Proceedings of the Tenth International AAAI Conference on
Web and Social Media (ICWSM 2016)

Are Social Bots on Twitter Political Actors? Empirical Evidence from a Ukrainian Social Botnet

Simon Hegelich
Technische Universität München,
Bavarian School of Public Policy,
80539 Munich, Germany
hegelich@hfpm.de

Dietmar Janetzko
Cologne Business School,
50677 Cologne, Germany
d.janetzko@cbs.de

Abstract

A considerable amount of data in social networks like Twitter is not generated by humans but by automatic programs (bots). Some of these bots are mimicking humans (social-bots) and can hardly be identified. In this article, we analyze a social botnet involved in the Ukrainian/Russian conflict. Based on text mining and unsupervised learning, we can identify three different behaviors: mimicry, window dressing, and reverberation.

Introduction

Social media like Twitter can be easily accessed and used. This is true for humans – but it also applies to automated programs (bots). The research presented here has been prompted by discussions on social botnets. Its focus is a social botnet involved in the current conflict in Ukraine. Its far-reaching structure and size is invisible to most Twitter users, but appropriate analysis techniques reveal that it consists of a complex network of thousands of Twitter accounts. It cannot be ruled out that the enormous size of botnets helps them to disguise their behavior and that their behavioral patterns are intentionally blurred by noise, i.e., Tweets unrelated to the social botnet's objectives. Distinctive behavioral patterns only surface if large amounts of data are analyzed. This leads to a methodological challenge. On the one hand, a large sample is required to identify and to characterize different types of bot behavior. On the other, it is possible that only a very small subsample is different from chitchat (noise), by having a distinctive political quality (signal). Using the methodology of a single-case study (Yin, 2013), we address two research questions: (1) Has the social botnet a political agenda? (2) Which kinds of behavior can be identified in the botnet?

Methods

The initial dataset contains Tweets with the hashtag "#Ukraine", collected for 24 h via the Twitter streaming API on February 22th, 2014 – the day of the impeachment of former Ukraine President Janukovitsch. This data included Tweets with identical texts not labelled as "retweet" in the meta-data, which is unlikely for normal users. We were able to expose 86 of the Twitter accounts as social bots: Twitter-meta-data contains information about the status source, i.e., the URL from which the Tweet was sent. Following this URL, we could identify *twifarm* – a program designed to manage huge numbers of social bots on Twitter. Therefore, we know that these Twitter accounts have been controlled by social bot software. To identify other social bots from the same botnet, we traced friends and followers of the social bots, checked their status sources, too. Several repetitions of this procedure revealed a huge social botnet. Often, bots follow other bots, and hence the sample includes duplicate Tweets as evidenced by the unique Tweet ID. Removal of duplicates led to a sample of 1.740 social bots. We downloaded the latest Tweet of each social both, translated each of them via the Google translation API from Ukrainian and Russian to English and pre-processed all Tweets (removing punctuations, transferring to lower characters, removing stopwords). The resulting document-term-matrix was then subjected to the following statistical analyses.

- Creation of a word cloud (Fig. 1).
- Analysis of the most frequent words (Fig. 2).
- Calculation of word-word correlations and visualization as a network, which revealed connections between terms across the Tweets (Fig. 3).
- Conducting different methods of cluster-analysis.
- K-means cluster analysis with the Hartigan-Wong-algorithm (Hartigan and Wong 1979)
- Hierarchical cluster-analysis

Figure 1: Wordcloud of Bot Content

While the first 3 descriptive analyses are straightforward, the cluster-analyses merit a discussion. Firstly, we run a k-means cluster analysis with the Hartigan-Wong-algorithm. Well-known problems with this analysis are (i) the choice of the "right" number of clusters and (ii) its dependency on the starting point. However, by the sum of squared error (SSE) for different numbers of clusters the solution found can be assessed. We ran the cluster analysis with different numbers of clusters k (2 to 15). For each k, the best cluster was chosen from 10 random starting points. For each repetition, the SSE was calculated, and 8 emerged as a good number of clusters as the SSE decreases only lightly with an additional cluster (Fig. 4). Secondly, we carried out a hierarchical cluster analysis, to examine whether there is a stable separation between political and non-political content. Due to the Twitter limitation of 140 characters, a document-term-matrix from Tweets usually contains zeros. In each Tweet, the number of words is very limited while the number of different words in the whole corpus is increasing with sample-size. Shared absence of a word therefore is not a good indicator for clustering. The binary distance (or asymmetric binary) is based on shared presence of features and should therefore represent the structure in the data properly. The results of the two different cluster algorithms are compared in Figure 4. These results have been used to identify different types of bot behavior by interpreting the content of each cluster. Hierarchical clustering has been successful in different settings in political science (Grabau and Hegelich 2016).

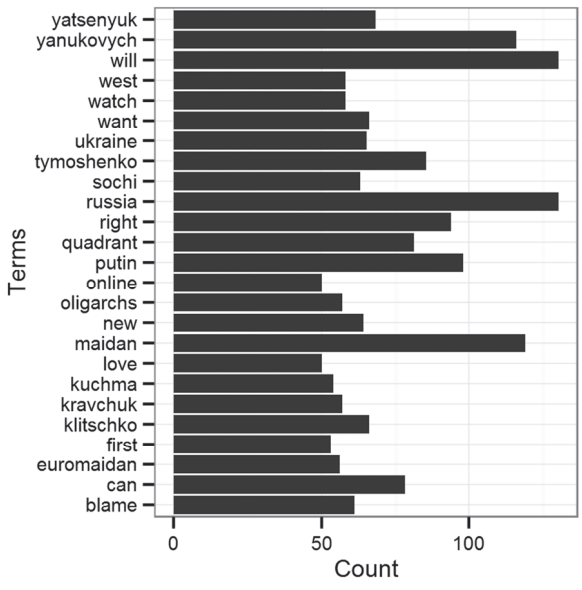


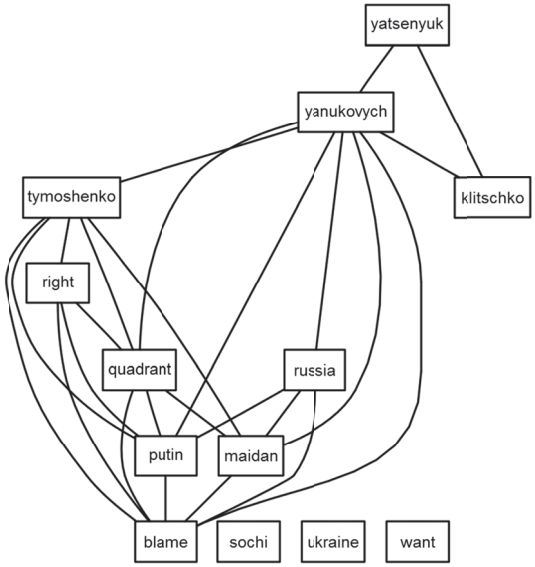
Figure 2: Most Frequent Words

Results

Descriptive Textmining

Fig. 1 shows a word cloud of the corpus with the word frequency been color-coded. The most frequent words are of political relevance. The bots are tweeting about "Russia", "Yanukovych", "Maidan", "Putin" etc. But there are also words like "cinema", "Christmas", "Godzilla", "Hercules" etc. that do not seem to be of political relevance.

Fig. 2 shows only terms that occur at least 50 times



which underscores the dominance of political topics.

Figure 1: Correlation Plot

Fig. 3 has the correlations between the most frequent terms. The term "сектор" (quadrant, sector) strongly correlates with "right" (Pearson's r = .94). We conclude that in most cases the original term has been "Правий сектор" (ultranationalist Ukrainian party Right Sector). Likewise, the terms "west", "Putin", "square", "blame" etc. all correlate (Pearson's r >= .94). Terms like "online", "Sochi", or "Russia" seem to fall in a different group. The patterns suggest that the botnet has a political agenda, though its specific intention is ambiguous.

k-Means Cluster-Analysis

Next is a description of the clusters and their corresponding Tweets revealed by k-means cluster-analysis.

Cluster 1 (almost 80% of all Tweets) was the largest one. Typical Tweets are: "Hollywood stars did not come to the party of the British Prime Minister" "RT @ ArkadijDR5C: Hercules beginning of legends download http://t.co/VvSjK0d2YP"

These Tweets have in common that the content is changing a lot. There are news reports (mainly from Russian news pages), retweets of other users, jokes (many of them sexist), and links to (probably illegal) downloads for films. This cluster is the main reason, why the bots are very hard to identify. Its strategic purpose seems to be twofold: camouflaging the true intentions of the botnet and presentation of content of interest to "normal" users.

Cluster 2 (1.5%) consists of just one Tweet, which is sometimes posted as original Tweet and sometimes as retweet: "RTElizavetaUkd: Has this revolution exhausted? Do you still have hope?" Ending with "#евромадан #євромадан #майдан #euromaidan", the original Tweet reveals hashtags both in Cyrillic and Latin for "Maidan" and "Euromaidan". On Twitter, messages become more influential when hashtags are widely used. The strategic purpose of this Tweet might be to increase the importance of these hashtags.

Cluster 3 (about 3%) consists mainly of Tweets with calendar mottos. We assume that the botnet has a corpus of different texts from which the bots glean content. Typical Tweets are: "If you want to know the person, do not listen to what others have to say about it, better listen to what he says about others"

Cluster 4 (6 Tweets) is the smallest cluster with retweets of a news report about a statement of Putin on World War I: "RTalulpholitwai: RTFake_MIDRF: Putin: Russia has almost won the First World War. In Russia almost to build democracy and a society of welfare ..."

Cluster 5 (7.2%) is mainly about news taken from Ukrainian and Russian news channels like Ukraine Today and NewsUA24. Typical Tweets are: "Poland's economy will suffer because of sanctions against Russia" "RT @ JGlotov32: Opening of the International Festival of Theatre and Film "Family Circle" in the East Siberian Irkutsk will take place on the shore of Lake Baikal ..." "Ukraine prepares sanctions against Russian companies."

Cluster 6 (1.6%) contains only one message that is retweeted many times: "RT leinidetelre: RT dazaruzepuq: Maidan, Russia, Ukraine, USA, Golden Eagle, Euromaidan, Yatsenyuk, Tyagnibok, Klitschko, the EU Yanukovych" "Golden Eagle" is the name of former Ukrainian special forces "Berkut" (Бе́ркут). Probably, here the objective is again to increase the popularity of these hashtags.

Cluster 7 (1.2%) contains Tweets that are derived from headlines from the webpage of *Right Sector*: "RTqubolacu: Right Sector fighting piracy: Right Sector made a statement about the illegal use of its symbols. - ... Http: //t.co..."

Cluster 8 (4 %) has only one Tweet that is retweeted by many different users: "Who is to blame? Maidan? Putin? Yanukovych? right sector? Oligarchs? Russia? West? Tymoshenko? Kravchuk? Kuchma? Yushchenko ..." This message can be seen as the signature Tweet of the social botnet. A Google search for the original Tweet lead to 66.000 hits, which suggests that the botnet has been successful in spreading this message.

In sum, the analysis reveals that the social botnet studied is capable of at least three kinds of distinctive behaviors. The bots try to *hide their bot-identity;* by being *interesting to normal users* whilst *promoting topics* via pushing hashtags and retweeting selected Tweets.

Hierarchical Clustering

The k-means cluster analysis clearly distinguished between political ("signal") and non-political Tweets ("noise") in the social botnet. This is remarkable as the biggest cluster is characterized by non-political content. But are these results robust? The k-means algorithm strongly depends on the number of clusters and the randomly selected starting points. Hierarchical clustering does not suffer from this weakness, and using the binary distance and complete linkage method leads to unambiguous and compact clusters. The clustering is not biased by randomness, and all observations within one cluster are very close together. However, there are two disadvantages associated with hierarchical clustering: the number of clusters cannot be controlled, and vis-à-vis minor changes the stability of the clusters remains unclear. To overcome these weaknesses, bootstrapped p-values have been calculated for each cluster. To compare the results of the two clustering approaches, we adopted a visualization approach invented to

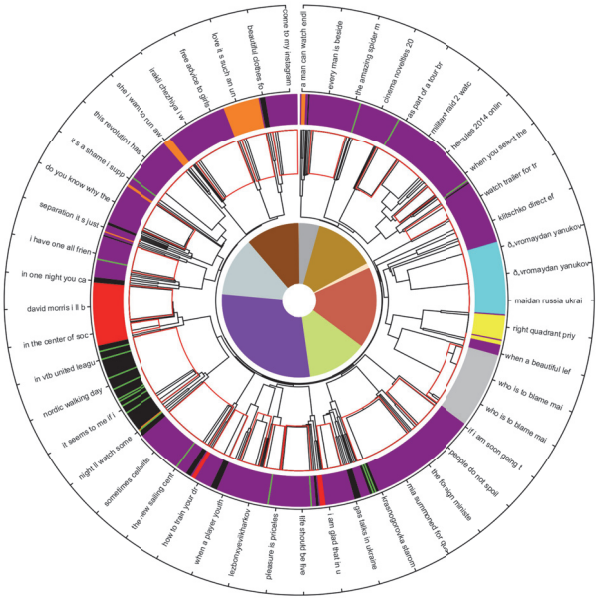


Figure 4: Circle-Dendogram

compare clustering of genomes by ordering information in a circle structure (Gu et al 2014).

In Fig. 4 the Tweets are ordered according to the hierarchical clustering. The outer ring shows the first 20 characters of every 20th Tweet The second ring in strong colors shows the results of the k-means clustering. Every color represents one of the 8 clusters. The third ring contains the dendogram of the hierarchical clustering, bended like a fan. In addition to a normal dendogram, stable clusters as indicated by bootstrapping (procedure adjusted p-value >= .95) are highlighted red. Hence, hierarchical clustering starts with each observation in a separated cluster and combines these clusters until only one is left, we can identify the results for 8 clusters and therefore compare this clustering approach directly with the k-means clustering. This is shown in the fourth ring in pastel colors.

From this visualization we derive the following insights:

- Both clustering approaches lead to comparable results.
- The results are mainly robust in the bootstraping procedure.
- The k-means approach finds one big class of "noise" and several classes of political "signals" while the hierarchical clustering differentiates the "signal" in useful subgroups.

Conclusions

Has the social botnet a political agenda? Which kinds of behavior can be identified in the botnet? The analysis presented in this paper provided answers to both research questions: (1) The social botnet studied has a political agenda. Even though the overwhelming majority of Tweets generated social botnet is often inconspicuous, political themes emerge. (2) The social botnet studied exhibits three distinctive patterns of behavior.

- Mimicry: The bots try to *hide their bot-identity.*
- Window Dressing: To be *interesting to normal users* they are *promoting* topics by pushing hashtags and
- Reverberation retweeting selected Tweets and messages.

The study revealed that the behavior of the bots is not guided by a simple deterministic structure of command and obedience between a human botmaster and an army of bots. Instead, we can show that the politically relevant behavior results from complex algorithms leading to a high degree of *autonomy* of the bots: The bots are not doing something they have been directly told. Instead they follow abstract rules like "Take a popular tweet and add the following hashtags". In addition, most of the time, the bots are not following their direct "mission". Instead, they use algorithms mirroring the behavior of normal users. This makes it extremely hard to identify the bots and to understand their political aim. At first glance, it might seem ineffective to send thousands of irrelevant messages before something of importance is transmitted. But taking into consideration the scalability of social botnets, this seems to be a very effective strategy because the bots are increasing their audience all the time.

Social media is already affecting the political and economic environment (Janetzko 2014). Social botnets are a new development with high political relevance. To understand political propaganda in social networks machine learning and data science methods are an essential tool for political scientists (Hegelich, Fraune, and Knollman 2015).

References

Yin, R. K. 2013. *Case study research: Design and methods*. Los Angeles: Sage Publications, 2013.

Grabau, M. and Hegelich, S. 2016. The Gas Game: Simulating Decision-Making in the European Union's External Natural Gas Policy. *Swiss Politcal Science Review* doi:10.1111/spsr.12202.

Hartigan, J. A., and Wong, M. A. 1979. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28 (1): 100-108.

Hegelich, S., Fraune, C., and Knollmann, D. 2015. Point Predictions and the Punctuated Equilibrium Theory: A Data Mining Approach—U.S. Nuclear Policy as Proof of Concept. *Policy Studies Journal* 43 (2): 228-256.

Janetzko, D. 2014. Predictive modeling in turbulent times – What Twitter reveals about the EUR/USD exchange rate. *NETNOMICS* 15 (2): 69-106.