**Behavior of Twitter Bot in California Fires**

09/12/2018  
Takeshi Oda

**Introduction**  
Twitter Bot is a type of twitter user account which is programmed to post messages into twitter autonomously. Improper uses of Bot are causing harmful effect on public communication in social media since they are sometimes used to manipulate public opinion.

The recent and worst wild fire in California ‘Campfire’ killed around 85 people. This analysis tried to answer two questions: Is there any different pattern of tweets between bots and normal users? Which type of accounts are more likely to post political or environmental mentioning? This analysis would be suggestive for governmental organizations such as department of fires in each state. The management team in those organizations would know how to distinguish improper use of tweets.

**Data Collection**  
REST API in tweepy was used to search past tweets about California Fires. The hash tag #CaliforniaFires was used as a key word. Tweets from 17/11/2018 to 20/11/2018 were collected. Features to represent tweeting behaviors were defined. (e.g. number of question mark or exclamation mark, number of political term or environmental term, number of hashtags in tweeted text)

Probability of Bot in tweets  
To classify each tweet into Bot or normal tweets, ‘Botometer’, a free API to ‘check a Twitter account and give it a score based on how likely the account is to be a bot’ (OSoMe project), was be used. Tweets having Bot Probability larger than 0.5 were flagged as Bot tweet.

**Data Analysis**  
11,018 random samples from collected tweets were analyzed. Statistical significance between two groups, i.e. Bots and Non-Bots were tested. Average of seven numerical features were compared through two-sample t-test. One categorical feature was compared through chi-squared goodness-of-fit test.

**Result/Conclusion**  
We found statistical significance in several behaviors between Bots and Non-Bots.

* + Tweets by Bot are more likely to be retweeted.
  + Screen name of Bot account is more likely to have digits (Many accounts for professional use?)
  + Tweets by Bot tend to like to comment about environmental aspects of California fire
  + Normal tweets are more likely to embed hashtags in their tweets
  + Tweets by Bot are more likely to be retweets of other tweets

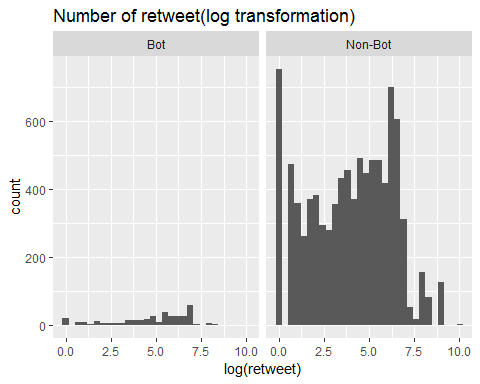
These results tell us that quantitative indicator of tweeter behavior will help distinguish problematic

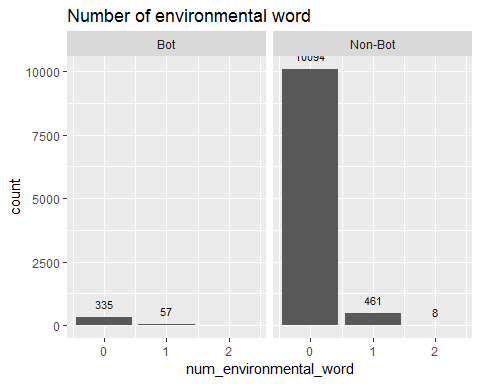
tweets without time consuming topic analyses in large amount of texts. Specifically, we can see

indication of Bot in highly large number of retweets, many digits in screen name and many

environmental terminologies while we can see normal user’s motivation to express their own

sentiment and be caught by others.





**Reference**  
https://www.usnews.com/news/top-news/articles/2018-11-25/number-of-missing-in-deadly-

california-wildfire- revised-down-more-rain-on-the-way

https://botometer.iuni.iu.edu/#!/

Varol, O., Ferrara, E., Davis, C., Menczer, F., & Flammini, A. (2017). Online Human-Bot Interactions:

Detection, Estimation, and Characterization. Accepted paper for ICWSM’17.