Assessing the influence of sockpuppets on the Twitter Transgender community

Supervisor: Professor Luc Berthouze
Candidate Number: 215758

October 30, 2021

Summary

1 Introduction

On Twitter, a user can easily create an account using an email address or phone number, and this opens up the issues of users exploiting this to create malicious user accounts which are either controlled automatically via the Twitter API or a human user who has another account which they use as an anonymous pseudonym to hide behind and express their true views online [1]. Both can be labeled under the category of a sockpuppet. Sockpuppets have a puppet master, the human who owns the accounts credentials for the account and can orchestrate multiple of them [2].

Sockpuppets can affect online social media platforms, such as Twitter, to spread disinformation; these sock puppets could just be random users or more organised such as state-sponsored [3]. Twitter in 2019 had 290.5 million users [4]; this is a huge attack vector for puppet masters to use. In 2017, it is thought that between 9 and 15 percent of Twitter accounts are bots and that bots use retweeting strategies to target specific communities

of people [5]. Not all bots are bad; some are neutral or designed to be helpful, however, they have been used to infiltrate political discourse and manipulate them [6][7]. It was estimated that $\frac{2}{3}$ of tweeted links were posted by these automated accounts [7]. Past twitter data show similar margins to polls collected during the election period in 2016 from three million tweets [8]. During the COVID-19 pandemic, many conspiracies were propagating on Twitter (and as of writing there still is); there exists a set of bots that mainly posts conspiracy theories of the political type around the COVID-19 topic [9].

We are very interested in how social media affects people and society as a whole, and how ideas can propagate within communities. There is a rise in reported hate crimes in the UK. Race, sexual orientation, disability, and transgender hate crime categories have risen over the year. Race was the highest at a 12% increase from just this year alone (76,158 to 85,268). Over the span of five years, there has been a 120% (in 2016/2017 1,195, in 2020/2021 2,630) increase in reported transphobic hate crimes [10]. This is only the reported hate crimes. Transphobia online, as well as any other kind of bigotry, is abundant, with collected tweets which span from 2016 to 2019 discussing transgender people on Twitter where 12% ($\frac{789615}{5547445}$) are being abusive [11].

We want to try to detect sockpuppets within the transgender community and to assess the influence of the sockpuppets if they do exist. This research project is split into separate questions; we will look at the influence of users within the transgender community; try to detect sockpuppets within the community, and see how much influence they have on the community. If we finish our main research questions, we will then assess the particular topics the users discuss within the community.

2 Background research

2.1 Twitter

Twitter is an online social network, specifically a micro-blogging platform. Users can sign up with an email or phone number. Each user has their own display name and username. A username is a unique string identifier for a user which is attached to a Twitter user id. A display name is a public-facing name that acts like a nickname for a user, for example, you could have the username as "@HelloWorld" and have a display name of "foo bar".

A user can post tweets on their account; tweets are 280 characters max length text, which can have additional attachments such as up to four images, or one video, or one audio recording. Tweets can link to external web pages such as news articles, but this takes up characters. Users can follow other users to have their tweets appear on their timelines. The count of the users following and the followers are publicly viewable. Each user has their timeline, where tweets are displayed from the users they follow, this also includes retweets.

A user has several actions they can perform onto a tweet: like, reply, and retweet. Likes are a way of showing one likes the content of the tweet, and replies allow one to comment on a tweet, and a reply in itself is a tweet that can be replied to as well, creating long threads of them. Retweets are when you amplify a tweet by showing their tweet as a part of your collection of tweets, users can also quote retweet, which allows you to reference a tweet and add your own 280 character max length text to it too, usually adding on top of the tweet. For each type of action, a numerical value is attached showing the amount of that action that has been performed, one for likes, tweets, retweets, and quote retweets.

Within tweets, a user can mention a hashtag; hashtags are keywords or phrases which have a suffix of #. They are used to help with searching for a particular topic, such as *#javascript* for tweets around the JavaScript programming language which a user has tagged it as.

Twitter supplies three types of APIs. API stands for (A)pplication (P)rogramming (I)nterface. APIs allow a programmer to interface and exchange data with another service, such as a third-party service like Twitter [12]. The three API types are Standard, Academic Research, and Business. We will be using the Academic Research API, as this provides a higher tweet cap per month, 10 million tweets a month for academic research API while the standard only provides 500,000 a month. It allows access to "real-time and historical public data with additional features and functionality that support collecting more precise, complete, and unbiased datasets" [4].

The API allows us to collect data on the number of likes, retweets (both normal and quote retweet), and replies for each tweet. For user accounts, the user creation time, bio, followers, following, and more are available. We can use the historical API to get old tweets from given users or a search term or use the real-time stream of tweets which can then be filtered through for specific keywords.

2.2 Sockpuppets in Social Media

Estimates show that between 9% and 15% of active Twitter accounts are bots [5]. Not all are malicious, some are even helpful tools for people, however, there are still many which are harmful which use tactics such as to try and manipulate users without them knowing as well as exploiting the Twitter algorithm [6].

Social media has been exploited by people to push certain ideas to the platform as a whole, such as to push users into groups such as ISIS sympathisers in the Syrian revolution, the alt-right, and the activists of the Euromaidan movement. They exploit twitters algorithms to gain more awareness for their groups; this all occurred on Twitter [2]. Governments have also exploited social media networks using cyber troops[3] to try to skew the general populations' opinion such as for election periods. Cyber troops are government military or political party groups to manipulate public opinion via social media such as Twitter.

Data collected after the 2016 election showed that one can get the consensus of the population's political leaning as the data showed had similar margins to polls which commenced during that election period [8]. This shows the potentiality of the network having an effect in the real world.

2.3 The Transgender Community

The Transgender community is made up of trans individuals who include binary and nonbinary gender identities.

The Twitter Transgender community includes the Transgender community as a subset of this community; additionally, it includes cisgender allies who either could be a part of the LGBTQ+ community or are heterosexual; in general, they are people who are affected by or interested in transgender issues.

In networks, a community is defined as a subset of nodes within the graph where connections between the nodes are denser than connections with the rest of the network, this is not just a computational problem but a social and biological one too [13].

The Transgender community is an online social network but an in-person social network too, while the Twitter Transgender community is predominately online due to it being based around Twitter as a platform, however, this does not mean that users do not know each other in-person.

Relationships between trans people as a support group helped with the shared experience of the psychological stress of being Transgender. [14]. Trans youth online use unique strategies for moving through a binary gendered online world, creating their own communities [15].

2.4 Community Detection

We want to try to detect, meaning to capture, a specific community which we want: the transgender community.

Community structure detection is usually a set of techniques in network science to detect multiple communities within one large set of data from a network/graph, but these attempt to capture all of the communities within the network [16] when we only want to capture one. To be able to map the community we need to apply specific capturing techniques for a single community.

One method is collecting tweets around terms of interest for the detection of a community around a certain topic. In [17], they use snowball sampling [18] at first which collected 119,156 users, which then with certain keywords in tweets to limit down the results to the users affiliated with ISIS only.

Another technique is snowball sampling [18], this is where a selection of users are chosen as **seed agents**, the user's followers from these agents are then added to the selection of users. This can be iterated, called **hops**, such as finding all seed agent followers (1-hop), and then finding all followers of the seed agents followers (2-hop). This technique was used to capture three communities [2]: the alt-right, ISIS sympathizers in the Syrian revolution, and activists of the Euromaidan movement. This method was used over the first method as it may potentially make it easier to observe the behaviours of the sockpuppets. They collected 106k users and 268 million tweets on the alt-right alone. This could collect mutual users who are not a part of the community however we can experiment with this.

Using the Twitter stream API is useful for research in general [19], giving us all the real-time created tweets. We could potentially use the technique of the typical meaning of community detection itself, however, typical techniques are not to detect a specific community but to find all the communities within the network. It may capture the specific community we want to collect data on, however, this may be difficult from the vast amount of data generated on Twitter from users.

2.5 Measuring Influence

To know how much influence a sockpuppet has, we need some way to measure influence over the network. There is no definitive definition of what an influential user in a social network is within this research area. New research papers with proposed influence measures give different perspectives and methods which define an influential user. Influential users can be split into several categories, such as opinion leaders, influencers and, discussers [20], there are many more, see [21] for survey. In general, we can define a user as influential if their actions within the network can affect the behaviour of the other users in the network [21].

Twitter has its own metric system to see how well a tweet has performed; they simply count the number of times the tweet was seen, the number of times someone engaged with tweets, and more specific metrics seen in

Figure 1. Note, however, these details are not available via the Twitter Research API, which means that we cannot make influence measures using these metrics.

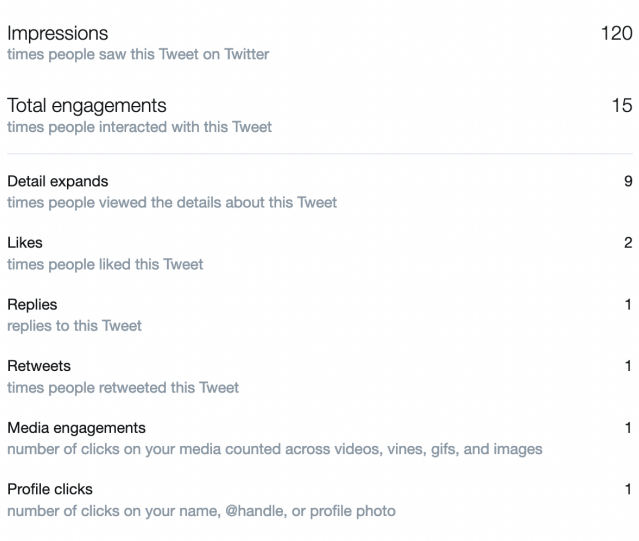


Figure 1: Official Twitter metrics

However, as mentioned in 2.1, the metrics available to us are the number of likes, normal retweets, quote retweets, and replies for each tweet, as well as the tweet text itself. This data can be a metric for measuring influence as this is the way a user can interact with a particular influential tweet or account [22] [23] [24] [25] .

The general measures researched are indegree, retweet, and mention influence. Indegree is the number of inbound connections to a particular node. where the node would be a user and the inbound connections are the users who follow the user. Popular users who have a high indegree are not as influential as people may think for making users retweet or mention[22]; the topological measures such as indegree alone reveal very little about the influence of a user.

In [23], for the time published, there were pre-existing closed source tools to measure influence, one of which is Klout which uses more than 25 variables, they see influence as the "ability to drive people to action", which makes replies and retweets the most important factors of their measure. Another service is Twitter Grader; this measure scores out of 100 and is also closed source, but factors that contributed to its algorithm were the number of followers, Twitter Grader score of those followers, the number of tweets the user has made, update recency, follower/following ratio and engagement (such as retweet and mention ratio for when a user has interacted/engaged with a tweet).

In [21], an active user is defined as a user who participates in the network consistently over a period of time. We can define the activity of a user as the probability of a user seeing a tweet. However, Twitter users who exclusively read who may be active on the network will not be captured as there is no visible way they interact with the network unless they use an observable metric such as retweeting, if they do we can assume they have read the tweet, meaning that if they have done more active users are likely going to exposed to said new tweets, in turn, have the possibility to interact with them.

Influence can be split into two paradigms: influence is predominately from a small number of users who are very connected or persuasive, or, many users can accidentally become influential depending on unpredictable factors [26]. Twitter influence measures in current research have usually used metrics related to retweets, mentions, and less used, followers. Some researchers have used passive topology of twitter such as a follower-s/following graph or retweets and mentions graph. Other authors have looked into the problem of influential users given a certain topic. [21]

One method is *TwitterRank*, which is a topic-sensitive measure. It is an addition to the PageRank algorithm (an algorithm that determines how relevant a web page is given a search term [27]). TwitterRank measures the influence taking both the topical similarity between users and the link structure into account [25].

Another two measures are *Retweet Impact* (**RI**) and *Mention Impact* (**MI**) [24]. **RI** estimates the impact of the content created by the user in the aspect of retweets: $RI = RT2 \cdot \log(RT3)$. "The logarithms moderate the impact of overly enthusiastic users who retweet the same content many times" is mentioned in [21], however a user cannot retweet a tweet twice using the current version of Twitter. *RT2* is the number of unique tweets retweeted by other users, *RT3* is the number of unique users who retweeted author's tweets. **MI** estimates the impact of the content created by the user in the aspect of mentions: $M3 \cdot \log(M4) - M1 \cdot \log(M2)$. M1, M2,

M3 and M4 can be seen in Table 1.

ID	Feature
OT1	Number of original tweets
OT2	Number of shared links
OT3	Self-similarity score of similarity to recent tweet to the users previous tweets
OT4	Number of hashtags used
CT1	Number of conversational tweets
CT2	Number of conversational tweets started by the user
RT1	Number of retweets of other's tweets
RT2	Number if unique tweets (*OT1*) retweeted by other users
RT3	Number of unique users who retweeted author's tweets
M1	Number of mentions of other users by the user
M2	Number of unique users mentioned by the user
M3	Number of mentions by other of the user
M4	Number of unique users mentioning the user
G1	Number of topically active followers
G2	Number of topically active bidirectional following including the user
G3	Number of followers tweeting on topic after the author
G4	Number of friends tweeting on topic before the author

Table 1: List of potential metrics [24][21]

Next is *Social Networking Potential* (**SNP**) [23]; the equation is: $\frac{Ir(i)+RMr(i)}{2}$, where $Ir(i)$ means Interactor ratio, and is defined as: $Ir(i) = \frac{RT3+M4}{F1}$, meanings of the separate variables are in table 1. $RMr(i)$ is Retweet and Mention Ratio, and is defined as: $RMr(i) = \frac{\#tweets\ of\ i\ retweeted+\#tweets\ of\ i\ replied}{\#tweet\ of\ i}$. **SNP** considers many kinds of actions except from likes [21], this seems like a good measure to use because of accounting for those different actions rather than just retweets or mentions. 25% of the total importance is the number of published tweets and follow-up relationships while the other 75% conciders the numbers of replies, and the number of followers related to the user through retweets and mentions. The time complexity uses $O(T \cdot k)$ [21] where T is the number of tweets and k is the length of an auxiliary vector.

Two more are *TunkRank* [28] and *UserRank* [29]. They are both adapted from PageRank [27]. TunkRank was the first PageRank translation to be applied to Twitter, and is defined as [28][21]:

$$TunkRank(i) = \sum_{j \epsilon followers(i)} \frac{1 + p \cdot TunkRank(j)}{\#followees\ of\ j}.$$

This method only uses followers/followees only. UserRank was created to measure the influence of a user from their tweets relevance [29][21]:

$$UserRank(i) = \sum_{j \epsilon followers(i)} \frac{1 + \frac{\#followers\ of\ i}{\#tweets\ of\ i} \cdot UserRank(j)}{\#followers\ of\ j}.$$

The benefit of using UserRank versus TunkRank is that they "calculate dynamic coefficient for each user based on a number of his followers and tweets"[29], essentially extending upon Tunkrank to consider other influence metrics like retweets, while TunkRank is an adapted PageRank algorithm without any additional metrics used on-top of that; UserRank requires more data while TunkRank only needs the topological data.

2.6 Sockpuppet Detection

As we mentioned before, sockpuppets reside within online social media platforms; we need a set of methods to be able to detect them to measure their influence.

A formal definition of what a sockpuppet is a fake online identity; a puppetmaster is a person who controls multiple sockpuppet identities and could also be automating posts [30]. They usually are used to try and unfairly support a user's point of view on a topic [1].

Bots are observed to exploit mentions of each other creating a network of these bots to manipulate their influence on the platform [2], and others exploit retweets [5].

There are three main groupings of sockpuppet detection which are: verbal behaviour analysis; non-verbal behaviour analysis and similar-orientation network [1]. Verbal behaviour analysis is based on the textual content

a user tweets out, trying to detect a user from a similar writing style, formally known as authorship attribution (**AA**) [31]. Non-verbal behaviour uses extracted features that capture a user's activity or movements such as times they tweet or tweets that are temporally close together [32]. Finally, a similar-orientation network is based on evaluating the similarity of sentiment orientations among user account pairs to construct a similar-orientation network [33][1].

There has been research using **AA** on textual content from Wikipedia users and their edits to detect sockpuppets [34]. They evaluated 239 features that capture grammatical, stylistic, and formatting the writer used; a state vector machine (**SVM**) was created to perform classification. This method has a high time complexity of $O((NR)^2)$; it is also a verbal behaviour analysis technique [1]. The F-Measure of this particular algorithm is 72%; F-Measure indicates the fraction of valid classifications.

In [30], they identified three features to distinguish legitimate accounts from illegitimate accounts: activity, community, and post features, however, it needs the IPs of the users which we cannot get. The method still could be useful to draw from.

SocksCatch [35] is a non-verbal behaviour analysis technique. It is comprised of three phases: data collection & selection; detection of the sockpuppet accounts using machine learning and finally grouping of sockpuppet accounts using graph theory. This process is complex and lengthy, as seen in figure 2. SocksCatch has a true positive rate of 92.6%; a false positive rate of 7.8%; and an F-Measure of 92.6%. When using an SVM machine learning algorithm, but it can range in between 89% and 95% for correct detection. It is better performing than similar algorithms such as [30] or [36].

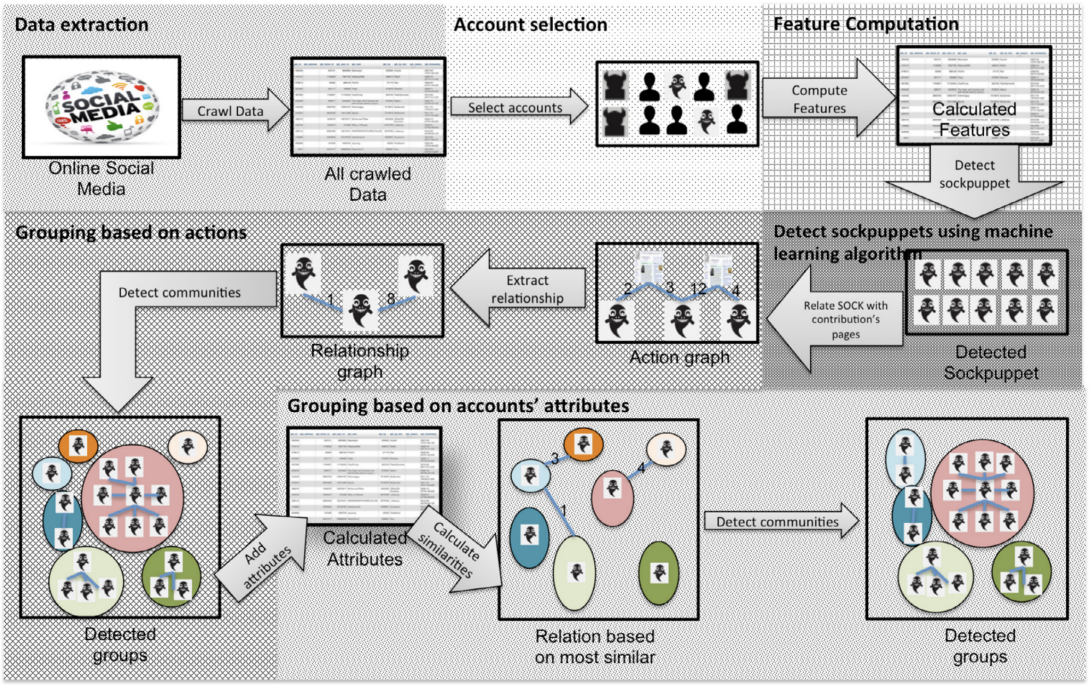


Figure 2: SocksCatch's main steps, image taken from [35]

In [33], they use a similar-orientation network method. The paper mentions how verbal and non-verbal analysis can be avoided by trying on purpose to not fall into either category of analysis by changing their language and more, this makes it difficult to gauge how many sockpuppets that are really within a dataset. They observe that sockpuppets want to recover similar social structures as the sockpuppet user can guarantee a similar propagation impact. They turn the problem into a subgraph similarity problem, which slightly outperforms other detection techniques made in the time the paper was published, 2018. They used a dataset from the social media site *Sina Weibo* for their detection algorithm which is a micro-blogging platform so it can still be applied for other micro-blogging platforms such as Twitter. Three users, a, b, c, each user would have the users they follow and interact with such as mentions or retweets. User a and b are a sockpuppet pair, as they share almost all of the same users they follow and interact with. a and c are not a pair of sockpuppets and only have two mutual users they are following, which shows they are not trying to gain the same social network structure to reach the same propagation impact. Because we are collecting a specific community where there will be a lot of users who follow the same people, but it would seem to be unlikely to have exactly all the same users they are following. This method has the average F-Measure score of 83.5%, which is better than the **AA**

method used on Wikipedia [34], but worse than the SocksCatch [35] method.

In another paper [37], they review current sockpuppet detection methods and if they can be used for real-time detection of bots. An example of a bot post points out a tweet that has very high retweets but two likes, which is explicitly a sockpuppet account that is amplified by the 412 other retweeters which show suspicious activity that would be an indicator that they are a bot or sockpuppet. Their results show that detecting bots or a sockpuppet campaign in real-time is impossible with current researched methods. Fortunately, for this research project, we will not need real-time detection.

3 Research Questions

The overarching question is how much influence do sockpuppets have; we are limiting this to a potential community that could very likely have sockpuppets residing within to then assess the influence the sockpuppets have, and we could compare the score of the sockpuppet users to the average score of the non-sockpuppet detected users.

Our research questions can be split into two main sections, which are the minimal viable questions to complete the main aim of the research, and the extensions to gain more information from the data if we have enough time.

For the main questions, there are four: how do we detect the Twitter Transgender community; how do we measure influence; how do we detect sockpuppets and how much influence do the sockpuppets have. We will be implementing methods covered in the background section and trying out more experimental methods.

How do we detect the Twitter Transgender community? A community can be found from who follows who and who interacts with who explicitly. From 2.4, we will attempt to use the snowball sampling technique with a max of 15 users to be sampled from, we will need to identify large accounts which are part of the trans community. We could also try to search for tweets containing words or short phrases associated with the community. Experimenting with a combination of both where we sample users who have a keyword or phrase in their bio could be another avenue we could explore to see its effectiveness. The keywords and phrases we will use for tweets are: "#TransRightsAreHumanRights", "#TransLiberationNow", "#trans", "#transgender", "#enby" and "#nonbinary"; we chose this as a starting point as each of the hashtags are relevant to the trans community and they would be discussed by users within the community mostly.

How do we measure influence? We want to know how much influence a user has over the network to determine who is leading the conversation the most within the network and who users interact with the most. In 2.5 we discuss different methods of measuring influence, we would like to test using basic measures: RI and/or MI [24], as well as more complex methods: **SNP** [23] and/or **UserRank** [29] depending on time constraints. We can apply these measures to users in the collected network to grasp the general influence of users within the network, and then once sockpuppet detection has been complete, we shall select the measures for those users specifically.

How do we detect sockpuppets? To be able to measure the influence of sockpuppets we need to discover them. Sockpuppets try to alter the discourse in a community to try to harm them or to push a particular view. From 2.6, we shall implement methods researched; these include subgraph similarity matching technique [33], and depending on time constraints try to use authorship attribution [34][31] and/or a non-verbal analysis such as SocksCatch [35]; however, we do not necessarily need to know the puppetmaster groupings unless we want to measure a particular groups influence as an extension.

How much influence do the sockpuppets have? This is a combination of the influence question and the sockpuppet question, once both are complete we can use the best performing method for them to then collect the statistics of the influence the sockpuppets have over the community.

Extension questions are based around natural language processing (**NLP**): What topics are the tweets about? What are the most used words? This would be using a simple approach, using frequency analysis to examine the most used words used additionally to explore a singular keyword topic associated with each tweet to see what topic is used. This is to see if the sockpuppets may only tweet about one topic area on a higher frequency than other members of the transgender community which could be a rather simplistic approach to detecting sockpuppets within the network using NLP.

4 Professional Considerations

4.1 BCS Code of Conduct

The BCS Code of Conduct can be seen in [38].

4.1.1 Public Interest

a. have due regard for public health, privacy, security and wellbeing of others and the environment.

We will be collecting data on a sensitive community without the direct consent of those users, what we will do is pseudo-anonymise the data and only publish aggregate data so no content can be reversed searched to target a particular individual. The data collected will be stored on secure servers hosted by the Univesity of Sussex.

b. have due regard for the legitimate rights of Third Parties*.

Yes, we will follow the Twitter developer policy while using their API.

c. conduct your professional activities without discrimination on the grounds of sex, sexual orientation, marital status, nationality, colour, race, ethnic origin, religion, age or disability, or of any other condition or requirement

We will not discriminate on any basis, we are looking to protect transgender people computationally through detecting sockpuppets, not to further promote disinformation or discrimination.

d. promote equal access to the benefits of IT and seek to promote the inclusion of all sectors in society wherever opportunities arise.

Yes, we want to include all. If this research goes well, this could be used to detect potential sockpuppets who may harm the transgender community or people within the transgender community who are trying to harm others. The data collected will not parse out particular identities to discriminate on background, however, users' data collected on all will have to have internet access to create a Twitter account.

4.1.2 Professional Competence and Integrity

a. only undertake to do work or provide a service that is within your professional competence.

This project will be implementing already existing algorithms but to a new problem area or even the same problem area for instance influence measures for Twitter. Our professor/supervisor has said this is challenging but believes we can implement this.

b. NOT claim any level of competence that you do not possess.

We shall not claim competence in any area I do not possess, the course I have taken touches on each area within the research questions to give a good foundation as well as my research into the separate areas gives us an acceptable amount of competence.

develop your professional knowledge, skills, and competence on a continuing basis, maintaining awareness of technological developments, procedures, and standards that are relevant to your field.

We shall continue to develop our professional skills, knowledge, and competence and maintain awareness of advancements, procedures, and standards that are relevant to our field.

d. ensure that you have the knowledge and understanding of Legislation* and that you comply with such Legislation, in carrying out your professional responsibilities.

We shall not break any legislation during the research's lifetime.

e. respect and value alternative viewpoints and, seek, accept and offer honest criticisms of work.

We shall give updates regularly to our professor and other peers on our work and will take on criticism of our work to further improve the piece of research.

f. avoid injuring others, their property, reputation, or employment by false or malicious or negligent action or inaction.

We shall avoid damaging others' reputation, property, employment by false or negligent action or inaction. No data displayed to readers will be reversible to a specific user.

g. reject and will not make any offer of bribery or unethical inducement.

We shall not be tempted through bribery or unethical inducement as well as not participate in encouraging it.

4.1.3 Duty to Relevant Authority

The relevant authority for this research is the informatics department at the University of Sussex.

a. carry out your professional responsibilities with due care and diligence per the Relevant Authority's requirements whilst exercising your professional judgement at all times.

We shall ensure the research is submitted by the deadline to meet the requirements set by the university.

b. seek to avoid any situation that may give rise to a conflict of interest between you and your Relevant Authority.

We shall meet this by communicating between us and our supervisor through the project and handing in the relevant work when needed for the research questions.

c. accept professional responsibility for your work and for the work of colleagues who are defined in a given context as working under your supervision.

We accept full responsibility for my work.

d. NOT disclose or authorise to be disclosed, or use for personal gain or to benefit a third party, confidential information except with the permission of your Relevant Authority, or as required by Legislation.

We will not reveal any third party or confidential information, especially due to the sensitive group and community members within which will be pseudo-anonymised. I shall only discuss such information with those associated with the project.

e. NOT misrepresent or withhold information on the performance of products, systems, or services (unless lawfully bound by a duty of confidentiality not to disclose such information), or take advantage of the lack of relevant knowledge or inexperience of others.

We shall not misrepresent or withhold information on the performance of the algorithms implemented or take advantage of the lack of relevant knowledge or inexperience of others. Data will be clearly displayed which shows an honest representation of the performance of the implemented algorithms.

4.1.4 Duty to the Profession

a. accept your personal duty to uphold the reputation of the profession and not take any action which could bring the profession into disrepute.

We shall accept our personal duty to uphold the reputation of the profession and not take any action which could bring the profession into disrepute.

b. seek to improve professional standards through participation in their development, use and enforcement.

We shall seek to improve professional standards through participation in their development, use, and enforcement.

c. uphold the reputation and good standing of BCS, the Chartered Institute for IT.

We shall uphold the reputation and good standing of BCS.

d. act with integrity and respect in your professional relationships with all members of BCS and with members of other professions with whom you work in a professional capacity.

We shall treat all BSC members in a professional. respectful and integral way.

e. encourage and support fellow members in their professional development.

We shall encourage and support other members in their professional development.

4.2 Ethical Considerations

Ethical considerations within this project include:

1. Collecting data from users who do not know (but have accepted the Twitter ToS)
2. A large number of tweets being collected
3. The data collected contains data of a particularly sensitive community, transgender individuals

Due to these ethical considerations will have to go through a full ethical review and is classified as high risk. The ethical application is soon to be submitted after the data management plan and protocol document are fully complete.

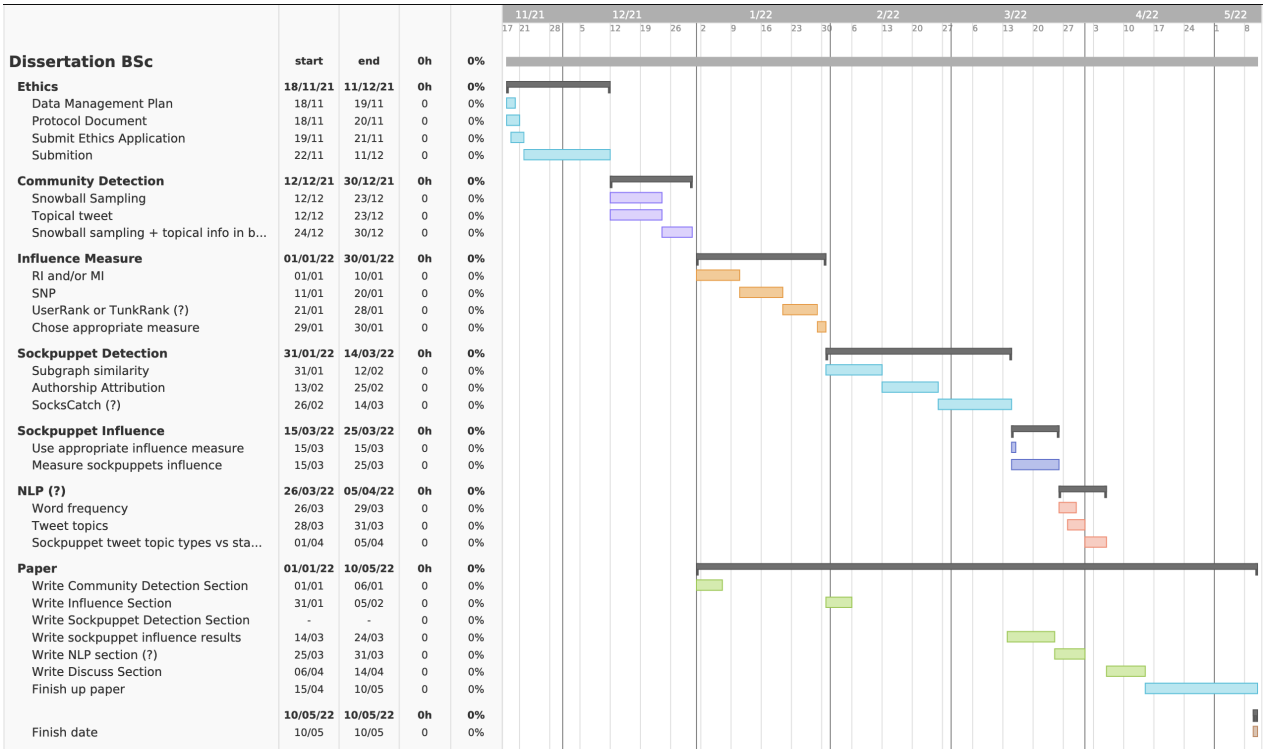
The data collected will involve human participants who will not know their data is being processed, and data collected is from a minority, unrepresented, and vilified group, and in no way do we want to harm the community.

The data will be from Twitter with their API. Data collected could be all types of data surrounding the Twitter trans community, this could be potentially identifying data or abusive content, this content will not be displayed within the data as it will exclusively be aggregate data. All data being pseudo-anonymised and secured on University of Sussex servers; when the project is complete, all collected data will be deleted.

We do not have an estimate for the number of tweets that will be collected, however, tweets will be collected from between two points in time from the community of users captured; this will most likely be in the hundreds of thousands. All data stored is relevant to the research, the other data will be discarded. The data which will be displayed within the paper will only show aggregate data, so there is no way for a person to reverse-search a particular piece of text to identify a user.

5 Plan

See Figure 3, *(?)* means extension.



Dissertation BSc	start	end	0h	0%
Ethics	**18/11/21**	**11/12/21**	**0h**	**0%**
Data Management Plan	18/11	19/11	0	0%
Protocol Document	18/11	20/11	0	0%
Submit Ethics Application	19/11	21/11	0	0%
Submition	22/11	11/12	0	0%
Community Detection	**12/12/21**	**30/12/21**	**0h**	**0%**
Snowball Sampling	12/12	23/12	0	0%
Topical tweet	12/12	23/12	0	0%
Snowball sampling + topical info in b...	24/12	30/12	0	0%
Influence Measure	**01/01/22**	**30/01/22**	**0h**	**0%**
RI and/or MI	01/01	10/01	0	0%
SNP	11/01	20/01	0	0%
UserRank or TunkRank (?)	21/01	28/01	0	0%
Chose appropriate measure	29/01	30/01	0	0%
Sockpuppet Detection	**31/01/22**	**14/03/22**	**0h**	**0%**
Subgraph similarity	31/01	12/02	0	0%
Authorship Attribution	13/02	25/02	0	0%
SocksCatch (?)	26/02	14/03	0	0%
Sockpuppet Influence	**15/03/22**	**25/03/22**	**0h**	**0%**
Use appropriate influence measure	15/03	15/03	0	0%
Measure sockpuppets influence	15/03	25/03	0	0%
NLP (?)	**26/03/22**	**05/04/22**	**0h**	**0%**
Word frequency	26/03	29/03	0	0%
Tweet topics	28/03	31/03	0	0%
Sockpuppet tweet topic types vs sta...	01/04	05/04	0	0%
Paper	**01/01/22**	**10/05/22**	**0h**	**0%**
Write Community Detection Section	01/01	06/01	0	0%
Write Influence Section	31/01	05/02	0	0%
Write Sockpuppet Detection Section	-	-	0	0%
Write sockpuppet influence results	14/03	24/03	0	0%
Write NLP section (?)	25/03	31/03	0	0%
Write Discuss Section	06/04	14/04	0	0%
Finish up paper	15/04	10/05	0	0%
	10/05/22	**10/05/22**	**0h**	**0%**
Finish date	10/05	10/05	0	0%

Figure 3: Project plan timeline

6 Log

- **01/10/2021** - Initial meeting happens to discuss the project and how to tackle the process of exploring the existing ideas within problem domain.
- **07/10/2021** - Meeting about the general project idea.
- **14/10/2021** - Meeting about the ethical issues.
- **21/10/2021** - Continuation of ethical issues and created questions to ask Lauren Shukru about twitter data storage.
- **22/10/2021** - Meeting with Supervisor and Lauren Shukru about ethical concerns with twitter data.
- **29/10/2021** - Talked about project slow progress issues and set deadline for the project proposal.
- **02/11/2021** - Interim report progress and paper locations for more background research.
- **09/11/2021** - Interim report update with new text, the sections to do next and reference issues (fixed with switching from BibTex to BibLaTeX in the LaTeX document).

7 References

[1] A. Alharbi, H. Dong, X. Yi, Z. Tari, and I. Khalil, "Social media identity deception detection: A survey," *ACM Comput. Surv.*, vol. 54, no. 3, Apr. 2021, ISSN: 0360-0300. DOI: 10.1145/3446372.

[2] M. C. Benigni, K. Joseph, and K. M. Carley, "Bot-ivistm: Assessing information manipulation in social media using network analytics," in *Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining*, N. Agarwal, N. Dokoohaki, and S. Tokdemir, Eds. Cham: Springer International Publishing, 2019, pp. 19–42, ISBN: 978-3-319-94105-9. DOI: 10.1007/978-3-319-94105-9_2. [Online]. Available: https://doi.org/10.1007/978-3-319-94105-9_2.

[3] S. Bradshaw and P. N. Howard, "Troops, trolls and troublemakers: A global inventory of organized social media manipulation," p. 37,

[4] Twitter. "Twitter API for academic research — products." (), [Online]. Available: https://developer.twitter.com/en/products/twitter-api/academic-research (visited on 11/08/2021).

[5] O. Varol, E. Ferrara, C. Davis, F. Menczer, and A. Flammini, "Online human-bot interactions: Detection, estimation, and characterization," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 11, no. 1, pp. 280–289, May 3, 2017, Section: Full Papers. [Online]. Available: https://ojs.aaai.org/index.php/ICWSM/article/view/14871.

[6] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Communications of the ACM*, vol. 59, no. 7, pp. 96–104, Jun. 24, 2016, ISSN: 0001-0782. DOI: 10.1145/2818717.

[7] S. Wojcik, S. Messing, A. W. Smith, L. Rainie, and P. Hitlin, "Bots in the twittersphere," Pew Research Center, Report, Apr. 9, 2018, Journal Abbreviation: An estimated two-thirds of tweeted links to popular websites are posted by automated accounts – not human beings. [Online]. Available: https://apo.org.au/node/141291.

[8] B. Heredia, J. Prusa, and T. Khoshgoftaar, "Exploring the effectiveness of twitter at polling the united states 2016 presidential election," in *2017 IEEE 3rd International Conference on Collaboration and Internet Computing (CIC)*, Oct. 2017, pp. 283–290. DOI: 10.1109/CIC.2017.00045.

[9] E. Ferrara, "What types of COVID-19 conspiracies are populated by twitter bots?" *First Monday*, May 19, 2020, ISSN: 1396-0466. DOI: 10.5210/fm.v25i6.10633. arXiv: 2004.09531.

[10] T. U. G. H. Office. "Hate crime, england and wales, 2020 to 2021," GOV.UK. (), [Online]. Available: https://www.gov.uk/government/statistics/hate-crime-england-and-wales-2020-to-2021/hate-crime-england-and-wales-2020-to-2021 (visited on 10/27/2021).

[11] Brandwatch. "The scale of transphobia online," Brandwatch. (), [Online]. Available: https://www.brandwatch.com/reports/transphobia/ (visited on 10/27/2021).

[12] I. C. Education. "What is an application programming interface (API)." (Oct. 15, 2021), [Online]. Available: https://www.ibm.com/cloud/learn/api (visited on 11/08/2021).

[13] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and identifying communities in networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 9, p. 2658, Mar. 2, 2004. DOI: 10.1073/pnas.0400054101.

[14] W. O. Bockting, M. H. Miner, R. E. Swinburne Romine, A. Hamilton, and E. Coleman, "Stigma, mental health, and resilience in an online sample of the US transgender population," *American Journal of Public Health*, vol. 103, no. 5, pp. 943–951, May 2013, ISSN: 0090-0036. DOI: 10.2105/AJPH.2013.301241.

[15] O. Jenzen, "Trans youth and social media: Moving between counterpublics and the wider web," *Gender, Place & Culture*, vol. 24, no. 11, pp. 1626–1641, Nov. 2, 2017, Publisher: Routledge _eprint: https://doi.org/10.1080/0966 ISSN: 0966-369X. DOI: 10.1080/0966369X.2017.1396204.

[16] T. D. Jayawickrama. "Community detection algorithms," Medium. (Feb. 1, 2021), [Online]. Available: https://towardsdatascience.com/community-detection-algorithms-9bd8951e7dae (visited on 11/06/2021).

[17] M. C. Benigni, K. Joseph, and K. M. Carley, "Online extremism and the communities that sustain it: Detecting the ISIS supporting community on twitter," *PLOS ONE*, vol. 12, no. 12, e0181405, Dec. 1, 2017, Publisher: Public Library of Science, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0181405.

[18] L. A. Goodman, "Snowball sampling," *The Annals of Mathematical Statistics*, vol. 32, no. 1, pp. 148–170, Mar. 1961, Publisher: Institute of Mathematical Statistics, ISSN: 0003-4851, 2168-8990. DOI: 10.1214/aoms/1177705148.

[19] A. Campan, T. Atnafu, T. M. Truta, and J. Nolan, "Is data collection through twitter streaming API useful for academic research?" In *2018 IEEE International Conference on Big Data (Big Data)*, Dec. 2018, pp. 3638–3643. DOI: 10.1109/BigData.2018.8621898.

[20] L. Ben Jabeur, L. Tamine, and M. Boughanem, "Active microbloggers: Identifying influencers, leaders and discussers in microblogging networks," in *String Processing and Information Retrieval*, L. Calderón-Benavides, C. González-Caro, E. Chávez, and N. Ziviani, Eds., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2012, pp. 111–117. DOI: 10.1007/978-3-642-34109-0_12.

[21] F. Riquelme and P. González-Cantergiani, "Measuring user influence on twitter: A survey," *Information Processing & Management*, vol. 52, no. 5, pp. 949–975, Sep. 1, 2016, ISSN: 0306-4573. DOI: 10.1016/j.ipm.2016.04.003.

[22] M. Cha, H. Haddadi, F. Benevenuto, and K. Gummadi, "Measuring user influence in twitter: The million follower fallacy," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 4, no. 1, pp. 10–17, May 16, 2010, Number: 1, ISSN: 2334-0770. [Online]. Available: https://ojs.aaai.org/index.php/ICWSM/article/view/14033.

[23] I. Anger and C. Kittl, "Measuring influence on twitter," in *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, ser. i-KNOW '11, New York, NY, USA: Association for Computing Machinery, Sep. 7, 2011, pp. 1–4. DOI: 10.1145/2024288.2024326.

[24] A. Pal and S. Counts, "Identifying topical authorities in microblogs," in *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, ser. WSDM '11, Hong Kong, China: Association for Computing Machinery, 2011, pp. 45–54. DOI: 10.1145/1935826.1935843.

[25] J. Weng, E.-P. Lim, J. Jiang, and Q. He, "Twitterrank: Finding topic-sensitive influential twitterers," in *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, ser. WSDM '10, New York, New York, USA: Association for Computing Machinery, 2010, pp. 261–270. DOI: 10.1145/1718487.1718520.

[26] D. Quercia, J. Ellis, L. Capra, and J. Crowcroft, "In the mood for being influential on twitter," in *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, Oct. 2011, pp. 307–314. DOI: 10.1109/PASSAT/SocialCom.2011.27.

[27] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks and ISDN Systems*, vol. 30, no. 1, pp. 107–117, Apr. 1998, ISSN: 01697552. DOI: 10.1016/S0169-7552(98)00110-X.

[28] "A twitter analog to PageRank," The Noisy Channel. (Jan. 13, 2009), [Online]. Available: https://thenoisychannel.com/2009/01/13/a-twitter-analog-to-pagerank/ (visited on 11/12/2021).

[29] T. Majer and M. Šimko, "Leveraging microblogs for resource ranking," in *SOFSEM 2012: Theory and Practice of Computer Science*, M. Bieliková, G. Friedrich, G. Gottlob, S. Katzenbeisser, and G. Turán, Eds., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2012, pp. 518–529. DOI: 10.1007/978-3-642-27660-6_42.

[30] S. Kumar, J. Cheng, J. Leskovec, and V. Subrahmanian, "An army of me: Sockpuppets in online discussion communities," in *Proceedings of the 26th International Conference on World Wide Web*, ser. WWW '17, Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, Apr. 3, 2017, pp. 857–866. DOI: 10.1145/3038912.3052677.

[31] R. M. Coyotl-Morales, L. Villaseñor-Pineda, M. Montes-y-Gómez, and P. Rosso, "Authorship attribution using word sequences," in *Progress in Pattern Recognition, Image Analysis and Applications*, J. F. Martínez-Trinidad, J. A. Carrasco Ochoa, and J. Kittler, Eds., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2006, pp. 844–853. DOI: 10.1007/11892755_87.

[32] M. Tsikerdekis and S. Zeadally, "Multiple account identity deception detection in social media using nonverbal behavior," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 8, pp. 1311–1321, 2014. DOI: 10.1109/TIFS.2014.2332820.

[33] J. Wang, W. Zhou, J. Li, Z. Yan, J. Han, and S. Hu, "An online sockpuppet detection method based on subgraph similarity matching," in *2018 IEEE Intl Conf on Parallel Distributed Processing with Applications, Ubiquitous Computing Communications, Big Data Cloud Computing, Social Computing Networking, Sustainable Computing Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom)*, Dec. 2018, pp. 391–398. DOI: 10.1109/BDCloud.2018.00067.

[34] T. Solorio, R. Hasan, and M. Mizan, "A case study of sockpuppet detection in Wikipedia," in *Proceedings of the Workshop on Language Analysis in Social Media*, Atlanta, Georgia: Association for Computational Linguistics, Jun. 2013, pp. 59–68. [Online]. Available: https://aclanthology.org/W13-1107.

[35] Z. Yamak, J. Saunier, and L. Vercouter, "SocksCatch: Automatic detection and grouping of sockpuppets in social media," *Knowledge-Based Systems*, vol. 149, pp. 124–142, Jun. 1, 2018, ISSN: 0950-7051. DOI: 10.1016/j.knosys.2018.03.002.

[36] M. Tsikerdekis and S. Zeadally, "Multiple account identity deception detection in social media using nonverbal behavior," *Trans. Info. For. Sec.*, vol. 9, no. 8, pp. 1311–1321, Aug. 2014, ISSN: 1556-6013. DOI: 10.1109/TIFS.2014.2332820.

[37] O. Beatson, R. Gibson, M. C. Cunill, and M. Elliot, "Automation on twitter: Measuring the effectiveness of approaches to bot detection," *Social Science Computer Review*, p. 08 944 393 211 034 991, Aug. 6, 2021, Publisher: SAGE Publications Inc, ISSN: 0894-4393. DOI: 10.1177/08944393211034991.

[38] "BCS code of conduct — BCS." (), [Online]. Available: https://www.bcs.org/membership/become-a-member/bcs-code-of-conduct/ (visited on 11/14/2021).

8 Appendix

Assessing the influence of sockpuppets on Twitter within the trans community: Project Proposal

Supervisor: Professor Luc Berthouze
Candidate Number: 215758

October 10, 2021

1 Introduction

2 Aims & Objectives

2.1 Aims

2.1.1 Purpose and intention

The purpose of this project is to see how sockpuppets influence others to do or think something, more specifically how well they can influence the current discourse in a community, we have chosen the trans community due to the reasons described in the motivations subsection.

The intention is to:

- Map the trans community
- Collect tweets in between two points in time
- Create aggregate data on the general statistics of the overall community
- Detect sockpuppets
- Create aggregate data on the general statistics on the sockpuppets
- Create aggregate data on the influence of sockpuppets on discourse within the community

2.1.2 Motivation

We are very interested in how social media affects people and society as a whole, and how ideas can propagate within communities.

There is a rise in hate crimes in the UK. Race, sexual orientation, disability, and transgender hate crime strands have risen over the year, with race being the highest at a 12% increase, and over the last five years, there has been a 120% increase in transphobic hate crimes[11].

Transphobia online, as well as any other kind of bigotry, is abundant, with Twitter having 12% ($\frac{789615}{5547445}$) of the conversation around trans individuals being abusive[5].

The first iteration of the idea was to look into transphobic anti-trans communities on Twitter. The next iteration, and current, is to access the trans community instead, for the sake of our mental health, to generally look at how the discourse is led, if there are sockpuppets within the trans community, as well as what topics/language is mostly used within the community.

2.2 Objectives

2.2.1 Primary objectives

Our primary objectives are:

- How to detect and map a community?
- How to detect sockpuppets?
- How to measure effectiveness and influence of sockpuppets?

2.2.2 Extensions

The extensions are:

- How has the community's engagement risen over time?
- What topics and keywords do the community mention the most?
- What are the sockpuppets' main topics and keywords used?
- How to visualise the data in an easily accessible way?

3 Background

3.1 General

On Twitter, a user can easily create an account using an email address or phone number, this opens up the issues of users exploiting this to create malicious user accounts which are either controlled automatically via the Twitter API or a human user who has another account which they use as an anonymous pseudonym to hide behind and express their true views online[2]. Both can be labeled under the category of a sockpuppet. Sockpuppets have a puppet master, the real human who owns the accounts credentials for the account and can orchestrate multiple of them [3].

Sockpuppets can affect online social media platforms, such as Twitter, to spread disinformation, these sock puppets could just be random users or more organised such as state sponsored[4]. Twitter in 2019 had 290.5 million users[7], this is a huge attack vector for puppet masters to use.

During 2017, within this paper, it is thought that between 9% and 15% percent of Twitter accounts are bots and that bots use retweeting strategies to target specific communities of people [13]. Not all bots are bad, some are neutral or designed to be helpful, however, they have been used to infiltrate political discourse and manipulate them[**?**][14]. It was estimated that $\frac{2}{3}$ of tweeted links were posted by these automated accounts[14].

Data on past twitter data has been able to collect statistics which show similar margins to polls collected during the election period in 2016 from three million tweets[10].

During the COVID-19 pandemic, many conspiracies were propagating on Twitter, and that there exists a set of bots that mainly posts conspiracy theories of the political type around the COVID-19 topic[8].

3.2 Methods

Community detection for Twitter can be approached in several different ways, two ways which have been implemented is snowball detection[3][9] and using search terms related to the community[1]. Sockpuppet detection has many methods, they typically fall under three different categories: Verbal Behaviour Analysis, Non-verbal Behaviour Analysis, and Similar-orientation Network[2], methods such as a case for detecting sockpuppets on Wikipedia uses authorship attribution[6][12] this particular method would be under Verbal Behaviour Analysis.

4 Relation to the course

The databases module will be utilised for the storage of user data, to use a relational database between users following each other and the tweets a user has made.

Data Structures & Algorithms module, I will be using the graph data structure for the community to be represented in, which can be used for visualisation and sockpuppet detection.

The Natural Language Engineering module will be used for the language and topic analysis, as well as contributing to the sockpuppet detection.

Fundamentals of Machine Learning module will be used for sockpuppet detection.

General statistics graphical representation from Natural Language Engineering, Fundamentals of Machine Learning, Acquired Intelligence & Adaptive Behaviour and Computer Vision.

General programming skills learned from the course as a whole.

5 Resources required

We will need access to the Twitter API to pull users and tweet data. The specific tweet data which needs to be stored is:

- User Data

- User id

  Reason: To be able to go through a users tweets and map out the connections between users.

- User followers

  Reason: For inbound connections on the graph for community detection.

- User following

  Reason: For outbound connections on the graph for community detection.

- Total tweets made

  Reason: A measure for influence of a user.

- User creation date

  A group of bots could be created on the same day, could be helpful for sockpuppet detection.

• Tweet Data

- Tweet ID

  Reason: Is required by twitter as well as used to map out a tweets replies.

- User ID

  Reason: To link back a tweet to a user, for example using the text content from a tweet linking back to a user for sockpuppet detection.

- In reply status ID

  Reason: If the tweet is in response to another tweet store it to collect even more tweets for sockpuppet analysation.

- Text content

  Reason: This is where we can analyse the language of the community and an vector to approach sockpuppets detection from.

- Date posted

  Reason: This is to map the tweets on a timeline to see how the conversation within the subcommunity changes over time.

- Number of likes

  Reason: A measure for engagement to see how much influence a tweet and inherently a user has, is positive only.

- Number of retweets

  Reason: This is a form of engagement measure to see how much influence a tweet and inherently a user has, is ambiguous for why.

- Number of replies

  Reason: This is another form of engagement measure to see how much users interact with it, can be for a variety of reasons a user replies.

Additionally, to comply with GDPR, we also need to securely store the collected data, as an encrypted SQLite database file where only the people involved with the paper can access such key.

Lastly, we need python, as it is easy to prototype with and particularly well suited for data science from notable libraries in its ecosystem, e.g., *NumPy*, *MatPlotLib* and *NLTK*.

6 Plan of study

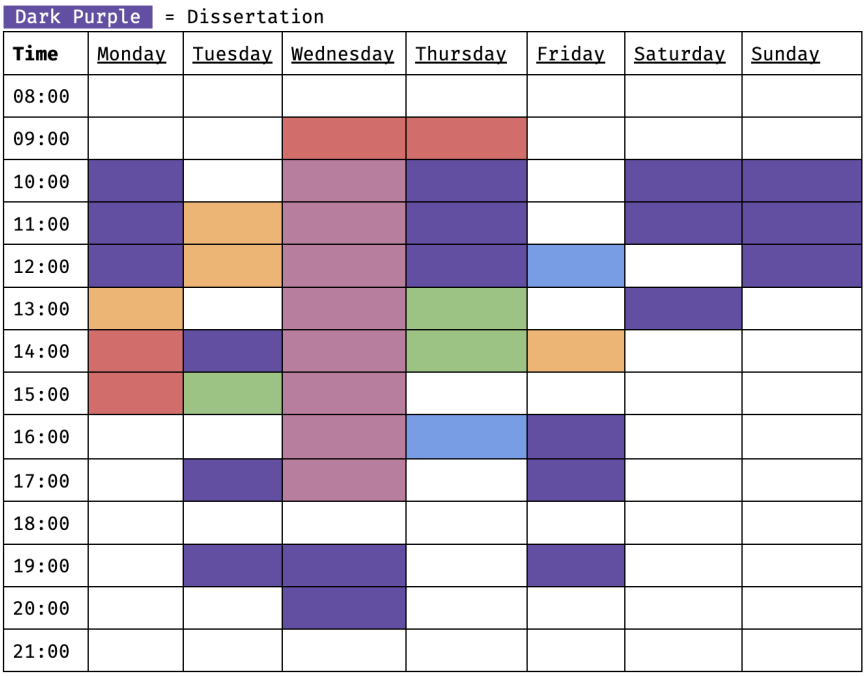


Figure 1: Study timetable. Dark purple is dissertation study.

7 Related projects

Simon Wibberley & Liv Livesey - **NLP and ML on Twitter for policy-making**

8 References

[1] Norah Abokhodair, Daisy Yoo, and David W. McDonald. Dissecting a social botnet: Growth, content and influence in twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, pages 839–851. Association for Computing Machinery.

[2] Ahmed Alharbi, Hai Dong, Xun Yi, Zahir Tari, and Ibrahim Khalil. Social media identity deception detection: A survey. 54(3):1–35.

[3] Matthew C. Benigni, Kenneth Joseph, and Kathleen M. Carley. Bot-ivistm: Assessing information manipulation in social media using network analytics. In Nitin Agarwal, Nima Dokoohaki, and Serpil Tokdemir, editors, *Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining*, Lecture Notes in Social Networks, pages 19–42. Springer International Publishing.

[4] Samantha Bradshaw and Philip N Howard. Troops, trolls and troublemakers: A global inventory of organized social media manipulation. page 37.

[5] Brandwatch. The scale of transphobia online. https://www.brandwatch.com/reports/transphobia/.

[6] Rosa María Coyotl-Morales, Luis Villaseñor-Pineda, Manuel Montes-y Gómez, and Paolo Rosso. Authorship attribution using word sequences. In José Francisco Martínez-Trinidad, Jesús Ariel Carrasco Ochoa, and Josef Kittler, editors, *Progress in Pattern Recognition, Image Analysis and Applications*, Lecture Notes in Computer Science, pages 844–853. Springer.

[7] Statista Research Department. Twitter: number of users worldwide 2019-2020.

[8] Emilio Ferrara. What types of COVID-19 conspiracies are populated by twitter bots?

[9] Leo A. Goodman. Snowball sampling. 32(1):148–170. Publisher: Institute of Mathematical Statistics.

[10] Brian Heredia, Joseph Prusa, and Taghi Khoshgoftaar. Exploring the effectiveness of twitter at polling the united states 2016 presidential election. In *2017 IEEE 3rd International Conference on Collaboration and Internet Computing (CIC)*, pages 283–290.

[11] The UK Government Home Office. Hate crime, england and wales, 2020 to 2021. https://www.gov.uk/government/statistics/hate-crime-england-and-wales-2020-to-2021/hate-crime-england-and-wales-2020-to-2021.

[12] Thamar Solorio, Ragib Hasan, and Mainul Mizan. A case study of sockpuppet detection in wikipedia. In *Proceedings of the Workshop on Language Analysis in Social Media*, pages 59–68. Association for Computational Linguistics.

[13] Onur Varol, Emilio Ferrara, Clayton A Davis, Filippo Menczer, and Alessandro Flammini. Online human-bot interactions: Detection, estimation, and characterization. page 10.

[14] Stefan Wojcik, Solomon Messing, Aaron W. Smith, Lee Rainie, and Paul Hitlin. Bots in the twittersphere. Journal Abbreviation: An estimated two-thirds of tweeted links to popular websites are posted by automated accounts – not human beings.

9 Interim Log

9.1 Materials

1. Botometer publications : https://botometer.osome.iu.edu/publications
2. Time series clustering : https://towardsdatascience.com/time-series-clustering-deriving-trends-and-arch
3. Community detection algorithms : https://towardsdatascience.com/community-detection-algorithms-9bd8951
4. Snowball sampling for Twitter Research : https://irepeat.wordpress.com/2010/12/07/snowball-sampling-for-t
5. Collecting tweets : https://mediaeffectsresearch.wordpress.com/collecting-tweets/
6. Tweepy : https://www.tweepy.org/
7. Twitter research API : https://developer.twitter.com/en/products/twitter-api/academic-research
8. Tweepy documentation : https://docs.tweepy.org/en/v3.5.0/index.html
9. Tweepy examples : https://github.com/tweepy/examples
10. NetworkX : https://networkx.org/
11. matplotlib : https://matplotlib.org/
12. pandas : https://pandas.pydata.org/
13. Extract someone's tweet using tweepy : https://fairyonice.github.io/extract-someones-tweet-using-tweepy.html
14. https://towardsdatascience.com/twitter-data-mining-measuring-users-influence-ef76c9badfc0

9.2 Meetings

- **01/10/2021** - Initial meeting happens to discuss the project and how to tackle the process of exploring the existing ideas within problem domain.
- **07/10/2021** - Meeting about the general project idea.
- **14/10/2021** - Meeting about the ethical issues.
- **21/10/2021** - Continuation of ethical issues and created questions to ask Lauren Shukru about twitter data storage.
- **22/10/2021** - Meeting with Supervisor and Lauren Shukru about ethical concerns with twitter data.
- **28/10/2021** - Talked about project slow progress issues and set deadline for the project proposal.