Analysis of bot behaviour on Twitter

Data Story Project

Mustafa Ourtashi

Winter 2022

CCT 416

***Overview***

Misinformation has become a rampant issue in modern day journalism and reporting as news spreads extremely quickly in our interconnected world. It has become increasingly important to understand the plethora of information we are taking in a digesting each day as it now shapes our world view furthermore, it is even more important that we are critical of the information that is being fed to us and rectify any false facts and statements that are being seen. This misinformation crisis has become an increasingly difficult issue for the current Russian-Ukraine war as information online is being misrepresented and falsified to justify and solidify different political views or deceive the general public’s perception of the current situation on the ground. One of the major challenges social media and news reporting sites are facing with regards to this is the spread of falsified images and videos that claim to be from current Russian military efforts (Silverman & Kao, 2022) reported that these informational attacks are so effective as they don’t even require the user to be convinced that they are true, as long as they become “uncertain” the desired misinformation effect is reached. One of the key reasons we still see a major amount of misinformation (especially on social media) is that it has become almost impossible to police. (Duffy & Metz, 2022) sates that verifying images and videos online that make claims is almost impossible for an individual as it requires expensive forensic analysis tools. Despite all this a twitter data analysis could help us in better understanding how this misinformation is spread and who is interacting with it, this is why I have chosen to answer the following research question: With respect to the current Ukrainian-Russian, do bots on Twitter have the capability to spread bots misinformation if so, what type of misinformation and to what extent?

***Method***

To collect the data necessary for my analysis I used Netlytics for data visualizations and text analysis. The data was collected from the Netlytics program directly using their built-in twitter API. The network analysis was used to detect bot behaviour within the network whist the text analysis was used to detect the type of misinformation bots were spreading onto the twitter-sphere. To conduct the network analysis, I first had to define all the relevant search query terms, for my purposes I had experimented with the following queries the Netlytics program: 1. Misinformation + Ukaraine+Russia 2. Ukaraine+Russia misinformation 3.Russia+Ukraine+Disinformation+Emotion and 4.Misinformation (+Ukraine OR Russia) +Disinformation.

I found the fourth search query to be the most accurate measure of bot behaviour on twitter as it outlined the type of information they are spreading, the emotions that the information evokes in human Twitter users and does not prioritize Ukraine or Russia due to the Boolean operators. This search then granted me accesses to both the data network visualization and a text analysis. For my network analysis I visualized two networks from the same search query, the first network included all the network discovery features (replies, quotes, retweets and mentions) this data visualization acted as the control group, here I was able to monitor all the behaviour within the network from bots and humans, positive and negative.

For my second data visualization I wanted to visualize the network without the heavy influence of bots, through a process of trial and error I was able to eliminate most of the bot behaviour within the network through the removal of the retweet discovery parameter. Now I am able to compare this new network visualization to the control one previously rendered. For my text analysis, different approach was used, instead of identifying control and experimental groups to discover bot behaviour, one large text analysis was conducted to identify the type of misinformation bots were spreading onto Twitter. I chose to remove the Taste, Touch and Shape dictionaries as results with them included rendered irrelevant information that produced anomalies for the remainder of the dataset. Note: The visualizations of both data networks can be found in the appendix below.

***Results***

The results obtained from the network and text analyses conducted can be found below. First, the network analysis. Through the process of trial and error by the elimination of network discovery features I was able to determine that bots use primarily the retweet function to spread their agenda on Twitter. Both networks that were visualized showed that most nodes are interlocked within their own ecosystem (this means there was verry little crossover between different clusters) but, some interconnectedness between clusters was observed on both networks visualized. The text analysis yielded some important results as well, for the dictionaries employed the word count is as follows. Swear Words – EN (225 Posts, 234 Terms) Feelings Good (67 Posts, 70 Terms) Feelings Bad (123 Posts, 129 Terms) Size (46 Posts, 51 Terms) Time (48 Posts, 48 Terms) Touch (41 Posts, 41 Terms). Whilst there were more dictionaries reported on the text analysis, they were too minuscule in size to deem relevant.

***Data story***

Prior to conducting the research there were three major questions that needed to be answered: 1. How do bots on twitter spread misinformation? 2. What type of misinformation is being spread? and 3. Do regular Human Twitter users interact with this misinformation, if so tow what extent? The results yielded from the data provided three main learnings each providing evidence to help answer the questions posed above.

First, looking at the data acquired from Netlytics, the network analysis helps us in understanding how bots navigate through twitter. Appendix 1.1 shows a network that is infiltrated with bot behaviour, all major clusters within this network have a bot at the centre of it with a collection of bots surrounding it mainly retweeting or quote tweeting the information from the centre of the cluster, I have deemed this type of cluster as a Bot Retweet Cluster (BRC). Appendix 1.2 depicts a network with a relatively muted bot population (relative to the network from appendix 1.1). For this network I used the exact same parameters and collected the same tweets but restricted retweets from showing up on the network visualization. Examination of this network and its clusters showed real human discussions on the topic instead of BRC’s, this was not observed on almost all the clusters within this network providing enough evidence for me to deem this a “Human focused network”. Despite the fact that this network showed a lot of human behaviour, bots where still able to find their way into the network. Appendix 1.3 illustrates a bot’s infiltration into a predominantly “Human” network. The data gathered from both network analyses allows us to answer our first major question. Bots on twitter are analogous to cells, the replicate rapidly within their own clusters through mainly retweets. Usually in a single cluster there will be anywhere from 20 to 50 nodes that are replicating and retweeting the same string of information. The rationale behind this is to get the misinformation Infront of as many eyes as possible as they don’t have to convince the human to reach their desired effect (Silverman & Kao, 2022).

The data collected from the text analysis allowed us to look further into what exactly these bot networks were trying to communicate/deceive the human user in. As reported in the results section an overwhelming majority of the information that is put across through bots is negative in some sort of way. In just the Swear words and Feelings Bad dictionaries, they accounted for a combined total of over 63% of the total words within the text analysis. Whilst the swear words dictionary is self-explanatory, I further explored the feelings bad category to gain further insight on what information bots are spreading. Words such as: Bad, Dangerous and Awful were frequent within this dictionary but surprisingly so were emotional and descriptive terms such as: Scary, Upset, Angry, Evil, Hurt and Arrogant. Prior to completing any data collection or analyses I was of the belief that bots on Twitter mainly use swear terms to evoke a response from human user, whilst this may be true as the Swear Words dictionary yielded the highest return of tweets and terms a major component of these bot’s tweet behaviours is the pursuit of evoking emotional responses within the human user. This could be due to a multitude of reasons that are outside the scope of this current study, but it is worth mentioning that emotional arousal is more clearly remembered than facts and evidence (Kensinger, 2009).

Finally, the interaction between bots and human users on Twitter. All the data collected from both the text and network analysis do not provide a great deal of evidence for the extensive connection between bots and humans on Twitter. Appendix 1.3 shows a small interaction from a bot who has happened to infiltrate a human network, but this connection is not too secure, only one node is directly attached to the parent network and can easily be cut off if a single user leaves the network. Similarly, the networks showed in appendix’s 1.1 and 1.2 both illustrate self-contained clusters that do not reach out to neighbouring networks. Despite this all, is important to note (Silverman & Kao, 2022) comment on the goal behind bot networks. Although direct contact might not be established between a human and bot cluster, the goal of a bot operator can be reached simply through impressions and views. When this is taken into account, we must then take into account the sheer volume of bots within the network for any given topic, if the bot network exceeds the human network in population, then it would be safe to assume that a heavy amount of human-bot contact is being made. We must assume here as the data collected for this study does not provide the required evidence to make a causal connection and answer the question with complete clarity. Future studies can further explore the question of human-bot connections in Twitter networks through the collection and analysis of both populations.

***Appendix 1.1***

***Background pattern

Description automatically generated***

***Appendix 1.2***

***Background pattern

Description automatically generated***

***Appendix 1.3A picture containing radar chart

Description automatically generated***

# Works Cited

Duffy, C., & Metz, R. (2022, March 15). *Why Ukraine war misinformation is so hard to police* . Retrieved from CNN: https://www.cnn.com/2022/03/15/tech/ukraine-russia-misinformation-challenges/index.html

Kensinger, E. A. (2009). Remembering the Details: Effects of Emotion . *National Library of Medicine* , 99-113.

Silverman, C., & Kao, J. (2022, March 8). *In the Ukraine Conflict, Fake Fact-Checks Are Being Used to Spread Disinformation* . Retrieved from ProPublica: https://www.propublica.org/article/in-the-ukraine-conflict-fake-fact-checks-are-being-used-to-spread-disinformation