Time Series Analysis of Hate Speech in Social Media

David J Corliss, Peace-Work, Plymouth, MI

ABSTRACT

Mining and analysis of social media text is a powerful tool for the analysis of thoughts, preferences, and actions of individuals and populations. While commonly used today in marketing and other business applications, Data for Good researchers have begun to apply these methods to the analysis of hate speech in social media. Different channels, including Twitter, Facebook, and Google searches, are found to have distinctive characteristics that affect the types of models and analyses each can support. This paper provides a step-by-step description of how to create and deploy a Twitter API and mine the data to extract tweets with user-selected search terms, including keywords and user name of the person or organization sending the tweet. This methodology is used to investigate hate speech, modeling the time series patterns with the aim of estimating the risk of subsequent acts of violence against persons targeted by the speech. These Data For Good analyses have been performed using SAS® University Edition, a free version of SAS® available to students, professors, and non-profit researchers.

INTRODUCTION

APIs now exist for reading social media from a number of important outlets, including Twitter and Facebook. Twitter considers posts on it site to be a public conversation - not exclusively between friends and connections – and supports some limited search capable through their website. Twitter also actively encourages developers to create APIs for user-written applications. A macro for using an API to search Twitter was described by in a paper by Isabel Litton and Rebecca Ottesen of California Polytechnic State University, presented at the Western Users of SAS conference in 2013. Earlier work by the author presented lag models analyzing hate speech on Twitter, investigating a possible connection between social media hate speech and subsequent acts of violence against persons targeted in the speech. While the earlier work focused on time series analysis using tweets sent over a period of several years. The present paper describes real-time analysis using maps dynamically updated at a user-selected time interval, with a period of 10 seconds used in the example.

SOCIAL MEDIA DATA

Different social media channels are found to have distinctive characteristics that affect the kinds of uses and questions that can be answered using the posts. Perhaps the most important distinction is between public and private media. Channels such as Facebook and LinkedIn are considered private because they require a personal connection to all parties seeing a post. This is usually mediated through the use of an invitation and acceptance system. By contrast, some other social media platforms are effectively public, allowing any subscriber on the platform to see all posts. Twitter in particular emphasizes this aspect, making a tweet the social media equivalent of sky writing: anyone (more exactly, any Twitter subscriber) can read any tweet without a personal connection to the tweeter. In this way, Facebook acts like a conversation with a friend in your living room: a private space, only accessible to people you have invited in. By contrast, Twitter is a like a conversation with a friend that occurs on a street corner: a public place where anyone passing by can hear everything said.

Another important feature is the time frame of the action. This is not as absolute as the public / private dichotomy: while most social media platforms can be used for either rapid response or durable posts and extended conversations, most will have a time frame characteristic of the particular social media channel. For example:

- Twitter: short duration and highly reactive to other posts and events

- Facebook: more medium term, with extended conversation threads

- Google Trends: Not usually thought of in this space, it has many characteristics of a public social media platform. While individuals are not identified, specific search counts and rates are available

to the public. Because the searches captured in Google Trends shows what people are thinking about, considering, and planning to do in the future, it functions as a leading indicator - the internet equivalent of Durable Goods – and therefore may have predictive power than other channels.

At this time, most tools re able to explore only text, not images. YouTube and Instagram – the #2 and #3 Social Media channels – have no effective means to mine images at this time

MINING TWITTER DATA WITH SAS

Twitter allows independent developers to create their own applications with search capability using an API. available at no cost. Twitter encourages these new applications by making the API available at no cost to registered developers, providing a web page to register for an API, and instructions analyzed implementation. SAS Support has a page on importing tweets: http://blogs.sas.com/content/sascom/2013/12/12/how-to-import-twitter-tweets-in-sas-data-step-using-oauth-2-authentication-style/. An excellent paper on setting up a Twitter API in a SAS program has been written by Isabel Litton and Rebecca Ottesen of California Polytechnic State University. The tweets analyzed in the following section were pulled using a modification of the Litton-Ottesen macro %GrabTweet.

In the mining and analysis of Twitter feeds, a five-step process is recommended:

1. Register with Twitter as a developer to get access

2. Create an API to deploy search terms and receive tweets

3. Create a program to deploy the API, obtain tweets, and parse the Twitter stream into a SAS dataset

4. Perform exploratory data analysis on a few tweets to find the best search terms for the investigation

5. Extract and analyze a large number of tweets pulled using the preferred search terms

The process begins with registering with Twitter as an app developer at https://dev.twitter.com/oauth/overview. Twitter, which encourages users to develop and implement an API in their web applications, provides detailed instructions for requesting APIs, which are available at no financial cost to the user. Once the on-line registration process is complete, Twitter will send three important articles by email:

- consumerKey – this serves as a Twitter User ID

- consumerSecret – this is the login password to the Twitter system

- Bearer Token: This combines the roles of consumerKey and consumerSecret to provide a single text string to be deployed in SAS code to access tweets.

It is strongly advised to keep all three highly secure. When deploying this information in SAS code, a FILENAME statement should be used to read this information from a secure location. These should not be hard-coded into SAS programs!

The following code can be used to deploy a registered Twitter developer ID to obtain tweets. It includes the parameter TYPE, which Twitter uses to determine the type of tweets to be extracted. Options for the TYPE parameter include

- Popular – the most popular tweets containing the selected search terms

- Recent – the most recent tweets containing the selected search terms

- Mixed – a mix of recent tweets and the most popular from further back in time

The first part of the Twitter access macro sets macro parameters extracts up to 100 tweets:

```
%MACRO grab_tweet_100(search_term, type, target_amount);


/* Location of the file with authentication information - edit this path */
filename auth "C:\dcorliss\SAS\Twitter\token.txt";


/* Location of the output file containing the tweets - edit this path */
filename twtOut "C:\dcorliss\SAS\Twitter\Tweets.txt";


/*
   These parameters need to be set when accessing the GET search/tweets url:
   COUNT = number of tweets to grab
   TYPE = Popular, Recent, or Mixed
*/


%IF &target_amount < 100 %THEN %LET num_tweet =
%NRSTR(&count=)&target_amount;
%ELSE %LET num_tweet = %NRSTR(&count=100);
%LET type = %NRSTR(&result_type=&type);


/* PROC HTTP to access the Twitter URL and write the output file */

PROC HTTP
   HEADERIN = auth
   METHOD = "get"

   /* Deploy the specified search terms in the Twitter search API */
  URL = "https://api.twitter.com/1.1/search/tweets.json?q=&search_term&type&num_tweet"

   /* Write the output to the designated location as one long text string */
   OUT = twtOut;
RUN;


%MEND;
```

While the text captured using the Twitter API is comma-delimited, there is no end of line marker: the text is transmitted as one continuous string:

```
Tweets.txt - Notepad
File  Edit  Format  View  Help
{"statuses":[{"created_at":"Thu Jun 08 08:52:55 +0000 2017","id":872738196777426944,"id_str":"872738196777426944","text":"@lsarsour You hate America &amp
ars","url":null,"entities":{"description":{"urls":[]}},"protected":false,"followers_count":2013,"friends_count":2511,"listed_count":26,"created_at":"Wed
idebar_fill_color":"DDEEF6","profile_text_color":"333333","profile_use_background_image":true,"has_extended_profile":false,"default_profile":true,"defaul
"recent"},"source":"\u003ca href=\"http:\/\/twitter.com\/#!\/download\/ipad\" rel=\"nofollow\"\u003eTwitter for iPad\u003c\/a\u003e","in_reply_to_status_
ges\/themes\/theme1\/bg.png","profile_background_image_url_https":"https:\/\/abs.twimg.com\/images\/themes\/theme1\/bg.png","profile_background_tile":fal
guess what...America is on to you and your sharia law promotion.  Your Alinsky tactics are transparent.","truncated":false,"entities":{"hashtags":[],"sym
one":null,"geo_enabled":true,"verified":false,"statuses_count":1015,"lang":"en","contributors_enabled":false,"is_translator":false,"is_translation_enable
s":null,"translator_type":"none"},"geo":null,"coordinates":null,"place":null,"contributors":null,"is_quote_status":false,"retweet_count":1,"favorite_coun
_reply_to_user_id":null,"in_reply_to_user_id_str":null,"in_reply_to_screen_name":null,"user":{"id":1216957579,"id_str":"1216957579","name":"Karen D. Scio
g.com\/profile_background_images\/799702831\/608c64c78f543cfda598bc9311553460.png","profile_background_image_url_https":"https:\/\/pbs.twimg.com\/profile
at":"Wed Jun 07 14:39:45 +0000 2017","id":872463091677540353,"id_str":"872463091677540353","text":"Agree our schools are run by commies &amp; traitors wh
n":"Libertarian Vice Presidential nominee, Best Selling Author, Youtube: https:\/\/t.co\/Cxln91HRMg WAR Now Radio: https:\/\/t.co\/WmcCUVjLrr","url":"htt
r":false,"is_translation_enabled":false,"profile_background_color":"222222","profile_background_image_url":"http:\/\/pbs.twimg.com\/profile_background_im
"place":null,"contributors":null,"is_quote_status":false,"retweet_count":36,"favorite_count":41,"favorited":false,"retweeted":false,"possibly_sensitive":
_reply_to_user_id_str":null,"in_reply_to_screen_name":null,"user":{"id":3290719405,"id_str":"3290719405","name":"Maine4Trumpism","screen_name":"Maine4Tru
s\/theme1\/bg.png","profile_background_image_url_https":"https:\/\/abs.twimg.com\/images\/themes\/theme1\/bg.png","profile_background_tile":false,"profil
k white men are most dangerous demographic, subscribe to primitive totalitaria\u2026 https:\/\/t.co\/r1CHYXsHsb","truncated":true,"entities":{"hashtags":
http:\/\/larrylockman.com","display_url":"larrylockman.com","indices":[0,23]}]},"description":{"urls":[]}},"protected":false,"followers_count":199,"frien
ink_color":"DD2E44","profile_sidebar_border_color":"000000","profile_sidebar_fill_color":"000000","profile_text_color":"000000","profile_use_background_i
2","indices":[95,107]},{"screen_name":"mainegop","name":"Maine GOP","id":22251316,"id_str":"22251316","indices":[120,129]},{"screen_name":"gehrig38","nam
te of Maine \ud83c\udf32 (DM US WITH TIPS) #mepolitics #mepol Recently Featured on @TuckerCarlson & @infowars","url":"https:\/\/t.co\/iuS68vcV1X","entiti
/7fqZZBA1_normal.jpg","profile_image_url_https":"https:\/\/pbs.twimg.com\/profile_images\/872561048095784961\/7fqZZBA1_normal.jpg","profile_banner_url":"
nt":0,"favorited":false,"retweeted":false,"lang":"en"},{"created_at":"Thu Jun 08 02:42:52 +0000 2017","id":872645070616489984,"id_str":"87264507061648998
coggin County GOP. Put Maine Citizens First! #mepolitics Award-Winner! https:\/\/t.co\/CItRSg8f8N","url":"https:\/\/t.co\/PjLJESlnmp","entities":{"url":{
file_background_tile":false,"profile_image_url":"http:\/\/pbs.twimg.com\/profile_images\/779527995501506560\/tqs6TWcE_normal.jpg","profile_image_url_http
,"truncated":true,"entities":{"hashtags":[],"symbols":[],"user_mentions":[],"urls":[{"url":"https:\/\/t.co\/r1CHYXsHsb","expanded_url":"https:\/\/twitter
ected":false,"friends_count":95,"created_at":"Wed Feb 19 03:09:53 +0000 2014","favourites_count":51,"utc_offset":-
color":"000000","profile_use_background_image":false,"has_extended_profile":false,"default_profile":false,"default_profile_image":false,"following":null,
[120,129]},{"screen_name":"gehrig38","name":"Curt Schilling","id":14945261,"id_str":"14945261","indices":[130,139]}],"urls":[{"url":"https:\/\/t.co\/NvM4
url":"https:\/\/t.co\/iuS68vcV1X","entities":{"url":{"urls":[{"url":"https:\/\/t.co\/iuS68vcV1X","expanded_url":"http:\/\/mainefirstmedia.wordpress.com",
qZZBA1_normal.jpg","profile_banner_url":"https:\/\/pbs.twimg.com\/profile_banners\/816402981851660288\/1496869729","profile_link_color":"ABB8C2","profile
9091027914753,"id_str":"872639091027914753","text":"RT @WayneRoot: Agree our schools are run by commies &amp; traitors who hate America &amp; Christianit
```

Figure 1: Raw Twitter Output

The long text string must be parsed. The comma-delimited fields appear in the following order:

- Timestamp at creation of original tweet
- Tweet ID (a long integer, successive, and unique to the tweet)
- Tweet text
- Twitter coded User ID
- User actual name
- User screen name (Twitter "handle")
- Location of user as registered on the Twitter account
- Timestamp for the creation of the Twitter account
- Language in which the tweet is written
- Retweet indicator

Code for parsing the tweets should be inserted into the macro. While the Twitter API is limited to 100 tweets, large quantities can be obtained by writing a macro to loop through repeated iterations of Twitter extraction.
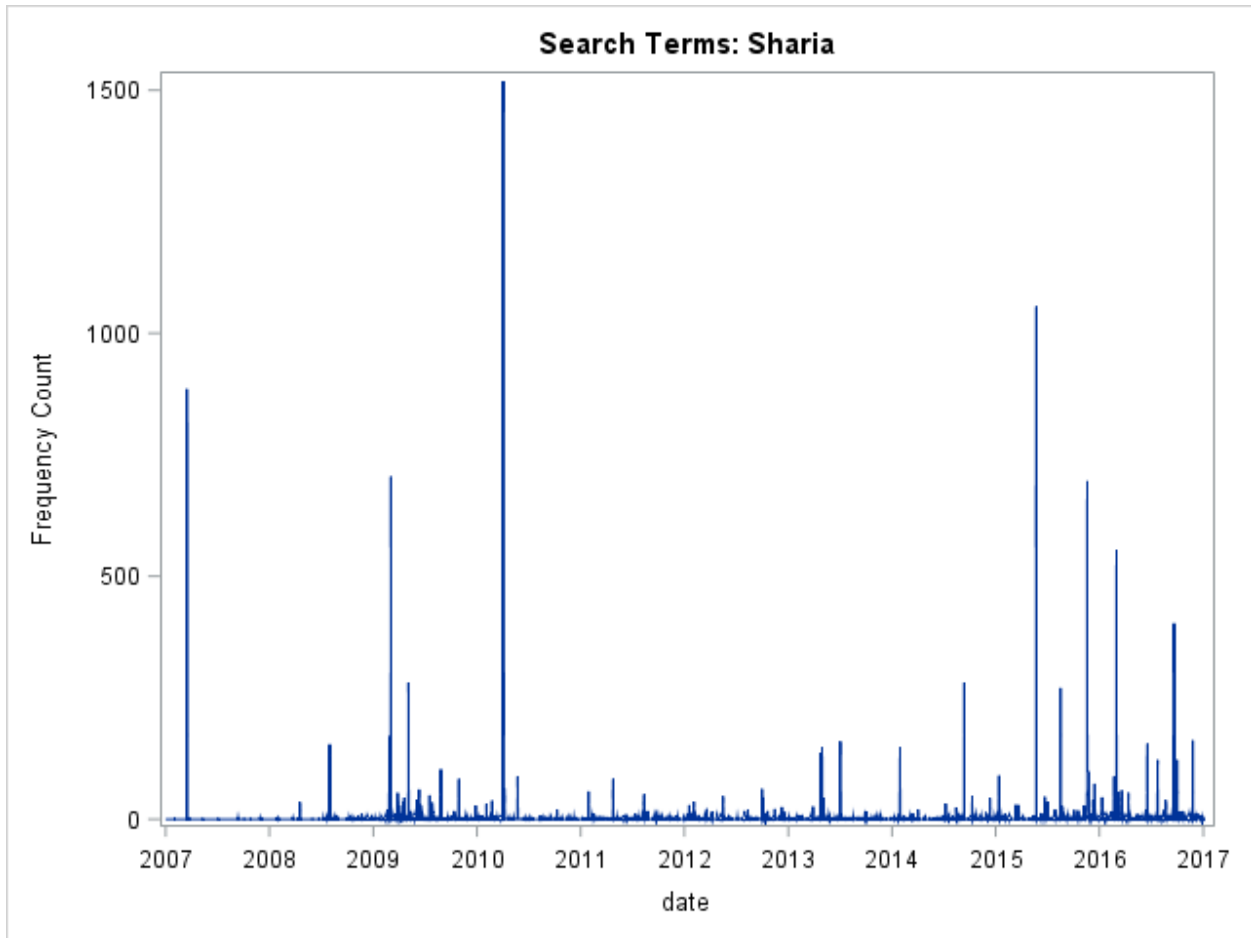
ANALYSIS OF HATE SPEECH IN TWITTER DATA

To perform analysis, it is recommended first to perform exploratory data analysis on a small number of tweets, examining individual tweets, to identify search terms yielding tweets best suited to the desired analysis. For example, in extracting hate tweets, it is found the search terms Sharia, America, and Hate produce a very pure stream of anti-Muslim hate speech.

Once a substantial body of tweets has been extracted, the time series of the number of tweets per day can be compared to either a list of discrete events or event counts. Many different methods of time series

analysis can be used investigate the hypothesized relationship between hate speech and subsequent violence against persons targeted by the speech. (Note: readers inexperienced in the analysis of hate speech are strongly cautioned regarding the use of language in this analysis. In order to analyze the harm of hate speech, the associated risks, and target appropriate responses, this paper deals with some of the very worst things human beings are capable of writing and saying.)

Tweets are characterized by sharp, short-term spikes in the time series of search term counts. For example:

Figure 2: Time Series of Extracted Twitter Data

While the search term "Sharia" is strongly associated with hate speech, other terms can be searched as well, with similar results. Exploratory data analysis reveals the combination of terms "Sharia", "America" and "Hate" produce a very pure stream of hate speech:

	created_at1	tweet_id	text	user_id	name	screen_name	location
1	Thu Jun 08 08:52:55 +0000 2017	872738196777426944	@Isarsour You hate America & support Sharia here. Your time will come.	1223623118	Michael Erwin	michaeloferwins	California, USA
2	Thu Jun 08 03:52:00 +0000 2017	872662466714624001	RT @LMARocks: @Isarsour Not a hate crime, and guess what...America is on to you and your sharia law promotion. Your Alinsky tactics are tr\	4223111532	PorkChop Express	China_is_here	Wing Kong Exchange
3	Mon Feb 23 03:30:48 +0000 2009	872660605043449856	RT @WayneRoot: Agree our schools are run by commies & traitors who hate America & Christianity? What exactly do you call this?\nhttp	1216957579	Karen D. Scioscia	KDScioscia	
4	Wed May 06 21:43:22 +0000 2009	872645595068067840	RT @lockman4mehouse: Many \"new Mainers\" hate America	3290719405	Maine4Trumpism	Maine4Trump	#ME02 Moscow, Maine
5	Wed Feb 19 03:09:53 +0000 2014	872598556640780291	#NewRelease \n\nMaine's Refugee Community: The Best And The Brightest \n\nhttps:\/\/t.co\/NvM4koRXXR\n\n #mepolitics @mainegop @	816402981851660	Maine First Media	MaineFirstMedia	Waterville, ME
6	Thu Jun 08 02:42:52 +0000 2017	872645070616489984	RT @lockman4mehouse: Many \"new Mainers\" hate America	779527044287848	Androscoggin GOP	AndroGOP	Maine, USA
7	Wed Feb 19 03:09:53 +0000 2014	872598556640780291	#NewRelease \n\nMaine's Refugee Community: The Best And The Brightest \n\nhttps:\/\/t.co\/NvM4koRXXR\n\n #mepolitics @mainegop @	816402981851660	Maine First Media	MaineFirstMedia	Waterville, ME
8	Thu Jun 08 02:19:06 +0000 2017	872639091027914753	RT @WayneRoot: Agree our schools are run by commies & traitors who hate America & Christianity? What exactly do you call this?\nhttp	2937854619	Sammie Snickers	Sammie_Snickers	
9	Wed May 06 21:43:22 +0000 2009	872610362209837056	RT @lockman4mehouse: Many \"new Mainers\" hate America	816402981851660	Maine First Media	MaineFirstMedia	Waterville, ME
10	Wed Feb 19 03:09:53 +0000 2014	872598556640780291	#NewRelease \n\nMaine's Refugee Community: The Best And The Brightest \n\nhttps:\/\/t.co\/NvM4koRXXR\n\n #mepolitics @mainegop @	816402981851660	Maine First Media	MaineFirstMedia	Waterville, ME
11	Thu Jun 08 00:24:21 +0000 2017	872610213991501828	Many \"new Mainers\" hate America	2351033034	Larry Lockman	lockman4mehouse	
12	Wed Jun 07 23:38:02 +0000 2017	872598556640780291	#NewRelease \n\nMaine's Refugee Community: The Best And The Brightest \n\nhttps:\/\/t.co\/NvM4koRXXR\n\n #mepolitics @mainegop @	816402981851660	Maine First Media	MaineFirstMedia	Waterville, ME

Figure 3: Parsed Hate Tweets

Reflecting the sort duration and often highly reactive nature of the Twitter channel, other hate speech search can show similar results.
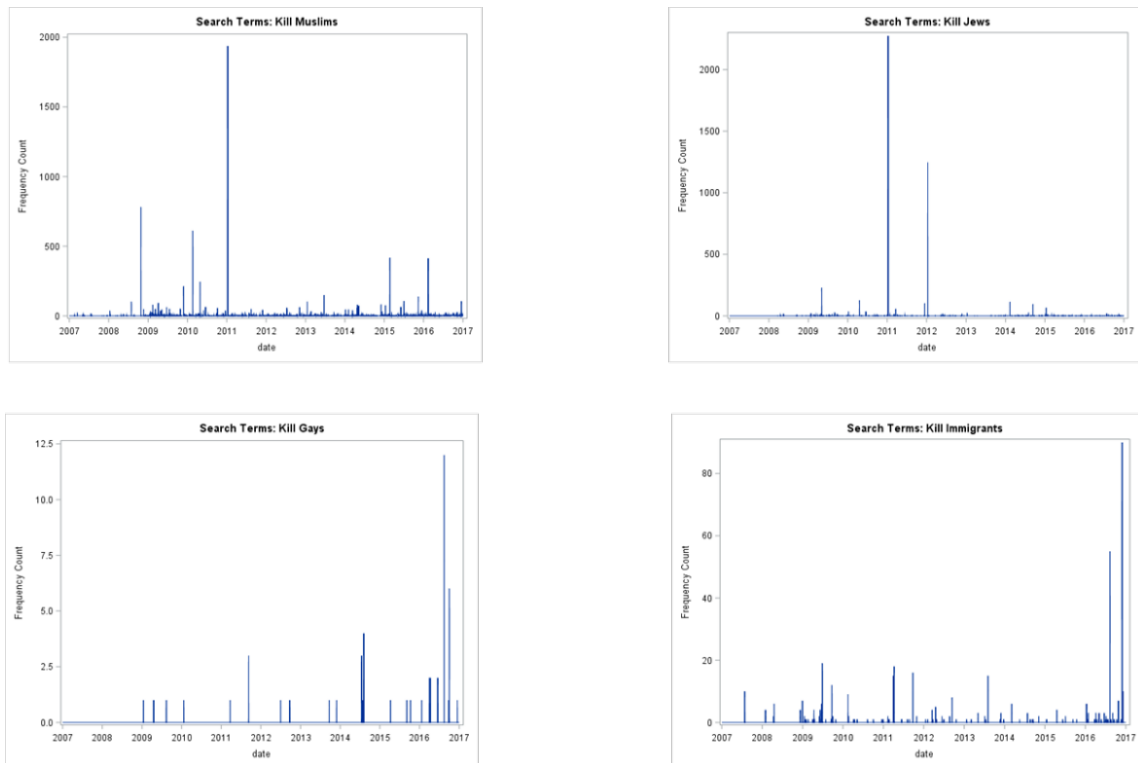

Figure 4: Comparison of Hate Tweets With Different Targets

In this analysis, spikes in Twitter data are not seen to closely precede attacks on minorities and other groups targeted by violence. A highly reactive pattern is seen instead, with sharp spikes in hate speech following terrorist attacks. The size of these spikes may be predictive of the degree to which reprisal attacks are seen.
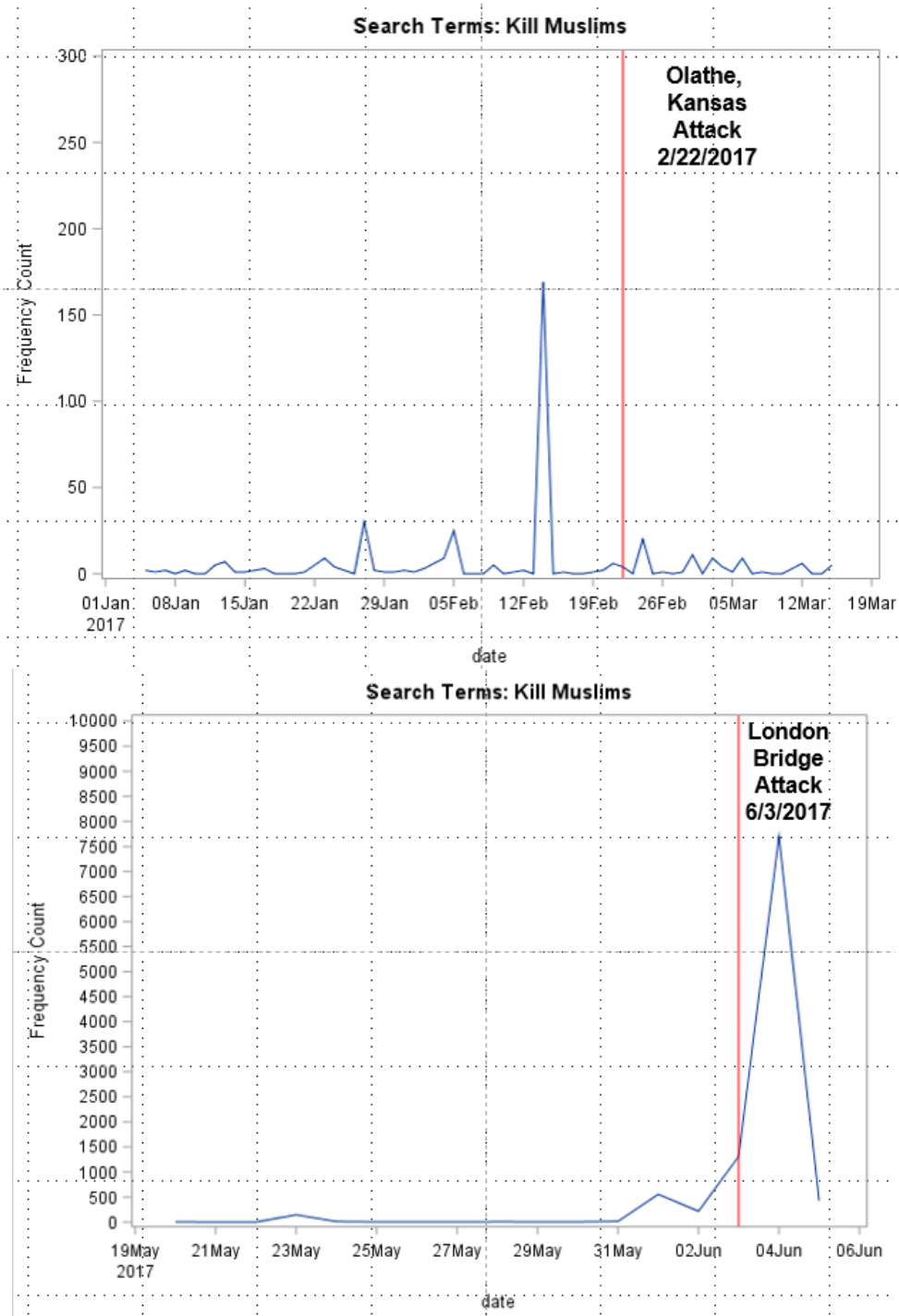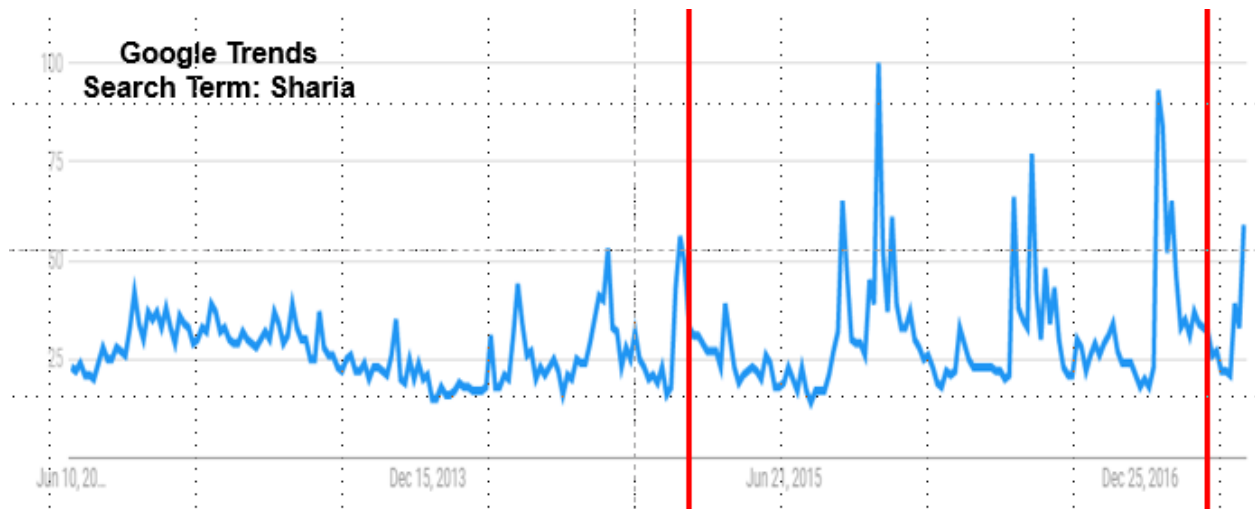
Figure 5: Reactive Tweets – A Lagging Indicator

Contrasting with the time series pattern of hate speech in Twitter, preliminary evidence may indicate peaks hate speech search terms in Google Trends can precede attacks on targeted groups. More data and analysis are need to investigate this.

CONCLUSIONS

Social media data can be matched to event data for time series analysis using a wide variety of methods. Because data can be mined and analyzed at high volume, analytic methods optimized for big data methods may be necessary.

Different social media channels have particular characteristics which should be remembered when selecting sources for analysis:

- Twitter: a short-duration lagging medium, highly reactive to new events. In the case of hate speech, it may be predictive of reprisal attacks

- Facebook is a medium-term channel with extended conversation threads. A general lack of access to data due to Facebook privacy terms limits its use in time series analysis on in-house threads.

- Google Trends functions as a leading Indicator, much like a Social Media equivalent of Durable Goods. Google Trends data may be predictive of subsequent events.

- YouTube and Instagram – the #2 and #3 social media channels as of this writing, offer no effective means to mine images at this time

Twitter data can be mined and analyzed at high volume. Consequently, Analytics methods optimized for big data methods may be necessary. Initial analysis of Hate Crimes with social media data shows promise but better crime data is needed.

REFERENCES

Isabel Litton and Rebecca Ottesen, 2013, WUSS

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

David J Corliss, PhD
Peace-Work
davidjcorliss@peace-work.org
peace-work.org

Time Series Analysis of Hate Speech in Social Media

David J Corliss (Corresponding Author)

Peace-Work
11580 Parkview, Plymouth, MI 48170

ABSTRACT

Mining and analysis of social media text is a powerful tool for the analysis of thoughts, preferences, and actions of individuals and populations. While commonly used today in marketing and other business applications, Data for Good researchers have begun to apply these methods to the analysis of hate speech in social media. Different channels, including Twitter, Facebook, and Google searches, are found to have distinctive characteristics that affect the types of models and analyses each can support. This paper provides a step-by-step description of how to create and deploy a Twitter API and mine the data to extract tweets with user-selected search terms, including keywords and user name of the person or organization sending the tweet. This methodology is used to investigate hate speech, modeling the time series patterns with the aim of estimating the risk of subsequent acts of violence against persons targeted by the speech. These Data For Good analyses have been performed using SAS® University Edition, a free version of SAS® available to students, professors, and non-profit researchers.

Keywords: Time Series Analysis, Social Media, Twitter, Hate Speech

APPENDIX B: CODE USED FOR THE EXTRACTION OF TWEETS FROM TWITTER

```
%MACRO grab_tweet_100(search_term, type, target_amount);


/* Location of the file with authentication information - edit this path */

filename auth "C:\dcorliss\SAS\Twitter\token.txt";


/* Location of the output file containing the tweets - edit this path */

filename twtOut "C:\dcorliss\SAS\Twitter\Tweets.txt";


/*

   These parameters need to be set when accessing the GET search/tweets url:

   COUNT = number of tweets to grab

   TYPE = Popular, Recent, or Mixed

*/


%IF &target_amount < 100 %THEN %LET num_tweet =
%NRSTR(&count=)&target_amount;

%ELSE %LET num_tweet = %NRSTR(&count=100);
```

```
%LET type = %NRSTR(&result_type=&type);


/* PROC HTTP to access the Twitter URL and write the output file */


PROC HTTP
   HEADERIN = auth
   METHOD = "get"


   /* Deploy the specified search terms in the Twitter search API */
  URL = "https://api.twitter.com/1.1/search/tweets.json?q=&search_term&type&num_tweet"


   /* Write the output to the designated location as one long text string */
   OUT = twtOut;
RUN;


%MEND;
```

APPENDIX B: CODE USED FOR REAL-TIME UPDATES FROM A TWITTER FEED

```
%macro timeloop(iterations);
%local i;
%let i=1;
%do i=1 %to &iterations;


%grab_tweet_100(Sharia,recent,100);


data work.tweet;
   set work.tweet;
   count = 1;

chardate=CATT(substr(date_time_text,9,2),substr(date_time_text,5,3),substr(da
te_time_text,27,4));
   timestamp=dhms(input(chardate,date9.), substr(date_time_text,12,2),
substr(date_time_text,12,2), substr(date_time_text,18,2));
   date = datepart(timestamp);
   time = timepart(timestamp);
   ten_sec = round(time, 10);
   ten_sec_est = ten_sec - 18000;
   if input(chardate,date9.) = today();
   if upcase(text) contains 'SHARIA' and upcase(text) contains 'AMERICA' and
upcase(text) contains 'HATE';
run;
```

2

```
data work.graph_data;
   set work.graph_data work.tweet;
   *if ten_sec_est GE hms(9,0,0);
run;

proc sort data=graph_data nodupkey;
   by time tweet_id;
run;


ods listing close;
ods html path='C:\dcorliss\SAS\Twitter' file='test.html';

proc sgplot data=graph_data;
   vbar ten_sec_est;
   *xaxis type=time interval=minute values=(47500 to 49000);
   xaxis type=time interval=hour;
   format ten_sec_est hhmm.;
   title "Tweets (10 Second Totals)";
run;

ods html close;
ods listing;


%end;
%mend timeloop;


proc delete data=work.tweet;
proc delete data=work.graph_data;
run;

DATA work.graph_data;
   LENGTH date_time_text $30 tweet_id $18 text $140 user_id $15 name $185
account_created $30 lang $2 screen_name $185 location $185;
   DELETE;
RUN;

%timeloop(20);
```