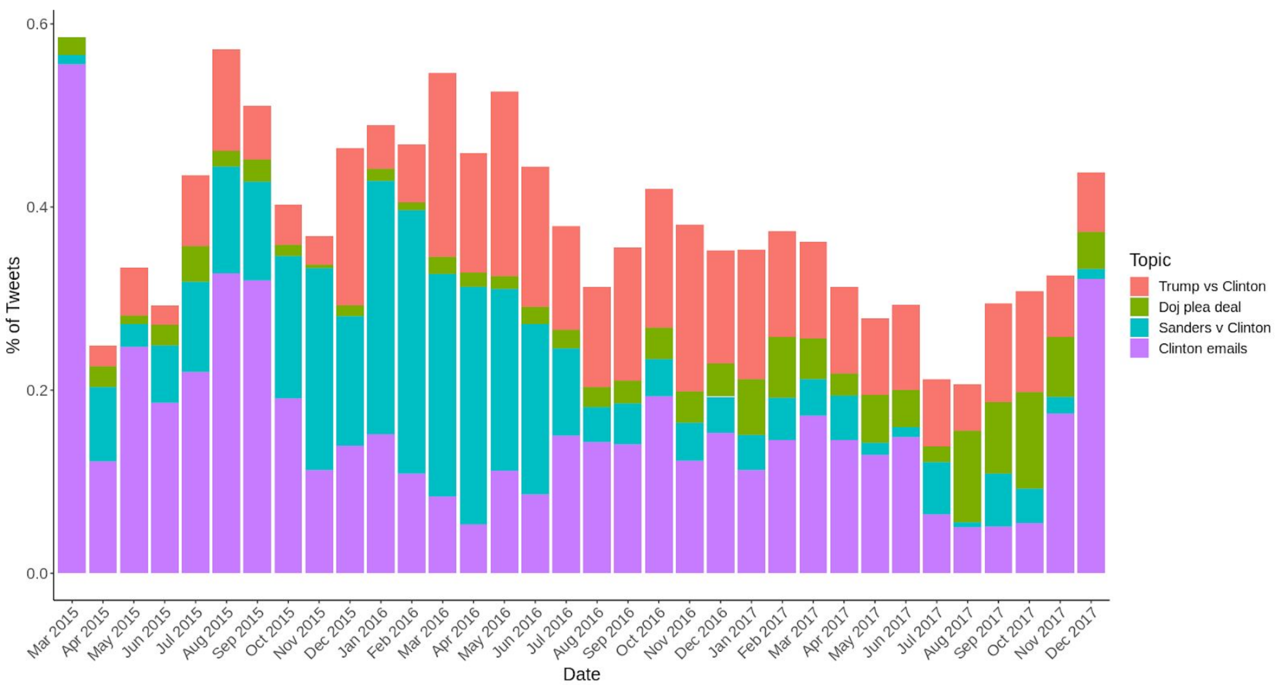
[15] Most common words: *president, qualified, would, never, next, one, want, run, like, become, endorses, becomes, woman, think, Obama*

[16] Most common words: *vote, women, people, didn't, black, voting, popular, would, want, i'm, never, get, please, said, America*

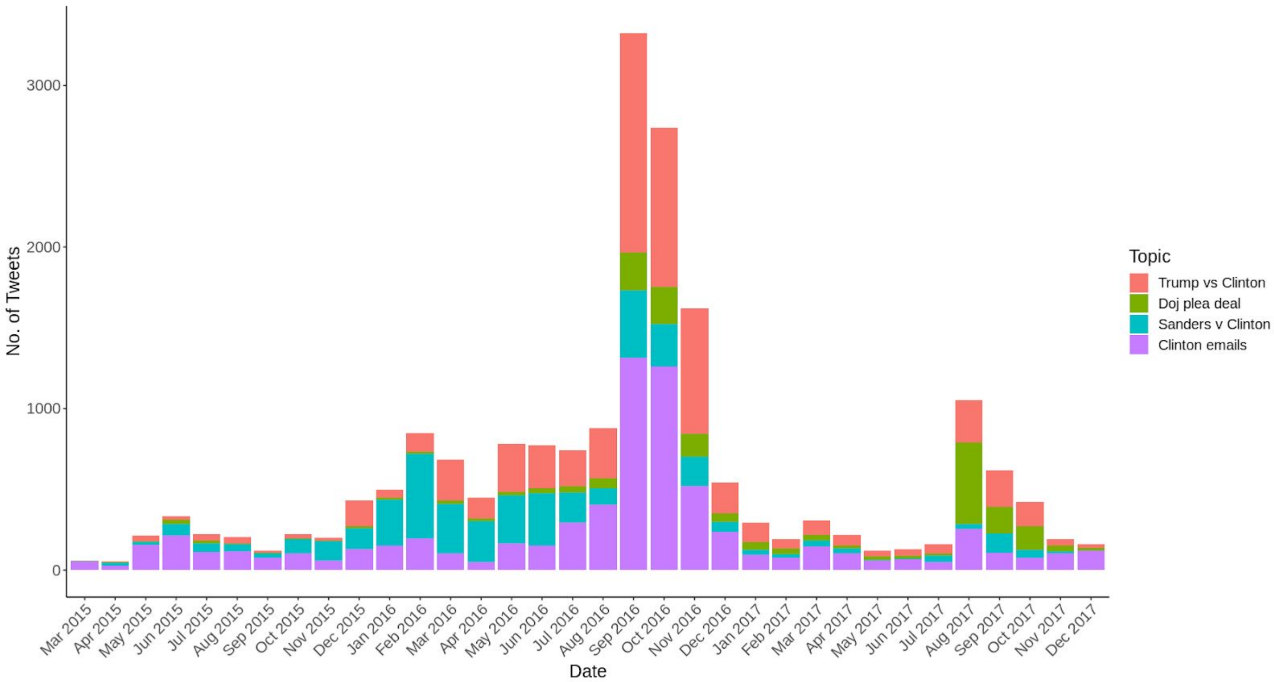
[17] Most common words: S*anders, Bernie, debate, Iowa, says, democratic, supporters, campaign, primary, race, voters, clash, California, fight, win*



Next, we look at the proportion of tweets from the IRA-linked accounts about Hillary Clinton that are defamatory in nature and how this changes over time. Note that the total proportion of defamatory tweets is always a large proportion of the total number of Hillary-related tweets. Moreover the topic of Clinton emails constitutes a large portion of tweets throughout most of the period from 2015 to 2017. Indeed, in reality, this represented a key issue for voters that voted against Hillary Clinton, indicating some level of success of the IRA campaign. Moreover, the Sanders v Clinton topic peaks in late 2015 to the first half of 2016, which also corresponds to reality, since this was the period of the Democratic primaries when Sanders and Clinton competed for the Democratic nomination. Similarly, the Trump vs Clinton topic picked up in early 2016 and lasted till the end of 2016, again corresponding to the period of the 2016 Presidential Election. On the other hand, we can infer that the DOJ plea deal was not as much of a key issue at the election. While the issue was constantly mentioned throughout 2016, it became a relatively larger part of IRA tweets only in 2017.



Following this, we looked at a related figure showing the total number of tweets by IRA-linked accounts that were classified as defamatory by our model. This plot confirmed many of our previous analyses about the levels of activity associated with the various topics and the overall change in the number of tweets over time. Clearly, the IRA accounts were most active in late 2016, which corresponds to the timing of the 2016 election. The number of tweets in all of the defamatory categories peaked at this time, demonstrating that the IRA indeed stepped up its campaign to influence the outcome of the election. After the election, the number of tweets fell sharply. There was again a smaller spike in late 2017, this time driven mainly by the DOJ plea deal. Again, this makes sense since the plea deal came under scrutiny during this period of time.



Conclusions

Overall, identifying and predicting defamatory tweets by the IRA-linked accounts is no mean feat, due to the high level of sophistication and added noise that these trolls have made use of. This certainly complicates the task of identifying foreign influence for a social network company like Twitter and especially for individuals that may interact with these trolls on a daily basis.

Having learnt this, we chose to focus on tweets related only to Hillary Clinton and the 2016 Presidential Election, which gave us a more manageable task. We were able to identify specific topics that IRA trolls were using in their defamatory campaign and analyze how these topics were changing over time. We could see the intensification of anti-Hillary tweets in the lead up to the election and how the issue of Clinton's private email server had a huge impact throughout the campaign period.

Our focus on defamation was motivated by the analyses in the literature that showed that defamation constitutes the most prevalent form of FIEs. From our analysis, we found that this is indeed true in the case of the IRA's influence operation in the 2016 Presidential Election. We saw that across the entire election year of 2016, defamatory tweets made up a large portion of all tweets related to the election and the number of these tweets grew rapidly close to the election.

Ultimately, the trolls are simply up to dragging people down. Despite hiding underneath the cover of anonymity and wielding the weapon of technology, they are in fact engaging in the oldest forms of electoral warfare - simple mudslinging and personal attacks that anyone could muster. The danger, though, lies in the scale and intent of its perpetrators, which threatens to bring the greatest countries to their knees.

Appendix

A.1 Complete list of Hillary Clinton-related hashtags

Pro-Clinton hashtags: #hillarystrong, #hillarysoqualified, #standwithmadampotus, #herstory, #bluewave, #imwithhillary, #hillafornia, #girls4hillary, #hrc2016, #iamwithher, #madamepresident, #connecttheleft, #momsdemandhillary, #imwither, #unitedagainsthate, #republicansforhillary, #hillaryclinton16, #voteblue, #turnncblue, #ilovehillary, #uniteblue, #werewithher, #imwiththem, #hillaryforamerica, #estoyconella, #hrcisournominee, #hillaryclintonforpresident, #hillyes, #hillaryforpresident, #whyimwithher, #readyforhillary, #hillaryaprovenleader, #strongertogether, #wearewithher, #hillstorm2016,bluewave2016, #heswithher, #vote4hillary, #clintonkaine2016, #bernwithher, #welovehillary, #imwithher2016, #imwithher, #gohillary, #hillarysopresidential, #itrusthillary, #ohhillyes, #itrusther, #republicans4hillary, #sheswithus, #voteblue2016, #madampresident, #hereiamwithher, #hillaryforpr, #votehillary, #yeswekaine, #hillary2016

Anti-Clinton hashtags: #hillary4prison, #hillaryliedpeopledied, #hillaryrottenclinton, #satanists4hillary, #deplorablehillary, #3wordhillary, #censoredforhillary, #releaseclintonsmedicalrecords, #madhillary, #hillaryforprison2016, #hypocritehillary, #imnotwithher, #whyimnotvotingforhillary, #lyinghillary, #criminalhillary, #crookedclinton, #heartlesshillary, #unforgivablehillary, #thingshillarywillneverhave, #hillarylostme, #moretrustedthanhillary, #crookedhilary, #sickhillary, #rejectedhillaryslogans, #queenofcorruption, #dropouthillary, #shelies, #crookedhiliary, #notwithher, #queenofcorrupton, #lyincrookedhillary, #neverclinton, #nevercrookedhillary, #wehatehillary, #clintoncrime, #lockherup, #democratliesmatter, #arrestclinton, #arresthillary, #impeachhillary, #corruptclintons, #indicthillary, #lyingcrookedhillary, #nomoreclintons, #deletehillary, #handcuffhillary, #hackinghillary, #libertynothillary, #nohillary2016, #zombiehillary, #anybodybutclinton, #corrupthillary, #sexoffendersforhillary, #crookedclintons, #thingsmoretrustedthanhillary, #hillaryforprison, #assad_clinton, #betternamesforhillarysbook, #saynotohillary, #billclintonrapist, #pedophilesforhillary, #thingshillarygoogles, #crookedhillaryclinton, #dontkillmehillary, #nohillary, #whatmakeshillaryshortcircuit, #clintoncrimefamily, #hillary4prison2016, #hillarylies, #makeamoviehillary, #theclintoncontamination, #neverhillary, #iwillneverstandwithher, #fbimwithher, #neverhilllary, #whatclintonwrites, #herpeshillary, #clintoncollapse, #stophillary2016, #hillno, #crookedhillary, #billclintonisrapist, #ohhillno, #clintoncrimefoundation, #clintoncorruption, #noclintonsever, #releasethetranscripts, #stopclinton, #abusivehillary, #lesshillarymorefun, #murderinghillary, #riskyhillary, #billclintonisarapist, #whathillarywouldforadollar, #hillary4jail, #killary, #hillary2jail, #stophillary, #lyinhillary, #heilhillary, #hillarysolympics, #hillaryliesmatter, #neverhilary, #losinglikehillary, #nevereverhillary

A.2 NMF Topics

Trump vs Clinton: trump, donald, poll, supporters, debate, russia, vs, says, voters, leads, lead, win, rally, wins, like
Fbi investigations: Obama, Comey, Lynch, FBI, calls, Mueller, congress, house, crimes, America, Russia, meeting, congressman, investigation, Abedin.
DOJ plea deal: breaking, doj, deal, plea, offers, expanding, probe, rumored, take, report, video, crooked, grounds, ample, prosecuting.
Hillary v Sanders: Sanders, Bernie, debate, Iowa, says, democratic, supporters, campaign, primary, race, voters, clash, California, fight, win.

Bill Clinton: Bill, remember, campaign, used, hate, flashback, symbol, made, fire, fury, comments, allwhite, club, joined, golf.
Vote for Hillary: vote, women, people, didn’t, black, voting, popular, would, want, i’m, never, get, please, said, America.
Clinton Emails: emails, state, foundation, FBI, email, department, via, clintons, dept, release, campaign, deleted, investigation, wikileaks, Benghazi.
Unclear Topic**:** rt, amp, debalwaystrump, https, Russia, America, realdonaldtrump, rr, Mueller, one think, republicans, nuclear, Benghazi, would.
Pro Hillary: President, qualified, would, never, next, one, want, run, like, become, endorses, becomes, woman, think, Obama.
Voting and polls: New, poll, York, Hampshire, excuse, shows, leads, go, voting, lol, lead, points, book, brand, fraud.
Foreign interference: election, us, going, interference, Ukraine, loss, know, media, wants, still, win, results, lost, blaming, would.
Unclear topic: video, watch, judge, Sessions, Napolitano, explain, it, kkk, mentor, reminder, Jeanine, calls, lol, go, thinks. (Jeanine Pirro (Judge Jeanine) and Napolitano were Fox News hosts at the time, while Sessions endorsed Trump during the 2016 elections and was briefly attorney general, so this could be a pro-Trump topic, but the word kkk makes the interpretation murky)