The Disinformation Ecosystem and the Frames Employed

Written By: Jonathan Schlosser
Prepared for: POLI 727 Framing
Instructor: Dr. Frank Baumgartner
Date: April 29, 2019

Introduction

During the 2016 US Presidential Election, an orchestrated disinformation campaign was undertaken to mislead the public and cultivate distrust in democratic institutions. This campaign employed social media to spread false content that played off of emotionally contested topics and divisive issues in the US in order sow ideological rifts (L. Bennett & Livingston, 2018; Parlapiano & Lee, 2018). The campaign was found to have reached upwards of 126 million people on Facebook alone, with the average American encountering 1-3 "fake news" stories during the month leading up to the election (Allcott & Gentzkow, 2017; Isaac & Wakabayashi, 2017). With the upcoming 2020 US Election, and evidence of malicious actors already working against Democrat candidates, research into the functioning of disinformation campaigns is necessary (Korecki, 2019).

The aim of this study is to examine the content included in the disinformation campaign surrounding the 2016 US election, while also trying to employ topic modelling as a way of identifying frames. With that being said, the examination here is more exploratory in nature and, as such, the aspiration is to see if the methodology can deliver valuable insights. Specifically, this work will assess a dataset of tweets identified by Twitter as being associated with the Internet Research Agency. While there are previous studies employing this dataset, this work takes a different approach in order to identify the frames employed and to gain new insights (Romano, 2018). This study will examine the IRA Twitter dataset and work to explore the frames used via the content and topics identified within the tweets.

Since the US election, the research area surrounding disinformation has been growing consistently while the conversation has become more nuanced. Therefore, the first step when studying disinformation is to clarify what the term actually means and what it does not. From there, certain aspects of disinformation will be discussed, and the elements that contribute to its spread and believability will be highlighted. The focus will then shift to discuss the Internet Research Agency and its role in the 2016 election. The methodology and results will then be presented, with the work culminating in a discussion concerning the application of topic modelling in this context and the frames found within the dataset.

Disinformation

When discussing disinformation, the term "fake news" has often been employed, but this term is not be the most appropriate (Kumar & Geethakumari, 2014; McCright & Dunlap, 2017; Ross & Rivers, 2018). "Fake news" has been defined as fabricated information that mimics news media in form, but not in content or intent; it consists of false claims that are designed to be shared and grab attention (Lazer et al., 2018; Pennycook & Rand, 2018b). With indiscriminate use though, a great deal of confusion has been sown around the term "fake news" (L. Bennett & Livingston, 2018; Cooke, 2017; Ross & Rivers, 2018). Ross and Rivers (2018) have shown that the Trump campaign in particular has blurred this term's definition by conflating disinformation from Russian actors with unfavorable reporting from CNN. Additionally, the term "fake news" fails to convey the variety and complexity of the different types of information referred to in the broad conversation, and is inadequate when it comes to characterizing the overall complex phenomena (Tandoc et al., 2017). Therefore, the terms "disinformation", "misinformation", and "malinformation" have been employed instead.

Disinformation has been conceptualized as patently false information with the intention to deliberately cause harm to either a person, social group, organization, or country (Stephan Lewandowsky, Ecker, & Cook, 2017; Marwick & Lewis, 2017; Wardle & Derakhshan, 2017). This type of false information is completely fabricated and malicious in aim. It can consist of deliberately manipulated audio, visual, or written information that is intentionally tied to conspiracy theories or rumors (Marwick & Lewis, 2017; Wardle & Derakhshan, 2017). Facebook has provided an alternative term for this type of information, “false news”, which has been described as news articles containing intentional misstatements to trigger passions, attract viewership, and/or deceive (Weedon, Nuland, & Stamos, 2017). Overall, disinformation is intentional and malicious, which differs from misinformation and malinformation.

Misinformation is information that is false but is not deliberately created with the intent of causing harm (Del Vicario et al., 2016; Stephan Lewandowsky et al., 2017; Wardle & Derakhshan, 2017). This type of information can be conceptualized as unintentional mistakes, inaccurate photo captions, incorrect dates, flawed statistics, poor translations, or satire (when mistakenly thought of as factual) (Wardle & Derakhshan, 2017). While this type of information may lead to erroneous conclusions, its intent is not to purposefully deceive or cause harm.

Malinformation, on the other hand, has been described as information that is used to inflict harm on a person, organization, or country (Wardle, 2017; Wardle & Derakhshan, 2017). This occurs when private information is deliberately published for personal or corporate interest rather than public interest (Wardle & Derakhshan, 2017). A perfect example of this type of information is “revenge porn” which occurs when someone posts nude photographs of a past romantic partner online with the intention of causing the past romantic partner harm (Wardle & Derakhshan, 2017). While this type of information is a serious concern and can cause significant damage, it will not be examined within this work.

This work will focus on disinformation because of its pervasiveness on social media, the likelihood of an average user encountering it, the deliberate use to manipulate, and the direct association this type of content has with socio-political issues.

The Spread of Disinformation

The spread of disinformation online has been shown to be massive, while also having the ability to attract significant user engagement. As mentioned previously, the Russian disinformation campaign reached upwards of 126 million people on Facebook with individuals encountering 1-3 "fake news" stories on average during the month leading up to the election (Allcott & Gentzkow, 2017; Isaac & Wakabayashi, 2017). In addition to this, a *Buzzfeed* analysis found that Facebook engagement (likes, comments, and shares) was greater for the top 20 "fake news" stories than for the top 20 "real news" stories in the three months leading up to the 2016 presidential election (Silverman, Strapagiel, Shaban, Hall, & Singer-Vine, 2016). Considering that 62% of US adults get their news from social media and that many people who see "fake news" stories report that they believe them, the challenges become even more evident (Gottfried & Shearer, 2016; Silverman et al., 2016).

Overall, though, disinformation has been found to flow "father, faster, deeper and more broadly" through social media networks when compared to factual information (Vosoughi, Roy, & Aral, 2018). When considering the extensive presence of disinformation on social media, it is almost guaranteed that at some point attention will be captured and engagement will occur. Past work has shown that individuals direct their focus to things they find pertinent, which is largely based on whether they find an item interesting, valuable, and/or socially significant (Kosicki & McLeod, 1990; Sears & Freedman, 1967). Disinformation poses a threat in this regard as malicious actors can orient their content to appeal to these determinants of attention and get people to notice false information.

Furthermore, it has been well studied that repeated information will not only capture attention but will reinforce itself and have a greater influence on an individual's cognition, decision-making, and behavior (L. Bennett & Livingston, 2018; de Zúñiga, Barnidge, & Scherman, 2017; Higgins, 1996). This idea stems from the concept of reinforcement, where interacting with content repeatedly works to increase its prominence in one's psychological processing (Ecker, Hogan, & Lewandowsky, 2017; S Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012; Pennycook, Cannon, & Rand, 2018). In a similar vein, increased exposure has been shown to make users more credulous to the information being reinforced (Bessi, Scala, Rossi, Zhang, & Quattrociocchi, 2014; Del Vicario et al., 2016; Mocanu, Rossi, Zhang, Karsai, & Quattrociocchi, 2015). In relation to disinformation, the false content can then be repeated ad infinitum so that people are forced to see it and interact with it. Bots have actually been shown to play a role here by amplifying the spread of disinformation and improving the likelihood of it being seen by users (Lazer et al., 2018).

Thus, the sheer volume of disinformation, combined with the reliance on social media for news, and the influence of cognitive biases, disinformation becomes a significant concern. In the next section, some of the prominent biases that come into play will be discussed.

Disinformation, Selective Exposure, and Confirmation Bias

There are a number of factors that relate to disinformation and facilitate its ability to grab attention and engagement but selective exposure and confirmation bias are often the focus (Bakshy, Messing, & Adamic, 2015; Barberá, Jost, Nagler, Tucker, & Bonneau, 2015; Garrett, 2009; Lazer et al., 2018; Messing & Westwood, 2012). Selective exposure has been identified as an individuals' likelihood to interact with information that falls in line with previously held beliefs and to ignore information that is contradictory (Bakshy et al., 2015; Garrett, 2009; Messing & Westwood, 2012). When choices are made concerning which information to interact with, there are purposeful decisions as to which information to ignore, avoid, or reject occurring simultaneously; often these decisions are made in order to maintain existing beliefs (Cooke, 2017; Mai, 2016). In this respect, it can be understood that people bring their own wishes and expectations to any communication situation, and that these elements can then influence what those individuals notice and remember (Neisser, 1976; Smith, 1998). Bode (2016) has further posited that the alignment of information with prior beliefs acts as a key factor guiding information-engagement decisions. A number of additional studies reaffirm this idea that prior beliefs effect media attention and engagement, and it is now generally accepted throughout the literature (Bakshy et al., 2015; Barberá et al., 2015; Garrett, 2009; Schmidt et al., 2017; Sears & Freedman, 1967; Stroud, 2008). Overall, though, selective exposure occurs when an individuals' habits, beliefs, and perceptions of benefits minimize the diversity of information they encounter (Messing & Westwood, 2012; Sears & Freedman, 1967).

Similar to selective exposure is the concept of confirmation bias -- the preference for confirmatory information as opposed to differing information (Bakshy et al., 2015; Lazer et al., 2018). In other words, confirmation bias refers to our preference for information that confirms prior-held views and beliefs. While confirmation bias and selective exposure are similar, confirmation bias refers to the preference for information that reinforces a previously held idea, whereas selective exposure refers to the exposure to information in line with previously held beliefs (Bakshy et al., 2015; Barberá et al., 2015; Passe, Drake, & Mayger, 2018; Schmidt et al., 2017; Sears & Freedman, 1967).

In an extreme case, the lack of information diversity could be attributed to a "filter bubble", where information is filtered by the social media platform to the extent that it creates a state of limited information exposure (Pariser, 2011). Facebook, Google, and many other companies, often employ algorithms to prioritize content that a user would likely prefer based on their past decisions and interactions with information (Messing & Westwood, 2012; Pariser, 2011). Thus, one's prior decisions about which content to attend to influences the platform's algorithm and results in the prioritization of similar information in the newsfeed (Pariser, 2011; Sunstein, 2007). In essence, "filter bubbles" affect the information a user encounters and minimizes content diversity (Messing & Westwood, 2012; Pariser, 2011). Similarly, when a person's network on social media primarily affects the information they encounter, it can create an echo-chamber (Bakshy et al., 2015; Barberá et al., 2015; Flaxman, Goel, & Rao, 2016; Garrett, 2009). An echo-chamber refers to the state when one's social media network is solely reflective of their beliefs and the only information they encounter is confirmatory (Bakshy et al., 2015; Barberá et al., 2015). This is partly due to similarities amongst connections in their network, but is also influenced by the preference for information sources that confirm one's beliefs (Bakshy et al., 2015; Bakshy, Rosenn, Marlow, & Adamic, 2012).

Disinformation Framing

The disinformation ecosystem is of particular importance because malicious actors can employ coordinated efforts to purposefully disseminate false information, exploit attention, and distort understanding (W. L. Bennett & Livingston, 2018; Weedon et al., 2017). Specifically, disinformation can be designed in such a way that hijacks cognition processes and makes individuals more susceptible to its message. As Pennycook and Rand (2018a) found, "lazy thinking" may be a cause of the formation of incorrect beliefs as individuals tend to employ analytical thinking when encountering "fake news" but that it may not be fully accurate. This finding deviates from the traditional view that a lack of media literacy is the culprit when individuals believe false information. Instead, individuals employ active decision-making when discerning the “truth” of the content but that decision-making process may not be as critical as is needed (Pennycook & Rand, 2018a).

One way in which disinformation can be designed is in the form of simplified messages; which can work to quickly grab attention and encourage abstract, rather than critically-evaluated, interpretations to be drawn. Horne and Adali reviewed the structure of numerous “fake news” articles and found that often the content is shorter and employs more repetitive language (Horne & Adali, 2017). In particular, “fake news” articles tend to contain very little information in the actual article and pack most of their claim in the title; essentially, the article is basically repetition of the content in the title (Horne & Adali, 2017). This short structure works to quickly convey the main message, while the repetition works to reinforce the simple message – this may then result in critical evaluation routes being bypassed.

Furthermore, disinformation may trigger specific types of thinking, and purposefully affect audience beliefs – essentially manipulating their understanding and leading to cognitive fallacies (Rapp & Donovan, 2017; Smith, 1994). Attention and engagement have been shown to be partially determined by environmental and social cues, as well as an individual’s prior beliefs and attitudes (Higgins & Bargh, 1987; Levine, Resnick, & Higgins, 1993; Sotirovic & McLeod, 2004). Thus, by activating certain ideas and feelings in a viewer, particular trains of thought can be triggered and the outcomes of beliefs and attitudes can be controlled (Price, Tewksbury, & Powers, 1997). This occurs because certain properties of the information, combined with certain contexts, trigger particular interpretive frames which then influence the individuals’ processing of the information (Kahneman & Tversky, 2013). Therefore, the way in which information is presented may directly influence how that information is incorporated into an individual’s ideologies. And, as such, the manner in which disinformation is framed can directly influence how it grabs attention, affects attitudes, and manipulates beliefs.

The Internet Research Agency

During the 2016 US Presidential election, a large portion of the disinformation came from a handful of organizations. The most prominently discussed is the Internet Research Agency, also known as the Russian Troll Factory (Romano, 2018). This was one of several Russian firms that trained and paid individuals to troll online (L. Bennett & Livingston, 2018). The IRA had upwards of 600 employees and an estimated budget of around US $10 million (L. Bennett & Livingston, 2018). The trolls were expected to: post on news articles 50 times a day; maintain six Facebook accounts and post new content at least three times a day; or maintain at least 10 Twitter accounts and tweet about 50 times on each one (L. Bennett & Livingston, 2018). All of the actions came with specific goals and targets that needed to be met (L. Bennett & Livingston, 2018). Therefore, the extent of the operation was massive and generated a significant amount of content.

Furthermore, analyses of the content stemming from the IRA have shown that it was designed to play on divisive topics and to sow distrust in societal institutions (L. Bennett & Livingston, 2018; Parlapiano & Lee, 2018; Romano, 2018). Some topics that have been discussed include Donald Trump, Hillary Clinton, Black Lives Matter, and Make America Great Again (Romano, 2018). In one ad, Hillary Clinton was compared to Satan, while another article stated that Pope Francis endorsed Donald Trump, both of which were patently false claims (Allcott & Gentzkow, 2017; Burkhardt, 2017; Kang, Fandos, & Isaac, 2017). With that, though, studies have shown that the most discussed fake news stories tended to favor Donald Trump over

Hillary Clinton (Allcott & Gentzkow, 2017; Silverman et al., 2016). Therefore, there is an expectation that these topics will present themselves in this work.

Continuing, in face of court proceedings and the publicity of the Russian troll factory, many companies have begun to respond. Facebook and Twitter have been working to remove numerous fake accounts and identify sources of fake information, while Facebook, Twitter and Google have all implemented mechanisms to flag questionable content (Fandos & Roose, 2018; Romano, 2018; Wakabayashi & Isaac, 2017; Weedon et al., 2017). Twitter has actually released a large data set of tweets linked to the Internet Research Agency and other prominent sources of disinformation during the 2016 US election. The IRA data, sourced directly from Twitter, will be analyzed in this work.

Research Aim

The aim of this study is to examine disinformation coming from the Internet Research Agency and to identify the frames employed. The work will also attempt to employ a methodology based on topic modelling and to assess its effectiveness. With that being said, this work is exploratory in nature and will be used to develop more rigorous research questions and to guide future methodological approaches. Therefore, the main objectives of this work are to:

1. Identify major topics within the IRA twitter data set.
2. Use topic modelling to identify the frames within those major topics.
3. Assess the effectiveness of topic modelling for the identification of frames.

Methodology

This study employed data sourced from Twitter, which included roughly 9 million tweets from 3,613 accounts associated with the Internet Research Agency (https://about.twitter.com/

en_us/values/elections-integrity.html#data). The data was collected in October 2018 and includes user information (screen name, reported location, profile description, and profile URL), tweet information (language, text, time, replies, quotes, hashtags, and retweets), and information concerning the spread of the tweet (quote count, reply count, and retweet count). The data is based on the information available at the time of account suspension. Since most of the accounts here have been deactivated or suspended, many of the tweets are no longer found online and, therefore, going back to the source material is not an option. For example, this can be problematic when attempting to look into the URLs used within tweets as shortened links may no longer be able to connect back to the original URL. Additionally, users that had less than 5,000 followers at the time of suspension have been anonymized throughout the dataset, which would prevent any granular examination of individual users. The data is well suited for this analysis, though, as it contains the text of the tweet, which is the main focus of the study. By examining the text, the topics discussed and the frames employed can be identified.

To conduct the analysis, the dataset was first cleaned and processed. This included narrowing down the dataset to only reflect tweets in English, which resulted in a new total of 2.9 million tweets. Then the text was cleaned with http, RT, @, usernames, punctuation, and common stop words removed. This step narrowed down the tweet text and focused it solely on content that would provide meaningful insights.

Next, the tweet text was analyzed on a number of levels to identify major topics contained within the cleaned-up dataset. First, the tweet text was broken into individual words and the words were then counted to identify their prevalence within the data set. This provided the most granular analysis and allowed for individual words to be assessed. With this though, individual words do not often convey a message. Therefore, the next step worked to identify paired combinations of words within the dataset and their prevalence. This contextualized the terms more and began to reveal the major topics within the dataset. This analysis was then conducted with sets of three words and sets of 5 words. For each of these analyses, a network of words was established in order to identify the major topics present.

From these analyses, five major topics were selected to move onto the topic modelling stage. These topics were based on their prevalence within the tweet dataset, the likelihood of providing different frames, and previous analyses of IRA tweets. The five major topics included Black Lives Matter, Obama, Trump, Make America Great Again, and Hillary Clinton. Each of these topics were demarcated by terms revealed in the word-network analyses and essentially filtered out large numbers of tweets. This focused the tweets under analyses and also made things a bit more manageable.

With the major topics identified, the analysis proceeded in order to examine the frames within each of these major topics. To achieve this, a Latent Dirichlet Allocation (LDA) topic model was employed to identify 5 frames within each topic. The number of topics was essentially picked arbitrarily in order to begin assessing this form of analysis. With that being said, this was decided to allow deviations within the topic to reveal themselves, while also keeping the analysis fairly manageable. Additionally, the γ for each tweet was the major focus as it reveals the likelihood of that tweet being included in that particular LDA topic. Therefore, it provides a good indication of representative tweets for that topic and, as such, can be used to determine the overarching frame.

With that, the results are presented in the next section. It begins by showing the word/word-network analyses and then moves into the LDA analyses. Discussion of these results follows afterward.

Results

The analysis was essentially taken in two parts; the first was an exploratory analysis of top terms within the twitter data set and their co-occurrence, and the second step was an attempt at topic modelling employing Latent Dirichlet Allocation (LDA). The first half of the analysis was broken down into four different approaches to identify the major concepts and to ensure that the main concepts had an opportunity to come out from the data. These included a single-word analysis, a paired-word analysis, a three-word analysis, and a five-word analysis.

The first level of analysis focused on the identification of top words within the data set. The top 20 words are presented in Figure 1 below. As shown in the image, the terms “news” (n = 267,616), “trump” (n = 109,371), “politics” (n = 83,605), “police” (n = 68,978), “Obama” (n = 51,815), and “black” (n = 13,974) came up pretty frequently. This analysis started to reveal the main topics in the data set based on the words present in the text of the tweets.

The next analysis was based on the co-occurrence of word pairs. To achieve this, the tweet text was processed in order to create the word pairs, and then counted to determine that pair’s frequency within the data set. While this analysis did bring in some noise, it helped identify major terms within the data set and identify links to other terms. The network image of these paired words is in Figure 2.

From Figure 2, we can see that there are some major sets of words that co-exist within the tweet data and can again see some major topics forming. There seems to be a pretty prominent cluster around the term “news” with connections to “police”, “fake”, “Chicago”, “local”, and “politics”. This cluster seems fairly diverse in regards to the terms that it incorporates, but that could be an indication of the diverse content associated with the term “news” and how frequently the term appears throughout the dataset. In this network, there are additional co-occurring terms that seem interesting including the connections between the terms “Donald” and “Trump” with the terms “Obama”, “Bernie”, and “Sanders”. In a different cluster, we can also see “Hillary” and “Clinton” come through. The other thing to note, is that there are a number of peripheral topics that can be seen in the network, including: “climate” “change”; “planned” “parenthood”; “supreme” “court”; “payday” “loans”; “social” “media”; “pleads” “guilty”; and “arrestobama”. While this begins to show more of the topics present in the data set, the next stages of the analysis worked to concentrate the focus on terms even further.

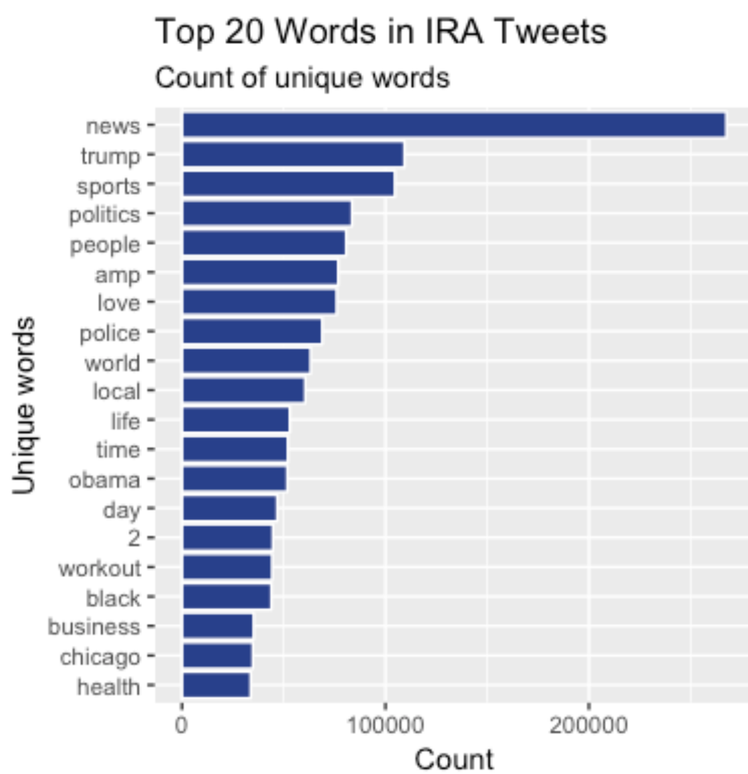


Figure 1: Top 20 Words in IRA Tweets. This image shows the top words in the IRA Twitter data set based on separating out each of the words within the tweet text and counting to identify their frequency.

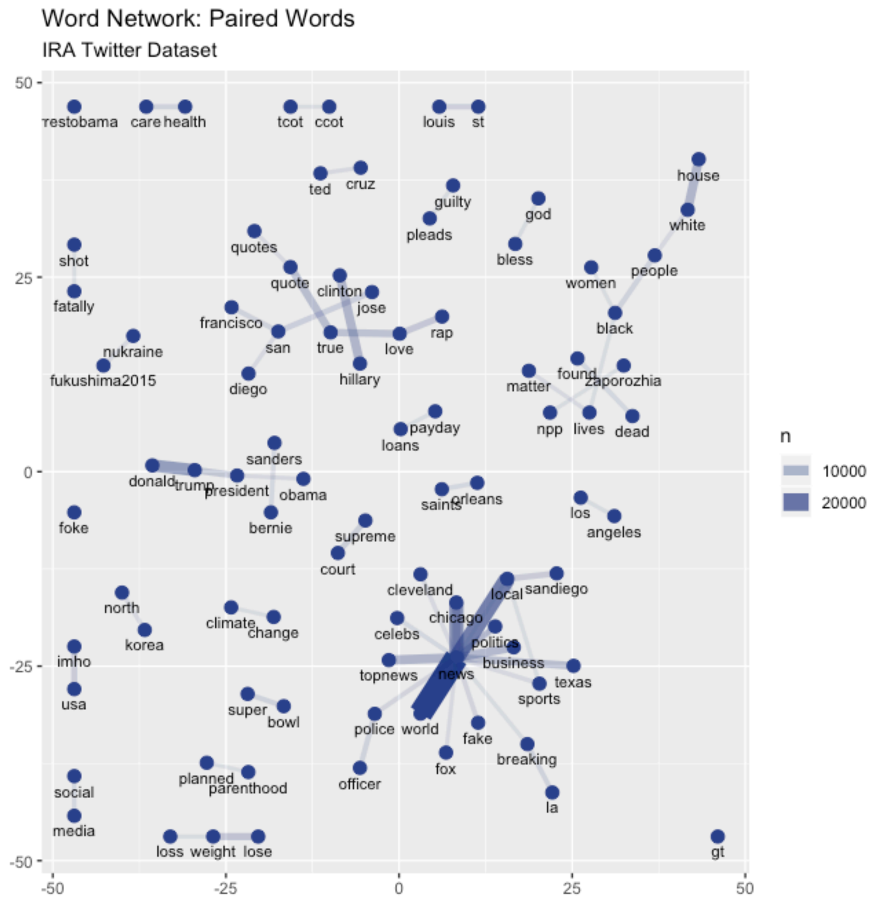


Figure 2: Word Network: Paired Words. This image shows the top paired-word associations in the IRA Twitter data set based on identifying the word pairs within the tweet text, counting their frequency, and identifying the strength of the word associations.

In the three-word analysis, there is still a prominent cluster around the word “news”, but here the connections are fewer with associations to the terms “world”, “Chicago”, “local”, “celebs”, and “breaking”. With that being said, there are still connections to the terms “politics” and “donald”, showing that these terms are often used in conjunction with the term “news”. Additionally, we see that the terms “ukranian”, “government”, “doesn’t”, and “care” associated with the terms “nuclear” and “power”, which indicates the diversity of topics discussed within the dataset. Another topic also seems to emerge from this analysis as seen in the terms “isis”, “targeted”, “accounts”, and “iceisis”. Lastly, many of the terms that were previously seen come up again, including “black” and “lives”, “americafirst” and “makeamericagreatagain”, and “arrestobama”. Figure 3, below, shows this term network and the term interlinkages.

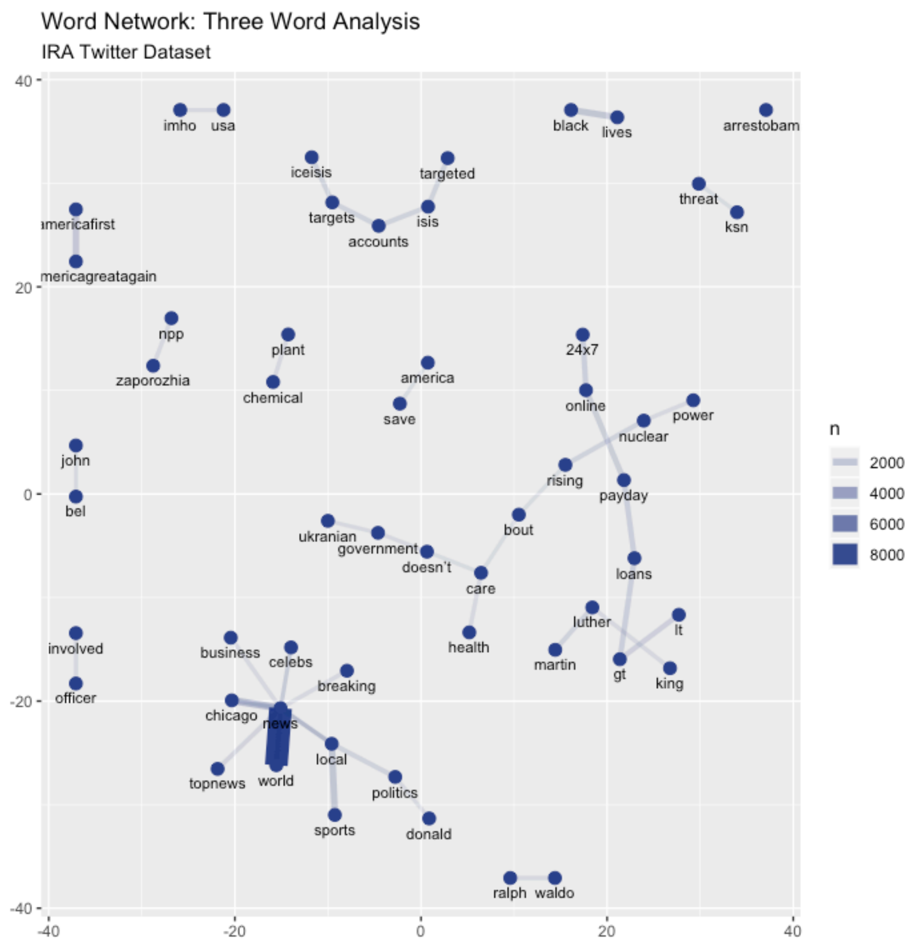


Figure 3: Word Network: Three Word Analysis. This image shows the top three-word associations in the IRA Twitter data set based on identifying the three-word connections within the tweet text, counting their frequency, and identifying the strength of the word associations.

The last stage was to conduct a five-word network analysis of the terms within the dataset. This term network was developed to see if themes emerged from longer chunks of text, as compared to the previous analyses. In this network, the “news” cluster that we have seen previously is no longer prominent. With this, we see other prominent topics emerge, such as one focused on payday loans (with the terms “24x7”, “online”, “payday”, and “loans”) and one focused on deep state and shadow government conspiracies (with the terms “Maga”, “War”, “WakeUpUSA”, “DeepState”, “ShadowGovernment”, and “Obamagate”). With that though, other topics that we have seen previously come up again, including: “ukranian”, “government”, “doesn’t”, and “care”; “isis”, “targeted”, “accounts”; and “arrestobama”.

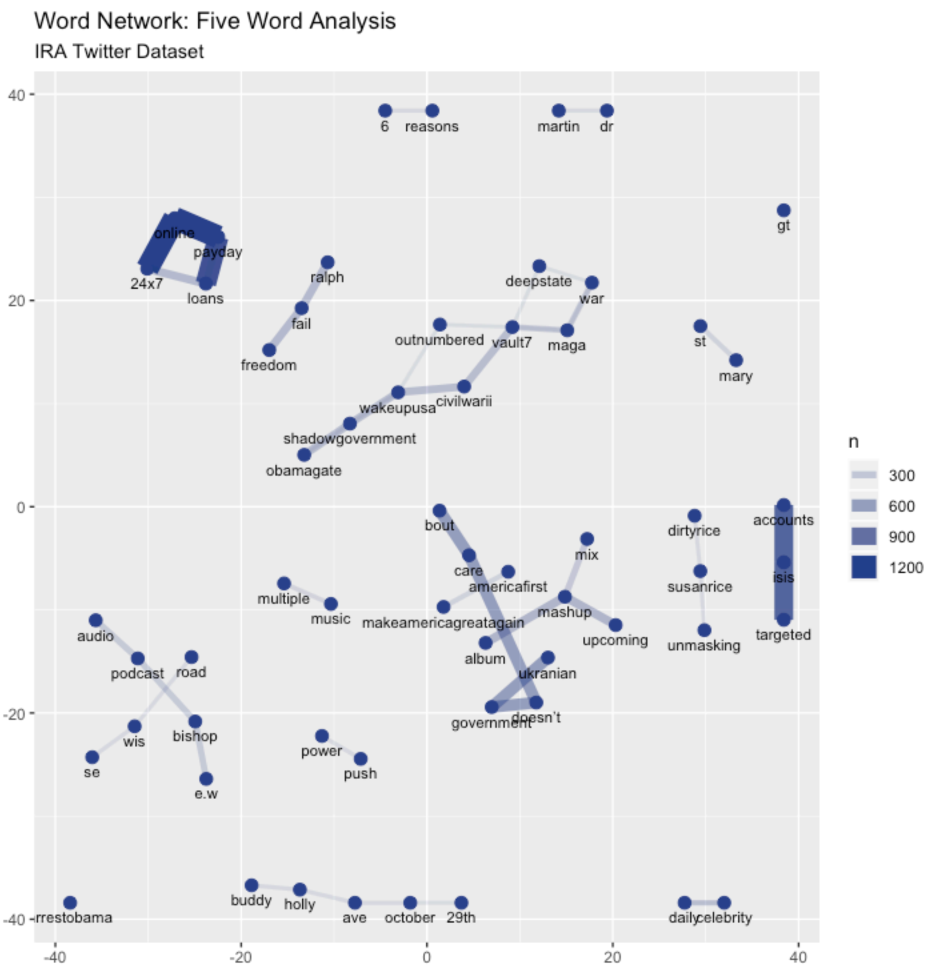


Figure 4: Word Network: Five Word Analysis. This image shows the top five-word associations in the IRA Twitter data set based on identifying the five-word connections within the tweet text, counting their frequency, and identifying the strength of the word associations.

The next stage of the analysis was to go in depth with a few of these topics and to test whether a topic model could reveal the frames employed. For this stage, 5 separate LDA topic models were employed each focusing on different key terms. The first topic was Black Lives, employing the terms “black” and “lives”. The second was focused on Obama and used the terms “Obama” and “obamagate”. The third was Make America Great Again, using the terms “makeamericagreatagain”, “americafirst”, and “maga”. The fourth was focused on Trump with the terms “donald” and “trump”. The fifth focused on Hillary Clinton with the terms “Hillary” and “Clinton”.

The first sub-analysis, focused on Black Lives Matter, identified 5 topics as determined from the LDA. The first topic consisted of 18,293 tweets out of 86,770 and included many tweets that showed support for black individuals. One tweet, stated “Senegalese designer, Sally Raby Kane, was just tapped by this major furniture company #ikea #blackdesigner…” (γ = 0.2323). While another stated “An 11-year-old black boy is found to have a higher IQ than Albert Einstien, Steven Hawking, and Bill Gates. Rahami Wilfred of Romford, U.K. has received a geniu-leveled Mensa IQ score of 162. His score places him in top 1 percent of IQs in the U.K.” (γ = 0.2184). With this though, there were a number of tweets that didn’t directly relate to an overarching theme. For example, a number of tweets encouraged the boycotting of certain companies that may have supported the Trump Campaign, such as “Boycott Companies that Supported Racist Trump Regime. Saks Off Fifth make racist afraid again…” (γ = 0.2252). Topic 2 contained 15,921 tweets out of 86,770. This topic largely seemed to be more of a “black power” sentiment with tweets such as “Stop Casting White Actors to Play People of Color #AkhenatonWasABlackPharoah #StopWhiteWashing” (γ = 0.2371), “Next highest rates of black disenfranchisement: - VA: 21.9% (vs. 7.9% among all adults) -FL: 21.3% (vs. 10.4%)…” (γ = 0.2224), and “This world was built on the backs of the black race. Don’t forget that. We are the queens and kings of this unpromising society” (γ = 0.2170). The third topic seemed to show a more direct pro-Black Lives Matter sentiment, with tweets like “A #BlackLivesMatter textbook is coming for middle and high schoolers #blmt #blacklivesmatter #blm2016 #school” (γ = 0.2354), but also focused on police officers, as seen in tweets like “#FOXNews Hosts Start Fund

To Help Counsel Officers Who Shoot #Black Suspects #BlackLivesMatter #BlackTwitter #riot" ($\gamma = 0.2234$) and "A #blackwoman slapped w/a $13,000 bill cause a #policeofficer didn't believe her" ($\gamma = 0.2232$). This third topic accounted for 16,460 tweets out of the 86,770 that encompassed the entire corpus. Topic 4 seemed to blend a bit with Topic 3, but concentrated more on high school students and police brutality with tweets like "Incident between Round Rock, #Texas PD & High School Student Caught on Camera #PoliceBrutality #BlackLivesMatter #BLM" ($\gamma = 0.2334$) and "#PoliceBrutality on a high school student #nyc #BlackLivesMatter #BlackTwitterr #BlackPeoplePrroblems #BlaclPower" ($\gamma = 0.2305$). Topic 4 encompassed 16,743 tweets out of 86,770. Topic 5, which is essentially a catch-all for the other tweets not previously demarcated, contained 17,823 tweets out of 86,770. Thus, Topic 5 contained tweets talking about the Black Hawks sports team and immigrants, along with those that actually related to Black Lives Matter. These tweets included "Blackhawks receive forwards Artem Anisimov, MarkoDano, Jeremy Morin, and Corey Tropp and a fourth-round selection in 2016…" ($\gamma = 0.2210$), "Speaking as someone who lives in NH, I'm not happy -> VT Admits Seven Refugees Diagnosed with Active TB – Breitbart" ($\gamma = 0.2165$), and "That's fucked up! Cops aggression & poor judgement is just terrifying! #StayWoke #BlackLivesMatter" ($\gamma = 0.2207$). Throughout these topics though, and the topic modelling analysis, there were tweets that didn't seem to be placed in the correct topic. For example, the tweet "Lost: $6 Billion Laptops, tablets, and cellphones 33,000+ emails four American lives Quite a 'record.' #NeverHillary" ($\gamma = 0.2244$) was designated to Topic 2 even though there is no clear connection to the overall topic and the rest of the tweets contained therein. This problem will be discussed later in more detail in the discussion section.

The second LDA focused on the terms “Obama” and “Obamagate” and aimed to discover the different frames surrounding these terms. The first topic from the LDA included 15,475 tweets out of 75,575 which mainly attacked Obama. Some examples of this content are seen in the tweets “Peters Calls Obama A ‘Total P***y”, Dash Says Obama ‘Couldn’t Give A S**t’ About ISIS, Both Get Suspended -BB$SP” ($\gamma$ = 0.2270), “Sailors BLAST Michelle Obama Over Navy’s Decision To Stop Serving French Fries, Fried Chicken, Whole Milk – BB4SP” (γ = 0.2198), and “Obama: It Wasn’t Just That We Survived a Marxist Pres, we Survived a Marxist Saboteur as Pres…” (γ = 0.2164). The second topic mainly promoted extreme views and conspiracy ideas, as seen in the tweets “WE ARE AT WAR! #MAGA #Vault7 #Outnumbered #WakeUpUSA #ShadowGovernment #ObamaGate #ARRESTObama #TheMessyTruth #Wiretap” (γ = 0.2391), “WE ARE AT #WAR! #DeepState #Vault7 #CivilWarII #WakeUpUSA #ShadowGovernment #ObamaGate #ARRESTObama #MAGA #Wiretap” (γ = 0.2336), and “#ObamaGate #ObamaWire tap IS true so #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama” (γ = 0.2347). This topic encompassed 12,639 tweets out of 75,757. The third topic, which contained 16,378 tweets out of 75,575, mainly showed positive support for Obama. Examples of this include “Barack Obama: 0 affairs 0 porn star payoffs 0 sexual assault allegations 0 divorces 0 drama Donald Trump: Many affairs Porn star payoffs (Stormy Daniels) 19 sexual assault allegations 2 divorces 100% drama” (γ = 0.2346) and “#ThankYouObama For being so inept that the Right now owns the Potus, Senate, House, Scotus Next 8 yrs will be awesome…” (γ = 0.2216). The fourth topic seemed to convey an anti-Trump sentiment with tweets like “Trump – over 50 scandals in one year Obama…. Tan suit… in 8 years” (γ = 0.3082), and “List of people Donald Trump has attacked: -LaVar Ball – Colin Kaepernick – Barrack Obama – Sgt. Johnson’s widow – Jemele

Hill – Marshawn Lynch – Khizr Khan – Stephen Curry – Frederica Wilson – London mayor Guess what all of them have in common? They're all people of color." (γ = 0.2442). With that being said though, there were tweets that had the direct opposite sentiment including "The 'Anti-Trump Wave' is a fabricated media lie. Counties won: Trump: 2,623 (84.3%) Hillary: 487 (15.7%) States won: Trump: 30 Hillary: 20 Democrat losses under Obama: 12 governorships 62 house seats 9 senators 958 state legislators NUMBERS DON'T LIE!" (γ = 0.2365). Thus, again showing the noise within the dataset. The fourth topic contained 16,427 tweets out of a 75,575. The fifth topic, which again was basically a catchall for tweets not included in other topics, contained 14,656 out of 75,575 tweets. This set contained a number of interesting tweets, including "#Obama May Be The Proverbial Weak Sister in Foreign Affairs, But He's RUTHLESS W/ Patriots, Constitutionalists, Tea Party…" (γ = 0.2225), "#Obama Claims #ClimateChange Greatest Threat – FLOTUS FLY SEPARATE PLANES #JunkScience #TCOT #PJNET" (γ = 0.2212), and "RETWEET if you think the policies made by Crooked Hillary and Weak Obama have caused the spread of ISIS. #MakeAmerricaGreatAgain..." (γ = 0.2139).

The third LDA focused on tweets related to the major topic of Make America Great Again. The first topic in this LDA seemed to be primarily Anti-Hillary with tweets like "Read Powerful Statement Repudiating Deplorables Racist Homophobes Islamaphobes Misogynists #TrumpsArmy #MAGA" (γ = 0.2241) and "He's unusual And unstoppable #Hillary's imploding America's rejoicing #MakeAmericaGreatAgain" (γ = 0.2201). This first topic contained 6,019 tweets out of 24,459. The second topic, which contained 4,312 tweets out of 24,459, seemed to be more of a mixed bag. It contained tweets calling for action, such as "Talk is cheap. We had 8 yrs of talk. Now is the time of ACTION. #MAGA OCare Repeal Economic Boon…" (γ = 0.2239), showing evidence of Trump not being racist, "President Trump signs proclamation honoring Dr. Martin Luther King Jr. But somehow he is still racist. Trump works to make life better for all Amerricans whatever your skin color is! #MAGA #MLKDay2018" ($\gamma = 0.2229$), and tweets calling for Obama's arrest, "ObamaGate wirretapping IS true so #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama" ($\gamma = 0.2195$). This arrest Obama sentiment largely comes through in the fourth topic model, but it also appears in this second topic, which again shows the noise in the dataset. The third topic seems to convey a pro-Trump sentiment with tweets like "How many RT's and Likes for the huge and beautiful border wall? Mexico|US Mexico|US Mexico|US Mexico|US Mexico|US Mexico|US Mexico|US #MAGA" ($\gamma = 0.2387$), and "PRESIDENT DONALD TRUMP We [heart emoji] YOU #MAGA #DescribeAGreenie #BuildTheWall #AmericaFirst #ANTIFA #EndSanctuaryCities #EarthDay" ($\gamma = 0.2275$). This third topic contained 4,788 out of 24,459 tweets. The fourth topic, as previously mentioned, seemed to be dominated with an "arrest Obama" sentiment and included tweets like "ObamaGate wiretap IS true so #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama" ($\gamma = 0.2347$) and "#ObamaGate #ObamaWire tap IS true so #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama" ($\gamma = 0.2255$). This topic contained 4,204 tweets out of the 24,459-tweet corpus. The fifth topic, was again a catchall, and contained tweets like "INVINCIBLE IMPERVIOUS 2 Leftist & #FakeNews ATTACKS I #StandWithFlynn 2 #MAGA #MAGAMARCH…" ($\gamma = 0.2309$), "#KeepBannon #KeepBannon #KeepBannon #KeepBannon #KeepBannon #KeepBannon #KeepBannon #AmericaFirst Where is Stephen Miller…" ($\gamma = 0.2237$), and "LOCK HIM UP -- #ObamaGate - #WireTap - #FakeNews - #WirreTapping - #ArrestObama - #DeepState - #TheHammer…" ($\gamma = 0.2153$). This fifth topic contained 5,136 tweets out of the 24,459 set.

The fourth LDA model focused on tweets relating to Trump and included the terms “donald” and “trump”. This analysis revealed some interesting topics that mainly oriented around pro-Trump and anti-Trump sentiments. The first topic, which contained 29,163 tweets out of 147,320, seemed to be pro-Trump but focused on continuing in the face of opposition. Some examples include “There will be no gotcha moment that will take down Trump we need to continue…” ($\gamma = 0.2729$) and “The ‘Anti-Trump Wave’ is a fabricated media lie. Counties won: Trump: 2,623 (84.3%) Hillary: 487 (15.7%) States won: Trump: 30 Hillary: 20 Democrat losses under Obama: 12 governorships 62 house seats 9 senators 958 state legislators NUMBERS DON’T LIE!” ($\gamma = 0.2497$). The second topic seemed to go the other way and contained anti-Trump sentiment. Some tweets that show this include “Trump marriage #1 – shutdown Trump marriage #2 – shutdown Trump Vodka – shutdown Trump Casino – shutdown Trump University – shutdown US Government – shutdown All this things were shutdown because of Trump. He’s doing the same thing to our country…” ($\gamma = 0.2319$) and “Donald Trump is what would happen if an email with the subject line FW: FW: FW: FW: FW: FW: were elected president” ($\gamma = 0.2242$). This second topic encompassed 28,479 tweets out of 147,320. The third topic seemed to also be anti-Trump but more focused on his failings. This topic contained 30,578 tweets out of 147,320, and included tweets like “First, they went after McCabe’s wife. Now it’s Ohr’s wife. The Nunes/Trump ghouls love looping people’s wives – noncombatants – into all this. Not breaking any news, but, yeah, Trump hates women. It’s a pattern” ($\gamma = 0.2384$), “Trump on Dec. 25: ‘Tomorrow it’s back to work!’ Trump on Dec. 26: Went golfing, no public schedule. Trump on Dec. 27: Went golfing, no public schedule. Trump on Dec. 28: Went golfing, no public schedule. Trump on Dec. 29: Went golfing, no public schedule.” ($\gamma = 0.2347$), and “Trump has fired or pressured: - Dep. FBI Dir. Andrew McCabe – US Attorney Preet Bharara – FBI Dir. Christopher

Wray – Dep. AG Rod Rosenstein – Acting AG Sally Yates – SC Robert Mueller – AG Jeff Sessions Only a fascist dictator would keep firing the people investigating him!" ($\gamma = 0.2237$). The fourth topic seemed to encompass tweets that directly and/or indirectly compared Trump to Hillary and Obama. This topic contained 26,952 tweets out of 147,320 and included tweets like "Trump – over 50 scandals in one year Obama…. Tan suit… in 8 years" ($\gamma = 0.2581$), "Lock her up? More like lock THEM up! Flynn Don Jr. Flynn Jr. Kushner Manafort Papadopoulos Donald Trump #Wikileaks" ($\gamma = 0.2326$), and "I can't wait until President Trump goes to Prison!...." ($\gamma = 0.2304$). Lastly, the fifth topic acted as a catchall and included tweets like "The nation-wide #March4Trump will be held on Saturday March 4th, in: DC NC FL MO MN NV CA CO LA TX AZ PA…" ($\gamma = 0.2329$), "Disgusting racist Trump has insulted Maxine Waters during the NRCC dinner: 'I watch this Maxine Waters. You ever see Maxine Waters? A low IQ individual. Low IQ." ($\gamma = 0.2244$), and "President Trump has arrived in the Philippines to attend the 31st ASEAN Summit and Related Summits. Welcome and mabuha…" ($\gamma = 0.2234$). This diverse set of tweets totaled 32,148 out of 147,320.

The last LDA focused on Hillary Clinton. In this analysis, all of the topics seemed to paint Clinton in a negative light. The first topic seemed to focus on corruption with tweets like "There's HUGE amount of corruption/criminal activities related 2 Clintons & Obama 2 investigate. Yet, Dems demanded Spcl Counsel 4 phony Russia…" ($\gamma = 0.2271$) and "#HillarysEmails FULL 55k "Us" #ClintonIt & State Dept Gowdy you get what they deem relevant, got it !? Take 500 & STFU! …" ($\gamma = 0.2184$). This topic contained 12,798 tweets out of 64,905. The second topic seemed to be more of a mixed bag with tweets like "The 'Anti-Trump Wave' is a fabricated media lie. Counties won: Trump: 2,623 (84.3%) Hillary: 487 (15.7%) States won: Trump: 30 Hillary: 20 Democrat losses under Obama: 12 governorships 62 house seats 9 senators 958 state legislators NUMBERS DON'T LIE!" ($\gamma = 0.2546$) and "Chelsea Clinton's publicists must be stopped for the sake of humanity" ($\gamma = 0.2405$). This topic contained 13,027 tweets but did not seem to convey a particular message or overarching theme. Topic 3 seemed to also be anti-Hillary, but also seemed to be more general. This topic contained 13,308 tweets out of 64,905 and included tweets like "If Carson loses, he'll return to his surgery. If Trump loses, he'll return to his business. If Hillary loses, she'll return to hell #GOPDebate" ($\gamma = 0.2224$), "America needs a real man to get the Job DONE! #MAGA #AmericaFirst #ImWithYou #illaryClinton" ($\gamma = 0.2223$), and "Emails show Ambassador Chris Stevens didn't want to stay in Bengazhi but Hillary made him and others stay, may they RIP…" ($\gamma = 0.2219$). The fourth topic seemed to go pro-Trump and continue the anti-Hillary messaging. This topic contained 13,646 tweets out of 64,905, and included tweets like "DONALD TRUMP WE [heart emoji] YOU MAKE AMERICA GREAT AGAIN TRUMP WILL SAVE USA LAW & ORDER CANDIDATE…" ($\gamma = 0.2398$), "1 year ago today, Trump stomped #CrookedHillary like a Narc at a Biker Rally… the look on the faces of the #FakeNews Media was priceless #TrumpIsAmericasPresident #BiteMeMSM" ($\gamma = 0.2297$), and "NATIONAL SECURITY In my opinion #HillaryClinton herself is perhaps the biggest threat to National Security #MakeAmericaGreatAgain…" ($\gamma = 0.2167$). The last topic contained 12,126 out of 64,905 and included a variety of tweets, such as "Net Worth before running for - \$4.5B - \$3M - \$480,000 Current Net Worth. Trump - \$3.5B Obama - \$40M+ Clinton - \$100M+ Very telling indeed… #ThingsPoliticiansShouldntSay #SundayMorning" ($\gamma = 0.2223$), "Saudi Prince #AlWaleed Bin Talal, 2nd largest shareholder #Twitter, Clinton's #1 Donor, has been arrested on money laundering…" ($\gamma = 0.2205$), and "List of who Trump has tried to blame for Russian Interference in last 24 hours: - Hillary Clinton – Barack Obama – Democrats – All Intel Agencies – Media –

FBI – A 400-pound genius sitting on a bed List of who Trump hasn't blamed: - Putin" (γ = 0.2204).

Overall, the LDA analyses highlighted a number of frames that were employed throughout the dataset, but there are some interesting things to note. The gammas throughout the analysis are quite small, which indicates that the likelihood of these tweets being in this particular topic is quite small. Thus, a fair amount of noise could have entered the analysis and caused errors in the topic modelling. Also, there are a number of tweets that appear multiple times throughout the different LDAs. This makes sense though as each LDA conducted was not mutually exclusive and, as such, if the terms associated with different major topics appeared in the same tweet, the tweet would be included in the associated LDAs.

The next section will dive deeper into these results and will discuss the implications of these findings.

Discussion

While this methodological approach did provide some key insights into the IRA Twitter dataset and the frames employed, it was not as clear as originally hoped. The word and word-network analyses did help to identify major topics within the dataset, but there were also many terms that did not necessarily relate to pertinent subjects. Also, the topic modelling approach did identify frames, but the results were fairly messy with many non-related tweets included. Thus, the topic modelling did work even though it was not perfect. With corrections and refinement, this methodology could be quite useful and could provide deeper insights into the frames used and the ways in which topics were discussed.

As mentioned previously, the word and word-network analyses helped guide decisions on which major topics to attenuate to within the dataset, but there were many terms that were not necessary. Primarily, the term “news” was way too broad to be able to identify a clear topic and essentially did not act as a good guide. Furthermore, topics that emerged, such as: “ukranian”, “government”, “doesn’t”, “care”, “about”, “nuclear”, “power”; and “24x7”, “online”, “payday”, “loans”, related more to issues outside of the US election and Trump and, therefore, were not useful. The former may have been targeted to provoke Ukranian response and sow division within that community, while the latter may have been spam from bots – this is the conclusion that Romano (2018) arrived at as well. With that being said, the terms for the 5 LDA analyses (“black”, “lives”; “Obama”, “obamagate”; “makeamericagreatagain”, “americafirst”, and “maga”; “donald”, “trump”; and “Hillary”, “Clinton”) did all appear in the word and word-network analyses, and were featured quite prominently. This is not surprising though as it confirms what other researchers have previously found.

Moving on to the LDA, the methodology did seem to provide some of the frames employed even though they were not as clear as originally hoped. Table 1, below, highlights the major frames as identified in the analyses. As mentioned previously, some analyses of the content stemming from the IRA have shown that it was designed to play on divisive topics and to sow distrust in societal institutions; this was confirmed here (L. Bennett & Livingston, 2018; Parlapiano & Lee, 2018; Romano, 2018). In the LDA topics, we can see that there are both pro- and anti- sentiments throughout the topics and that parallels can be drawn to many social and political issues. For the Black Lives Matter topics, we see this theme of promoting black lives matter and black power. We also see content framed as BLM vs. police officers, which was a tense topic throughout most of the election. This may show how the IRA worked to promote division and sow distrust in social institutions. Continuing, the Obama topic had both pro-Obama and anti-Obama sentiments contained within the tweets, as revealed by the topic model, and shows how the tweet played to both sides. This was also seen in the LDA focused on Trump with both pro-Trump and anti-Trump sentiments contained in the tweets. The "Make America Great Again" topic seemed to be a bit more mixed though with some tweets attacking Obama, some showing support for Obama, some promoting extremist views and conspiracy theories, and some containing Anti-Trump sentiment. This may be an indicator that the "Make America Great Again" topic did not successfully demarcate the content and thus revealed unclear results. The Hillary Clinton topic was also very interesting as it seemed to predominantly promote anti-Hillary sentiments rather than containing both pro- and anti- sentiments. This makes sense though as other analyses have shown that there was a major focus on promoting anti-Hillary content (Kang et al., 2017; Romano, 2018; Ross & Rivers, 2018).

Table 1: Topics for each of the LDA Analyses. This table shows the major sentiments conveyed in the tweets within each of the LDA analyses conducted.

Topics for each of the LDA Analyses					
	LDA Topic 1	LDA Topic 2	LDA Topic 3	LDA Topic 4	LDA Topic 5
Black Lives Matter	Support For Black Individuals	"Black Power" Sentiment	More Direct Pro-Black Lives Matter Sentiment	High School Students And Police Brutality	Catchall
Obama And "Obamagate"	Mainly Attacked Obama	Promoted Extreme Views And Conspiracy Ideas	Positive Support For Obama	Anti-Trump Sentiment	Catchall
Make America Great Again	Anti-Hillary	Not Clear – Mixed Bag.	Pro-Trump Sentiment	"Arrest Obama" Sentiment	Catchall
Trump	Pro-Trump - Focused On Continuing In The Face Of Opposition	Anti-Trump Sentiment	Anti-Trump - Focused On His Failings	Directly And/Or Indirectly Compared To Hillary And Obama	Catchall
Hillary Clinton	Focus On Corruption	Not Clear – Mixed Bag.	More General Anti-Hillary	Pro-Trump	Catchall

Continuing, throughout the analysis, a number of tweets were contained in topics that seemed to be counter the major frame. For example, in Topic 1 connected to the major topic of Make America Great Again, the tweet "President Donald Trump Rally THURSDAY in

#WestVirginia Huntington, WV 7PM EST #AmericaFirst #MAGA #Jobs" ($\gamma = 0.2383$) appeared with a fairly high gamma. This tweet did not relate to the Anti-Hillary sentiment which was the predominate topic as designated by the LDA. Also, in Topic 1 connected to Trump, the tweet "Grand Visions for America: Trump: Xenophobia Unlimited Clinton: 'I'm not Trump' Stein: Green New Deal" ($\gamma = 0.2221$) appeared, which again did not relate to the major topic as identified by the LDA. This noise in the analyses could be due to a number of things, though: it could be due to the LDA working to equally separate tweets into each topic; poor demarcation of the terms used for the major topic; or not enough topics selected in the LDA to clearly separate tweets. With this, it is clear that the LDA method needs further refinement for this application and may need some correcting to be fully accurate.

Also, a number of tweets reappeared in multiple LDAs, including the tweets "ObamaGate wiretap IS true so #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama #ARRESTObama" and "The 'Anti-Trump Wave' is a fabricated media lie. Counties won: Trump: 2,623 (84.3%) Hillary: 487 (15.7%) States won: Trump: 30 Hillary: 20 Democrat losses under Obama: 12 governorships 62 house seats 9 senators 958 state legislators NUMBERS DON'T LIE!". This may seem like an error on first glance, but because the LDA's approached the dataset independently it actually makes sense. Also, these tweets may have had a certain prominence in the major topic delimitated by the key terms that they reappear multiple times.

Often times, an LDA is designed to predict how things relate to one another. That aspect was not included in this work as this work was more exploratory in nature. Also, due to the noise in the LDA analyses, prediction would be hard to accomplish accurately. Overall, though, the exploratory approach revealed some interesting things and seemed to show indications that the process may be applicable for future research and analysis. Since there was a lot of noise in these analyses, significant refinement is still needed even though the methodology seems feasible. If corrections are made, future analyses could be more accurate and hopefully deliver results that are clearer and better able to deliver complete insights.

Conclusion

While this work has highlighted some benefits for this methodology, there is still room for improvement. This work was predominantly an exploration into the applicability of this methodology for Twitter analyses in relation to disinformation. While some major topics and frames were found in this analysis, there was a significant amount of noise throughout. With this, conclusions need to be taken cautiously and viewed as indications rather than significant findings. In future work, this method could be refined and finessed to be more accurate and draw more concrete conclusions. The work basically acted as a proof of concept, and in this regard, it succeeded. Future work will aim to achieve that refinement and to develop a sounder methodology for more in-depth analyses. Overall, though, the aims of the study were met and there is a path to guide future work.

Citations:

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130 LP-1132. Retrieved from http://science.sciencemag.org/content/348/6239/1130.abstract

Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web* (pp. 519–528). ACM.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, *26*(10), 1531–1542. https://doi.org/10.1177/0956797615594620

Bennett, L., & Livingston, S. (2018). *The disinformation order: Disruptive communication and the decline of democratic institutions*. *European Journal of Communication* (Vol. 33). https://doi.org/10.1177/0267323118760317

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139. https://doi.org/10.1177/0267323118760317

Bessi, A., Scala, A., Rossi, L., Zhang, Q., & Quattrociocchi, W. (2014). The economy of attention in the age of (mis)information. *Journal of Trust Management*, *1*(1), 12. https://doi.org/10.1186/s40493-014-0012-y

Bode, L. (2016). Political News in the News Feed: Learning Politics from Social Media. *Mass Communication and Society*, *19*(1). https://doi.org/10.1080/15205436.2015.1045149

Burkhardt, J. M. (2017). *Combating fake news in the digital age*. ALA TechSource.

Cooke, N. A. (2017). Posttruth, Truthiness, and Alternative Facts: Information Behavior and Critical Information Consumption for a New Age. *The Library Quarterly*, *87*(3), 211–221. https://doi.org/10.1086/692298

de Zúñiga, H. G., Barnidge, M., & Scherman, A. (2017). Social Media Social Capital, Offline Social Capital, and Citizenship: Exploring Asymmetrical Social Capital Effects. *Political Communication*, *34*(1). https://doi.org/10.1080/10584609.2016.1227000

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., … Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 201517441.

Ecker, U. K. H., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and Repetition of Misinformation: Helping or Hindering Its Retraction? *Journal of Applied Research in Memory and Cognition*, *6*(2), 185–192. https://doi.org/https://doi.org/10.1016/j.jarmac.2017.01.014

Fandos, N., & Roose, K. (2018, July 31). Facebook Identifies an Active Political Influence Campaign Using Fake Accounts. *New York Times*. Retrieved from https://nyti.ms/2Oxpkfi

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, *80*, 298–320. Retrieved from http://10.0.4.69/poq/nfw006

Garrett, R. K. (2009). Echo chambers online?: Politically motivated selective exposure among Internet news users. *Journal of Computer-Mediated Communication*, *14*(2), 265–285.

Gottfried, J., & Shearer, E. (2016). *News Use Across Social Medial Platforms 2016*. Pew Research Center.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience.

Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual Review of Psychology*, *38*(1), 369–425.

Horne, B. D., & Adali, S. (2017). This just in: fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. *ArXiv Preprint ArXiv:1703.09398*.

Isaac, M., & Wakabayashi, D. (2017, October 30). Russian Influence Reached 126 Million Through Facebook Alone. *New York Times*. Retrieved from https://www.nytimes.com/2017/10/30/technology/facebook-google-russia.html

Kahneman, D., & Tversky, A. (2013). Choices, values, and frames. In *Handbook of the Fundamentals of Financial Decision Making: Part I* (pp. 269–278). World Scientific.

Kang, C., Fandos, N., & Isaac, M. (2017, November 1). Russia-Financed Ad Linked Clinton to Satan. *New York Times*. Retrieved from https://www.nytimes.com/2017/11/01/us/politics/facebook-google-twitter-russian-interference-hearings.html

Korecki, N. (2019, February 20). ‘Sustained and ongoing’ disinformation assault targets Dem presidential candidates. *Politico*. Retrieved from https://www.politico.com/story/2019/02/20/2020-candidates-social-media-attack-1176018

Kosicki, G. M., & McLeod, J. M. (1990). Learning from political news: Effects of media images and information-processing strategies. *Mass Communication and Political Information Processing*, 69–83.

Kumar, K. P. K., & Geethakumari, G. (2014). Detecting misinformation in online social networks using cognitive psychology. *Human-Centric Computing and Information*

Sciences, *4*(1), 14. https://doi.org/10.1186/s13673-014-0014-x

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., … Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094 LP-1096. Retrieved from http://science.sciencemag.org/content/359/6380/1094.abstract

Levine, J. M., Resnick, L. B., & Higgins, E. T. (1993). Social foundations of cognition. *Annual Review of Psychology*, *44*(1), 585–612.

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era. *Journal of Applied Research in Memory and Cognition*, *6*(4), 353–369. https://doi.org/https://doi.org/10.1016/j.jarmac.2017.07.008

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction continued influence and successful debiasing. *Psychol Sci Public Interest*, *13*. https://doi.org/10.1177/1529100612451018

Mai, J.-E. (2016). *Looking for information: A survey of research on information seeking, needs, and behavior*. Emerald Group Publishing.

Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online. *New York: Data & Society Research Institute*.

McCright, A. M., & Dunlap, R. E. (2017). Combatting Misinformation Requires Recognizing Its Types and the Factors That Facilitate Its Spread and Resonance. *Journal of Applied Research in Memory and Cognition*, *6*(4), 389–396. https://doi.org/https://doi.org/10.1016/j.jarmac.2017.09.005

Messing, S., & Westwood, S. J. (2012). Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation When Selecting News Online. *Communication Research*, *41*(8), 1042–1063. https://doi.org/10.1177/0093650212466406

Mocanu, D., Rossi, L., Zhang, Q., Karsai, M., & Quattrociocchi, W. (2015). Collective attention in the age of (mis)information. *Computers in Human Behavior*, *51*, 1198–1204. https://doi.org/https://doi.org/10.1016/j.chb.2015.01.024

Neisser, U. (1976). *Cognition and reality. Principles and Implications of Cognitive Psychology*. San Francisco: WH Freeman and Company.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Parlapiano, A., & Lee, J. C. (2018, February 16). The Propaganda Tools Used by Russians to Influence the 2016 Election. *The New York Times*. Retrieved from https://nyti.ms/2C7mFpE

Passe, J., Drake, C., & Mayger, L. (2018). Homophily, echo chambers, & selective exposure in social networks: What should civic educators do? *Journal of Social Studies Research*. https://doi.org/10.1016/j.jssr.2017.08.001

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, *147*(12), 1865–1880. https://doi.org/10.1037/xge0000465

Pennycook, G., & Rand, D. G. (2018a). DNU - Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*. https://doi.org/https://doi.org/10.1016/j.cognition.2018.06.011

Pennycook, G., & Rand, D. G. (2018b). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*. https://doi.org/https://doi.org/10.1016/j.cognition.2018.06.011

Price, V., Tewksbury, D., & Powers, E. (1997). Switching trains of thought: The impact of news frames on readers' cognitive responses. *Communication Research*, *24*(5), 481–506.

Rapp, D. N., & Donovan, A. M. (2017). Routine Processes of Cognition Result in Routine

Influences of Inaccurate Content. *Journal of Applied Research in Memory and Cognition*, *6*(4), 409–413. https://doi.org/https://doi.org/10.1016/j.jarmac.2017.08.003

Romano, A. (2018, October 19). Twitter released 9 million tweets from one Russian troll farm. Here's what we learned. *Vox*. Retrieved from https://www.vox.com/2018/10/19/17990946/twitter-russian-trolls-bots-election-tampering

Ross, A. S., & Rivers, D. J. (2018). Discursive Deflection: Accusation of "Fake News" and the Spread of Mis- and Disinformation in the Tweets of President Trump. *Social Media + Society*, *4*(2), 2056305118776010. https://doi.org/10.1177/2056305118776010

Schmidt, A. L., Zollo, F., Del Vicario, M., Bessi, A., Scala, A., Caldarelli, G., … Quattrociocchi, W. (2017). Anatomy of news consumption on Facebook. *Proceedings of the National Academy of Sciences*, *114*(12), 3035 LP-3039. Retrieved from http://www.pnas.org/content/114/12/3035.abstract

Sears, D. O., & Freedman, J. L. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly*, *31*(2), 194–213.

Silverman, C., Strapagiel, L., Shaban, H., Hall, E., & Singer-Vine, J. (2016, October 20). Hyperpartisan Facebook pages are publishing false and misleading information at an alarming rate. *Buzzfeed*.

Smith, E. R. (1994). Procedural knowledge and processing strategies in social cognition. *Handbook of Social Cognition*, *1*, 99–151.

Smith, E. R. (1998). Mental representation and memory.

Sotirovic, M., & McLeod, J. M. (2004). Knowledge as understanding: The information processing approach to political learning. *Handbook of Political Communication Research*, 357–394.

Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior*, *30*(3), 341–366.

Sunstein, C. R. (2007). Republic. com 2.0. Princeton, NJ: Princeton University Press.

Tandoc, E. C., Ling, R., Westlund, O., Duffy, A., Goh, D., & Zheng Wei, L. (2017). Audiences' acts of authentication in the age of fake news: A conceptual framework. *New Media & Society*, *20*(8), 2745–2763. https://doi.org/10.1177/1461444817731756

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146 LP-1151. Retrieved from http://science.sciencemag.org/content/359/6380/1146.abstract

Wakabayashi, D., & Isaac, M. (2017, January 25). In race against fake news, Google and Facebook stroll to the starting line. *New York Times*.

Wardle, C. (2017). Fake news. It's complicated. *First Draft News*.

Wardle, C., & Derakhshan, H. (2017). Information Disorder: Toward an interdisciplinary framework for research and policymaking. *Council of Europe Report, DGI (2017)*, *9*.

Weedon, J., Nuland, W., & Stamos, A. (2017). Information operations and Facebook. *Version*, *1*, 27.

Appendix 1: R Code for Analysis

```
#Loading in a bunch of Libraries
library(tidyverse)
library(rtweet)
library(ggplot2)
library(dplyr)
library(tidytext)
library(tm)
library(igraph)
library(ggraph)

#Loading the full IRA Tweet Data Set
IRA_Tweets <- read_csv("ira_tweets_csv_hashed.csv")

#Filtering out all of the tweets that are not in English
IRA_Tweets_en <- IRA_Tweets %>%
  filter(tweet_language=="en")

#Saving the English Tweet Set
write_csv(IRA_Tweets_en,'IRA_Tweets_en.csv')

IRA_Tweets_en <- read_csv("IRA_Tweets_en.csv")


#Narrowing down IRA_Tweets_en variables
IRA_Tweets_en_Clean <- data.frame(
  IRA_Tweets_en$user_display_name,
  IRA_Tweets_en$follower_count,
  IRA_Tweets_en$following_count,
  IRA_Tweets_en$tweet_text,
  IRA_Tweets_en$tweet_time,
  IRA_Tweets_en$is_retweet,
  IRA_Tweets_en$tweet_client_name,
  IRA_Tweets_en$hashtags,
  IRA_Tweets_en$urls)

#Removing http elements, RT, and @___ from words list
IRA_Tweets_en_Clean$stripped_text <- gsub("http\\S+","",
IRA_Tweets_en_Clean$IRA_Tweets_en.tweet_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("RT ","",  IRA_Tweets_en_Clean$stripped_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("@\\S+","",
IRA_Tweets_en_Clean$stripped_text)

write_csv(IRA_Tweets_en_Clean,'IRA_Tweets_en_Clean.csv')

###Start of Actual Analysis###
IRA_Tweets_en_Clean <- read_csv("IRA_Tweets_en_Clean.csv")

#Creating a clean list of words for analysis
IRA_Tweets_Words <- IRA_Tweets_en_Clean %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(word, stripped_text)

#Removing stop words from list of words
data("stop_words")

Clean_IRA_Tweet_Words <- IRA_Tweets_Words %>%
  anti_join(stop_words)

#Making plot of Top 20 words.
Clean_IRA_Tweet_Words %>%
```

```
  count(word, sort = TRUE) %>%
  top_n(20) %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(x = word, y = n)) +
  geom_col(fill = "royalblue4",
           colour = "grey100") +
  geom_text(aes(label = n), vjust = 1.8, size = 3) +
  xlab(NULL) +
  coord_flip() +
  labs(y = "Count",
       x = "Unique words",
       title = "Top 20 Words in IRA Tweets",
       subtitle = "Count of unique words")

#change scientific notation
options(scipen=6)

###Exploring Co-Occurence of Words###

#Re-run this...
IRA_Tweets_en_Clean$stripped_text <- gsub("http\\S+","",
IRA_Tweets_en_Clean$IRA_Tweets_en.tweet_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("RT ","",  IRA_Tweets_en_Clean$stripped_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("@\\S+","",
IRA_Tweets_en_Clean$stripped_text)

#Installing and loading some more libraries
install.packages(devtools)
install_github("dgrtwo/widyr", force = TRUE)
library(devtools)
library(widyr)

#Identifying Paired Words
ira_tweets_paired_words <- IRA_Tweets_en_Clean %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(paired_words, stripped_text, token = "ngrams", n = 2)

ira_tweets_paired_words %>%
  count(paired_words, sort = TRUE)

#Loading another Library
library(tidyr)

#Separating the words into two columns
IRA_Tweets_Separated_Words <- ira_tweets_paired_words %>%
  separate(paired_words, c("word1", "word2"), sep = " ")

IRA_Tweets_Filtered <- IRA_Tweets_Separated_Words %>%
  filter(!word1 %in% stop_words$word) %>%
  filter(!word2 %in% stop_words$word)

IRA_Words_Counts <- IRA_Tweets_Filtered %>%
  count(word1, word2, sort = TRUE)

#Loading in some more libraries
library(igraph)
library(ggraph)

#Word Network Plot
IRA_Words_Counts %>%
```

```
  filter(n >= 2000) %>%
  graph_from_data_frame() %>%
  ggraph(layout = "dh") +
  geom_edge_link(color = "royalblue4", aes(edge_alpha = n, edge_width = n, colour = n))
+
  geom_node_point(color = "royalblue4", size = 3) +
  geom_node_text(aes(label = name), vjust = 1.8, size = 3) +
  labs(title = "Word Network: Paired Words",
       subtitle = "IRA Twitter Dataset",
       x = "", y = "")

#playing with graph
# IRA_Words_Counts %>%
#   filter(n >= 2000) %>%
#   graph_from_data_frame() %>%
#   ggraph(layout = 'kk') +
#   geom_edge_density(aes(fill = n)) +
#   geom_edge_link(aes(edge_alpha = n, edge_width = n, colour = n)) +
#   geom_node_text(aes(label = name), vjust = 1.8, size = 3) +
#   labs(title = "Word Network: Paired Words",
#        subtitle = "IRA Twitter Dataset",
#        x = "", y = "")
#Re-run this...
IRA_Tweets_en_Clean$stripped_text <- gsub("http\\S+","",
IRA_Tweets_en_Clean$IRA_Tweets_en.tweet_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("RT ","",  IRA_Tweets_en_Clean$stripped_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("@\\S+","",
IRA_Tweets_en_Clean$stripped_text)

#Network of 3 words
IRA_Tweets_Triad_Words <- IRA_Tweets_en_Clean %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(triad_words, stripped_text, token = "ngrams", n = 3)

IRA_Tweets_Triad_Words %>%
  count(triad_words, sort = TRUE)

IRA_Tweets_Separated_Triad_Words <- IRA_Tweets_Triad_Words %>%
  separate(triad_words, c("word1", "word2", "word3"), sep = " ")

IRA_Tweets_Filtered_3 <- IRA_Tweets_Separated_Triad_Words %>%
  filter(!word1 %in% stop_words$word) %>%
  filter(!word2 %in% stop_words$word) %>%
  filter(!word3 %in% stop_words$word)

IRA_Triad_Words_Counts <- IRA_Tweets_Filtered_3 %>%
  count(word1, word2, word3, sort = TRUE)

IRA_Triad_Words_Counts %>%
  filter(n >= 500) %>%
  graph_from_data_frame() %>%
  ggraph(layout = "dh") +
  geom_edge_link(color = "royalblue4", aes(edge_alpha = n, edge_width = n, colour = n))
+
  geom_node_point(color = "royalblue4", size = 3) +
  geom_node_text(aes(label = name), vjust = 1.8, size = 3) +
  labs(title = "Word Network: Three Word Analysis",
       subtitle = "IRA Twitter Dataset",
       x = "", y = "")
```

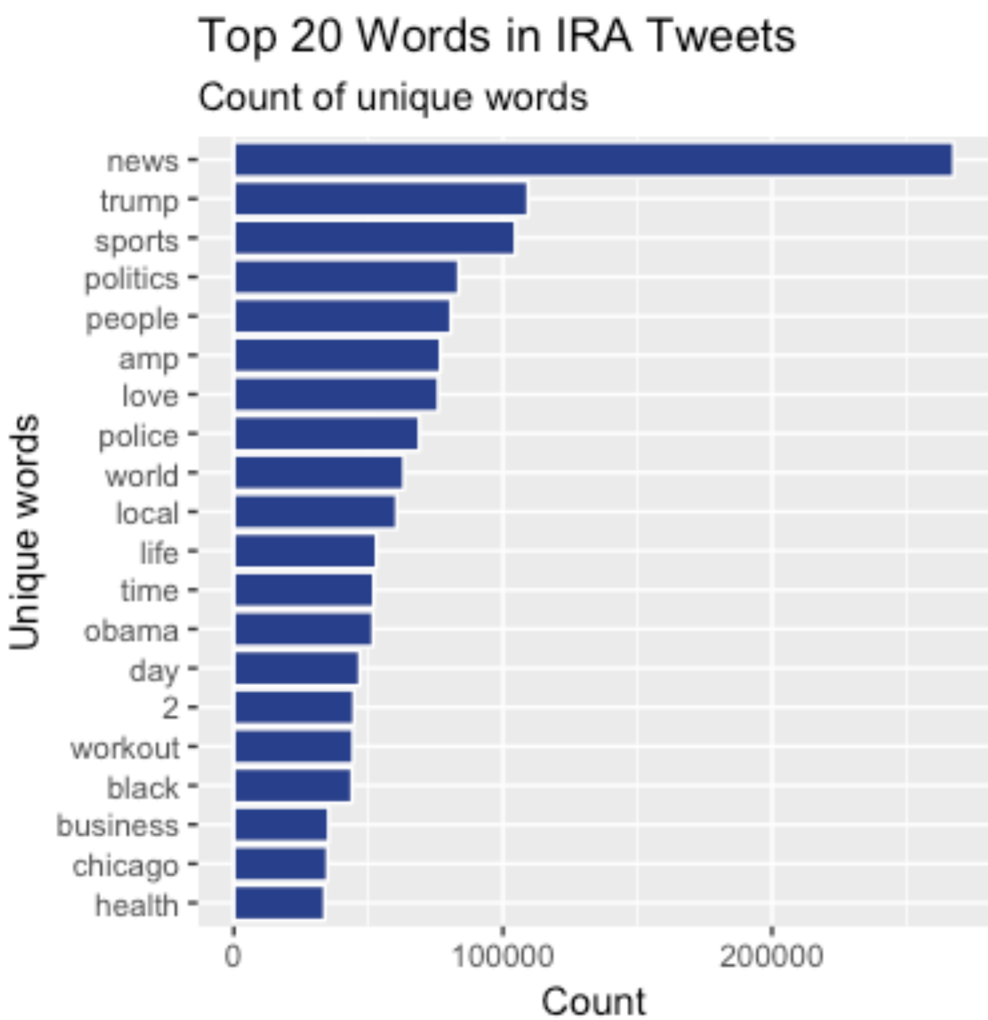
```
#Re-run this...
IRA_Tweets_en_Clean$stripped_text <- gsub("http\\S+","",
IRA_Tweets_en_Clean$IRA_Tweets_en.tweet_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("RT ","",  IRA_Tweets_en_Clean$stripped_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("@\\S+","",
IRA_Tweets_en_Clean$stripped_text)

#Network of 5 words
IRA_Tweets_Five_Words <- IRA_Tweets_en_Clean %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(five_words, stripped_text, token = "ngrams", n = 5)

IRA_Tweets_Five_Words %>%
  count(five_words, sort = TRUE)

IRA_Tweets_Separated_Five_Words <- IRA_Tweets_Five_Words %>%
  separate(five_words, c("word1", "word2", "word3", "word4", "word5"), sep = " ")

IRA_Tweets_Filtered_5 <- IRA_Tweets_Separated_Five_Words %>%
  filter(!word1 %in% stop_words$word) %>%
  filter(!word2 %in% stop_words$word) %>%
  filter(!word3 %in% stop_words$word) %>%
  filter(!word4 %in% stop_words$word) %>%
  filter(!word5 %in% stop_words$word)

IRA_Five_Words_Counts <- IRA_Tweets_Filtered_5 %>%
  count(word1, word2, word3, word4, word5, sort = TRUE)

IRA_Five_Words_Counts %>%
  filter(n >= 100) %>%
  graph_from_data_frame() %>%
  ggraph(layout = "dh") +
  geom_edge_link(color = "royalblue4", aes(edge_alpha = n, edge_width = n, colour = n))
+
  geom_node_point(color = "royalblue4", size = 3) +
  geom_node_text(aes(label = name), vjust = 1.8, size = 3) +
  labs(title = "Word Network: Five Word Analysis",
       subtitle = "IRA Twitter Dataset",
       x = "", y = "")


### Beginning Topic Modelling Analysis of Key Topics ###

#Just look at major concepts here... will definitely need to go back and fine tune this
code/ clean it up/
  #go even more in depth with analysis... hashtags, links, etc.


## List of Topics and Words by Level ##
    # One Word Analysis
        # News
        # Trump
        # Sports
        # Politics
        # People
        # Love
        # Police
        # World
        # Local
```

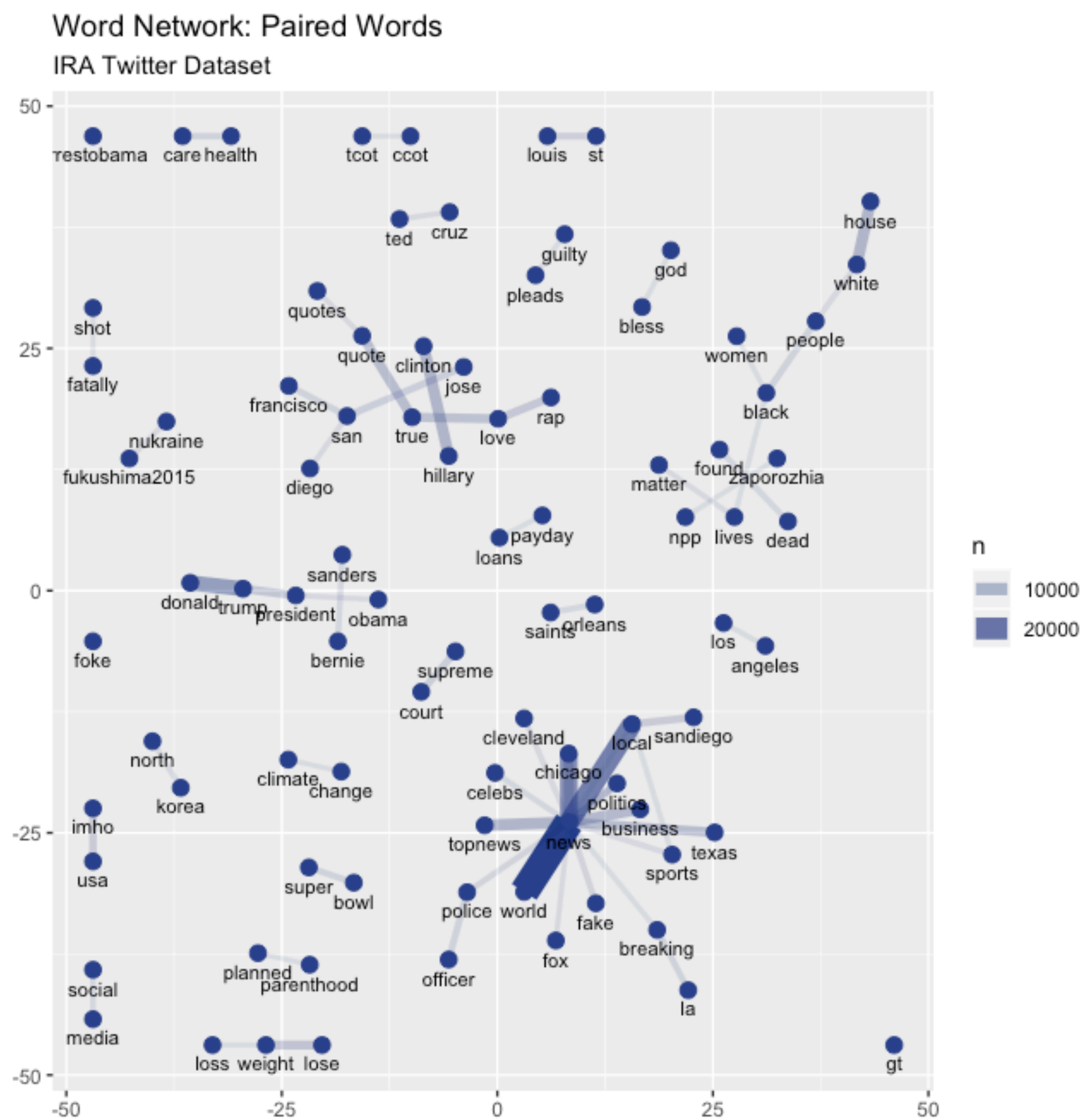
```
        # Life
        # Time
        # Obama
        # Day
        # Workout
        # Black

    # Two Word Analysis
        # Chicago World News
        # 24x7 Online Payday Loans
        # AmericaFirst MakeAmeericaGreatAgain
        # Black Lives
        # Sports Local

    # Three Word Analysis:
        # News World Local TopNews Breaking Donald
        # AmericaFirst MakeAmeericaGreatAgain
        # 24x7 Online Payday Loans
        # Black Lives
        # Isis Targeted Accounts IceIsis
        # Goveernment Doesnt Care Nuclear Power Ukranian

    # Five Word Analysis:
        # 24x7 Online Payday Loans
        # Ukranian Government
        # Targeted Isis Accounts
        # Maga War WakeUpUSA DeepState ShadowGovernment Obamagate
        # MakeAmericaGreatAgain AmericaFirst

# Loading in Libraries
install.packages("topicmodels")
install.packages("qdap")

library(tm)
library(tidyverse)
library(tidytext)
library(topicmodels)
library(qdap)


#
# hc <- IRA[str_detect(IRA$IRA_Tweets_en.tweet_text,"@HillaryClinton"),]
#
# IRA %>% group_by(IRA_Tweets_en.tweet_text) %>% summarize(count=n()) %>%
arrange(desc(count)) %>% head()

# Pulling out 500 to play with...
#corp.df <-
cbind(as.character(1:500),IRA[base::sample(1:nrow(IRA),500),"IRA_Tweets_en.tweet_text"])

# names(corp.df) <- c("doc_id","text")
#
# rm(IRA_1000)
# gc()
#
# write_csv(corp.df,"sample.csv")

#Reloading Dataset
IRA_Tweets_en_Clean <- read_csv("IRA_Tweets_en_Clean.csv")
```

```
#Re-run this...
IRA_Tweets_en_Clean$stripped_text <- gsub("http\\S+","",
IRA_Tweets_en_Clean$IRA_Tweets_en.tweet_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("RT ","",  IRA_Tweets_en_Clean$stripped_text)
IRA_Tweets_en_Clean$stripped_text <- gsub("@\\S+","",
IRA_Tweets_en_Clean$stripped_text)


# #Filtering out so that the five word topics are represented
# Filter_Five <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("ukranian","government"), collapse = "|"))
# Filter_Five <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("targeted", "isis", "accounts"), collapse = "|"))
# Filter_Five <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("maga", "war", "wakeupusa", "deepstate", "shadowgovernment",
#
"obamagate"), collapse = "|"))
# Filter_Five <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("makeamericagreatagain","americafirst"), collapse = "|"))

#Filtering based on main topics
Filter_BlackLives <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("black", "lives"), collapse = "|"))
Filter_Obama <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("obamagate", "obama"), collapse = "|"))
Filter_MakeAmericaGreat <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("makeamericagreatagain","americafirst", "maga"), collapse = "|"))
Filter_Trump <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("trump", "donald"), collapse = "|"))
Filter_DeepState <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("wakeupusa", "deepstate", "shadowgovernment","obamagate"), collapse = "|"))
Filter_Hillary <- str_detect(str_to_lower(IRA_Tweets_en_Clean$stripped_text),
paste(c("hillary", "clinton"), collapse = "|"))

# Setting Stripped Text as Corpus
data("stop_words")
c <- Corpus(VectorSource(IRA_Tweets_en_Clean$stripped_text[Filter_BlackLives]))

dtm <- DocumentTermMatrix(c,control = list(removePunctuation = TRUE,
                                           stopwords = TRUE,
                                           stemming = TRUE,
                                           wordLengths=c(0,Inf)))

# #Checking the dtm
# inspect(dtm)
#
# #Filtering out rows that are zero and making a new corpus
# clean.corp <- IRA_Tweets_en_Clean$stripped_text[!(which(rowSums(as.matrix(dtm))==0))]
# c <- Corpus(VectorSource(clean.corp))

#Ten Topic LDA on dtm
base_lda <- LDA(dtm, k=5, method = "VEM", control = NULL, model = NULL)

topics_beta <- tidy(base_lda, matrix = "beta")

top_terms <- topics_beta %>%
  group_by(topic) %>%
  top_n(5, beta) %>%
  ungroup() %>%
  arrange(topic, -beta)
```

```
as.data.frame(top_terms)

lda_gamma <- tidy(base_lda, matrix = "gamma")
tweet_class <- lda_gamma %>%  group_by(document) %>%
  top_n(1, gamma) %>% arrange(as.numeric(document)) %>% ungroup()

corp.df <- data.frame("text"=IRA_Tweets_en_Clean$stripped_text[Filter_BlackLives])
corp.df$doc_id <- as.character(1:nrow(corp.df))

tweet_class_text_BL <- merge(tweet_class,corp.df,by.x="document",by.y="doc_id") %>%
arrange(as.numeric(document))

View(tweet_class_text_BL %>% arrange(desc(gamma)))

table(tweet_class_text$topic)
```

Appendix 2: Large Scale Images of Figures

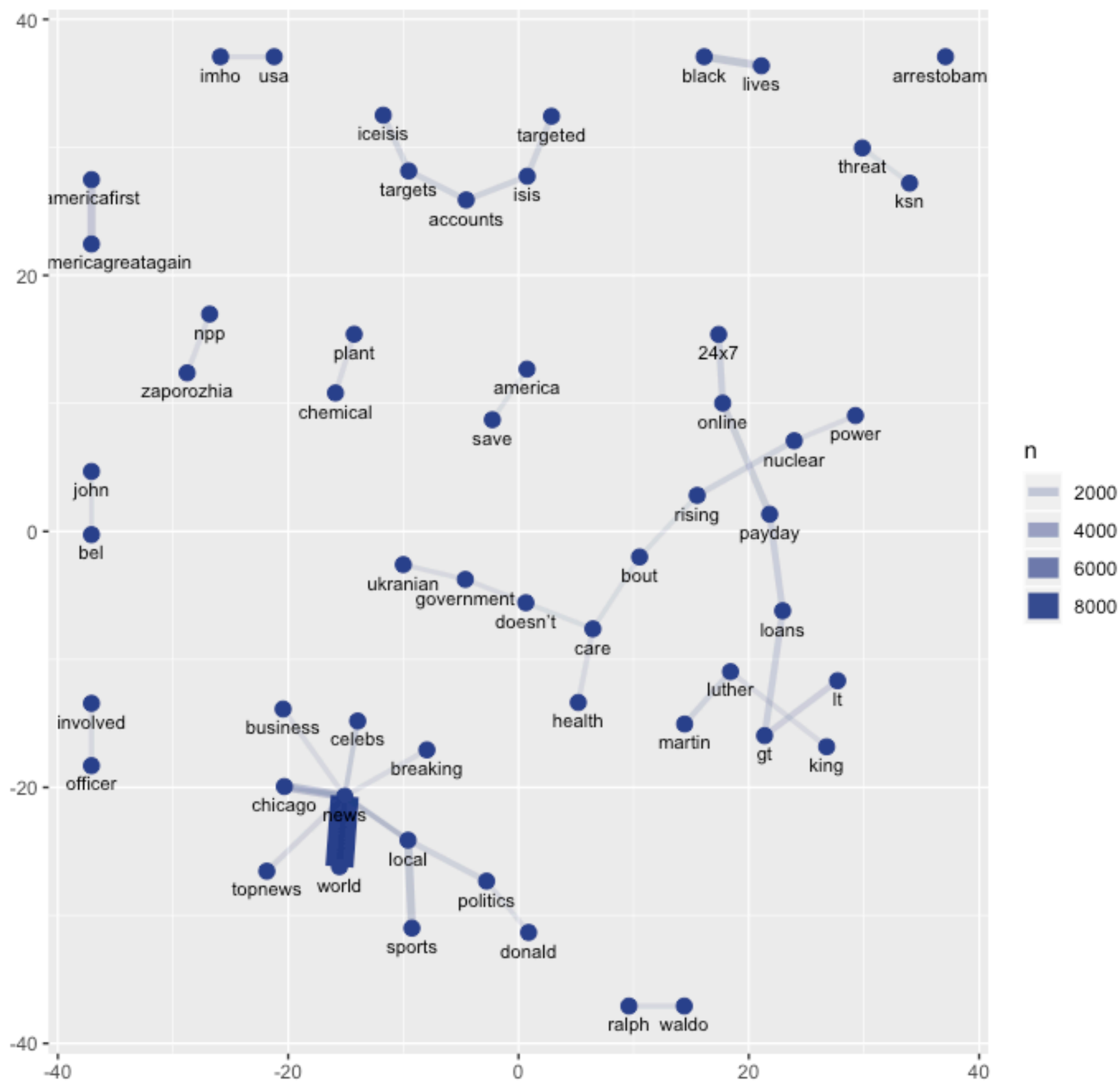




Word Network: Paired Words
IRA Twitter Dataset
restobama
care
health
tcot
ccot
louis
st
ted
cruz
house
guilty
god
white
pleads
bless
shot
quotes
people
women
quote
clinton
jose
fatally
francisco
black
rap
san
true
love
nukraine
hillary
found
zaporozhia
fukushima2015
matter
diego
payday
npp
lives
dead
loans
sanders
donald
trump
president
obama
saints
orleans
los
angeles
foke
bernie
supreme
court
cleveland
local
sandiego
north
chicago
climate
change
celebs
korea
politics
imho
business
topnews
news
texas
sports
usa
super
bowl
police
world
fake
fox
breaking
officer
social
planned
parenthood
la
media
loss
weight
lose
gt
n
10000
20000
50
25
0
-25
-50

Word Network: Three Word Analysis

IRA Twitter Dataset



Word Network: Five Word Analysis

IRA Twitter Dataset