# **How ISIS Uses Twitter**

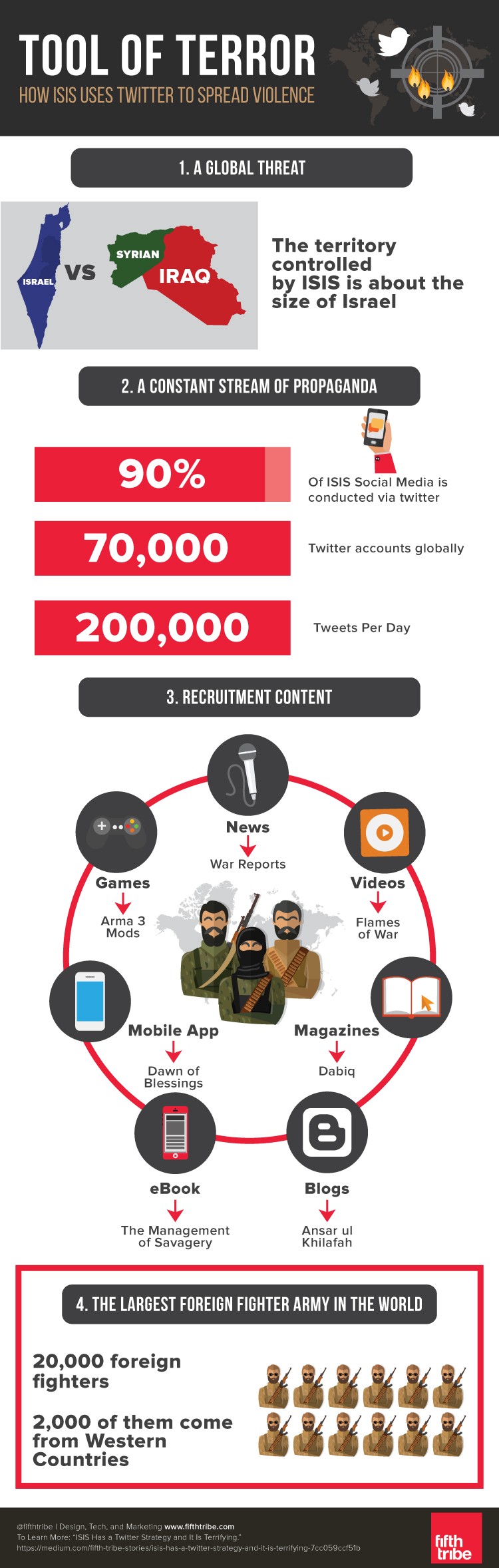
A data story for Springboard.com’s class “Introduction to Data Science” in R.

Alexander M. Babenchuk  
December 4th, 2016

**​Why?**

Terrorist organization ISIS (“Islamic State of Iraq and the Levant”) is referred to by many names such as ISIL, IS and Da’esh. These acronyms refer to different stages of the same infamous terrorist group which is occupying Iraq and Syria. Unlike ISIS’ predecessors such as Al-Qaida, ISIS has been extremely successful in infiltrating modern social media as a broadcast channel for its propaganda and to attract new recruits or jihadists. The group has been especially successful in radicalizing westerners into supporting, aiding, or performing acts of terror. Twitter is considered one of the primary methods ISIS uses to recruit new jihadists. It is estimated that ISIS had over 46,000 to 70,000 accounts on Twitter alone.

Originally, Twitter, Inc. was slow to react, since social media corporations support free speech and essentially do not want to pick a side. However, a good deal of content on Twitter surpasses free-speech and is considered by counter-terrorism groups as unlawful activity. However, due to public pressure Twitter ramped up the company’s efforts and has shut down over 360,000 accounts which is not an easy task considering the speed at which pro-ISIS supporters open and switch to new accounts.

  
*Infographic excerpt from FifthTribe.*

The goal of my project is to analyze the 17,000 twitter messages made available on Kaggle.com as part of FifthTribe’s effort to prevent radicalization, counter-message pro-ISIS tweets and limit the organization’s ability to broadcast, radicalize and recruit on Twitter. To highlight the importance of programs such as this you need to understand that this problem affects us all, no matter how far we are from the battle fields of Iraq and Syria.

Ali Shukri Amin was Northern Virginia seventeen-year-old high school student who was sentenced by a U.S. District court to over eleven years in prison. Amin ran a pro-Islamic State Twitter account, “@AmreekiWitnesss”. Amin had more than 4,000 followers including active fighters and recruiters in Syria and Iraq; tweeted over 7,000 messages in support of Islamic State and helped his high school friend, Reza Niknejad, an eighteen-year-old friend connect and become a part of ISIS in Syria. In a Washington Post article Amin described himself “a troubled kid who lost himself while seeking acceptance and respect in a sinister, virtual world.” However, the story of Amin is just one of hundreds, if not thousands of young westerners who are radicalized by ISIS every year. ISIS recruiters use social media such as Twitter to beginning with what seems as innocent conversations about Islam and the teachings of Muhammed. Eventually these conversations turn into Anti-American or Anti-Western ideology in an effort to recruit the individual.

  
*Ali Shukri Amin’s picture from Reuters.com*

**Who Cares?**

There is an array of government agencies, counter-terrorism groups and non-for-profit organizations which are intensely interested in eliminating or preventing radicalization by ISIS or similar terrorist organizations. For example, Google's Think Tank "Jigsaw" analyzed similar data sets and launched a project called “The Redirect Method” which according to Jigsaw’s website “is focused on reaching those who are actively looking for extremist content and connections. Rather than create new content and counter-narratives, [their] approach tries to divert young people off the path to extremism using pre-existing YouTube content and targeted advertising.”

On the opposite side of the spectrum is Anonymous, a self-proclaimed hacker vigilantes group which launched a campaign on Twitter called “OpsISIS”. The campaign aims at identifying and tagging pro-ISIS supporters for Twitter to disable the accounts. This campaign also generated thousands of counter Pro-ISIS messages which make it more difficult to find content for would be supporters of the terrorist organization.

**The Data**

The data set is very simple and is comprised of the self-assigned name, username, user’s Twitter biography, location if one was shared, number of followers at the time the tweet was downloaded, number of statuses by the user when the tweet was downloaded, date and timestamp of the tweet, and lastly the tweet itself. In total 17, 410 observations and eight variables.

**Limitations**

Twitter usernames of pro-ISIS fanboys are constantly changing because Twitter shuts them down. ​Thus, we cannot analyze long periods of tweet history and its effect on the user’s influence. It would be also beneficial if the data set contained the user's avatar. An avatar is a picture a twitter users assigns to his or her account as a way to graphically describe themselves. These are avatars are often re-cycled when a username is switched and can be used to identify and link new username to disabled account. Below is an excerpt of how to identify pro-ISIS users by [Khuram Zaman](https://medium.com/@zamandigital?source=post_header_lockup) of FifthTribe, the creator of the data set:

“Indicators of someone being pro-ISIS included: (a) keywords in the user’s name, description or tweets such as “Dawla” (which refers to the State), “Baqiyyah” (which denotes being part of the supposed ‘ever-expanding’ of the ISIS state), “Amaq” (the agency used by ISIS to issue official proclamations, “Wilayat” (used by ISIS to divide up the world into its provinces), etc. I also looked at imagery such as if a user had the ISIS flag or images of radical leaders like al-Baghdadi, Anwar Awlaki, etc. I also looked at who they were following and following them back. All of these factors helped me identify factors to collect the data.”  
(Source: http://blog.kaggle.com/2016/06/03/dataset-spotlight-how-isis-uses-twitter/)

I also compared this dataset to the information available when pulling tweets via the newly created “rtweet” library and realized that the raw data is much better suited than the published dataset for exploration. However, by the time FifthTribe’s data set was made available much of the original content has been removed by Twitter so I was not able to re-pull the data.

Another obvious limitation can be seen by analyzing the month and year of the data set. This clearly shows that the majority of information is in 2016 between January and May. To properly analyze ISIS’s affect we would need to monitor the tweet activity over a larger period of time.

**Tweet distribution by month and year**

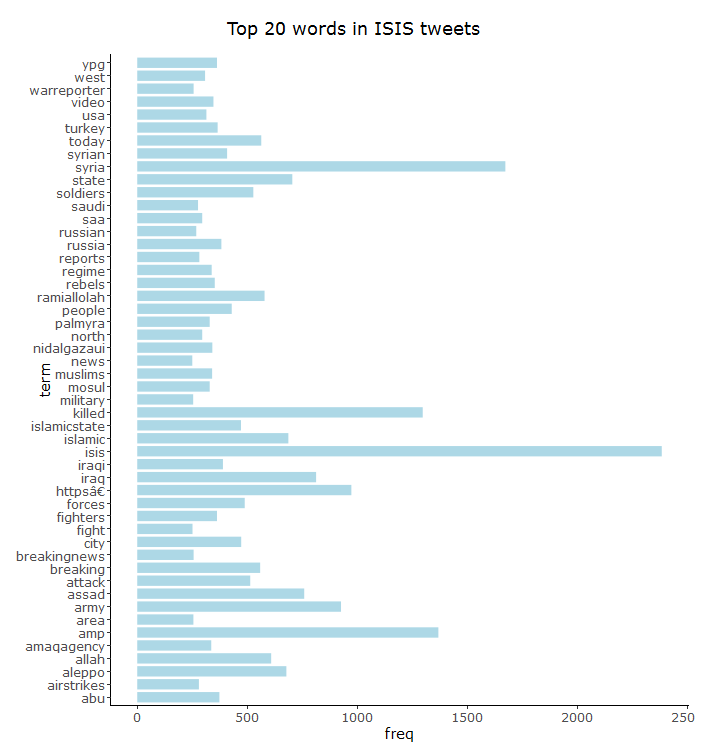


**Cleaning**

I used the “tm”, “dplyr”, “stinger” and “lubridate” packages to remove stopwords (i.e. the, is, and, etc.), I also removed links, retweets, hashtags, @people, punctuations, numbers, emojis and non-English characters.

For preliminary exploration, I looked at the data set tweets to gather what the dataset looks like. Once I understood the format and structure, I wanted to know who are the most active twitter users in this dataset, hashtags and words. You quickly begin to see that a handful of users are responsible for much of the content. You also see that the hashtags used are very general and would appear in pro-ISIS content as well as general publications about the war in Syria and Iraq.

**Word Cloud of Top Words and Users**



We can quickly see from the above to visualizations that much of the pro-ISIS content is being published as if it was a news source. This is supported by usernames such as “warrnews, WarReporter1” and hashtags “#breakingnews, #breaking, #amaqagency and #news.” Below is a data frame showing top users from the data set; included is the number of followers, number of tweets and percent of tweets.

|  |  |  |  |
| --- | --- | --- | --- |
| **username** | **n\_followers** | **n\_tweets** | **% of tweets** |
| Uncle\_SamCoco | 1809 | 1580 | 18% |
| RamiAlLolah | 34692 | 1475 | 17% |
| warrnews | 7566 | 1191 | 14% |
| WarReporter1 | 2632 | 1095 | 13% |
| mobi\_ayubi | 965 | 1056 | 12% |
| \_IshfaqAhmad | 1667 | 709 | 8% |
| wayf44rerr | 1398 | 405 | 5% |
| Nidalgazaui | 4662 | 397 | 5% |
| MaghrabiArabi | 353 | 366 | 4% |
| melvynlion | 81 | 348 | 4% |

**Top Hashtags**

|  |  |
| --- | --- |
| **tweets.hash** | **n\_hashtags** |
| #ISIS | 724 |
| #IS | 398 |
| #Syria | 348 |
| #AmaqAgency | 283 |
| #BreakingNews | 244 |
| #IslamicState | 180 |
| #Iraq | 172 |
| #Breaking | 164 |
| #Aleppo | 129 |
| #USA | 125 |
| #Russia | 119 |
| #BREAKING | 111 |

We can also see that a handful of users are posting majority of the content and have significant influence. Based on the chart below we can see that one user stands-out, RamiAlLolah. This is additionally supported by [Khuram Zaman](https://medium.com/@zamandigital?source=post_header_lockup) of FifthTribe who stated that:

The most interesting insight I learned was how a handful of users are essentially thought leaders and influence much of the network. I thought it would be a bit more evenly spread across the network, but its clear that 3-5 people are generating the lion’s share of content and serve as connectors between content producers and content receivers.  
(Source: http://blog.kaggle.com/2016/06/03/dataset-spotlight-how-isis-uses-twitter)





The cluster dendrogram above shows the association of words based on their cosine distance.

**Result**

Originally, I believed that it will be easy to identify pro-ISIS tweets and content on Twitter today, however, by attempting to search Twitter for the top hashtags or users from this dataset I was not able to find any new pro-ISIS content, instead the most frequent words and hashtags point to mainstream media and content. Additionally, many of the trigger words are no longer used by pro-ISIS supporters because of campaigns such as “#OpsISIS” content aimed to discourage radicalization. This clearly signals that campaigns against ISIS have been successful.

I also wanted to analyze the users who are still posting today, however all of the accounts have been disabled.

Is there any link between attacks and tweets?