

Understanding trust in social media: Twitter

Catherine Ives-Keeler¹, Oliver Buckley¹, and Jason Lines¹

School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ, UK
{c.ives-keeler, o.buckley, j.lines}@uea.ac.uk

Abstract. In this paper we investigate how users can be perceived on Twitter by looking at a selection of tweets, and how the type of personality traits and language can effect trust. We present participants with a selection of tweets and gather their initial opinions of the Twitter users presented by using a Likert scale¹ and free text box for participants to share their opinions, hosted on Microsoft forms. This paper presents preliminary results based on the data gathered from a questionnaire created by the researcher, highlighting factors that impact how participants perceived the Twitter users. It was found that participants valued content, profile pictures and the display name more than likes and retweets. They did not like the more aggressive or opinionated users and had more of a neutral or positive reaction to the more light-hearted users.

Keywords: Trust · Misinformation · Twitter

1 Introduction

With the increased popularity of social media platforms, such as Twitter, comes an increase in social media bots and the spread of misinformation[1]. Twitter is an open platform which has made the creation and automation of bots easy for users to set up and post tweets on their behalf. Misinformation is a growing problem on social media, especially during the pandemic era, as discussed by H. Rosenberg et al. [2], so it is important to understand factors that impact what users trust on social media, when and why they share information.

Understanding trust in social media is vital, including assessing factors which make users want to tweet, retweet or like certain pieces of information or news, and what are users more sceptical about. If we can pin-point different characteristics or personality traits, this could help determine factors that lead to how information spreads around social media. In this paper we want to investigate the spread of information and factors that make users more or less likeable and how a select set of tweets from a user can be perceived. We hope the outcome of this experiment will give us an insight as to the type of language and personality traits that participants find most trustworthy and engaging when using social media. The outcome of this experiment could lead to further work including how fake news spreads on social media, and how misinformation can spread due to a large number of users retweeting incorrect information.

¹ <https://www.simplypsychology.org/likert-scale.html>

2 Background and relevant literature

The motivation behind this study is to understand trust within social media, with a focus on why a Twitter user may be perceived as more or less trustworthy. One particular area of interest is how the language used can alter our perceptions of a user, and whether it can make a tweet more appealing to like or retweet. With the growth of social media in recent years, studies have emerged looking at personality and perception on social media, as discussed by Qiu L. et al [3]. In this paper they looked at measuring the big 5 personality traits by collecting tweets from users over a 1 month period to observe if personalities can manifest in micro-blogging, the chosen social media platform for this study was Twitter. It was found that Out of the 5 traits only agreeableness and neuroticism could be judged accurately, they judged these traits by associating posts with specific linguistic markers, and observers rated the participants' personality of the basis of their tweets.

Another study that focused on predicting the big 5 personality traits from social media posts was discussed by Azucar, D et al [4] in which their research aimed to develop automated methods to extract and analyse these digital footprints to predict personality traits. In this study they conducted a series of meta-analyses to determine the predictive power of digital footprints collected from social media over big 5 personality traits, finding that the accuracy of predictions is consistent across big 5 traits, and that accuracy improves when analyses include demographics and multiple types of digital footprints. Both of the papers discussed motivated our research and help to highlight the importance of personality on social media and how it can be measured.

Research conducted by D. Sterrett et al. [5] discuss what news to trust on social media and how social media platforms are becoming increasingly popular news sources. It was found that two cues could impact opinions of news on social media: the trustworthiness of the person who shares a story and the credibility of the news outlet reporting the story. This research was part of the motivation behind our study, highlighting how news can spread around social media and what factors can effect its trustworthiness. One of the interesting areas focused in this study is if people are more likely to trust an article on social media if it is shared by a public figure they trust than by one they do not trust.

These studies highlight the importance of understanding personality types on social media, by researching and understanding areas around predicting personality on social media [6] and who to trust we were motivated to create this study. Our hypothesis for this study is that users who come across as having strong personalities or strong views will generate stronger opinions and be more actively disliked than those who share more typical, 'mundane' posts.

3 Method

3.1 Procedure

For this study it was decided that a quick and effective way of collecting insightful results in a relatively short period of time was a questionnaire, this is the best

platform to gather the information we wanted to analyse. We decided to use Microsoft forms as a link to a form can be easily created and shared, results gathered on this platform can easily be exported to an excel file for analysis and it had the functionality to have a Likert scale and optional free text responses.

A questionnaire was designed and created by the researcher to gather opinions about how Twitter users are perceived by a number of random participants. To execute this, a range of fake users and tweets were created based on research on how users interact on Twitter and generating believable tweets, which were then created to look like a genuine screen shot from Twitter by using tweetgen². Research conducted behind the creation of these tweets was based on observing language and interactions between a range of users, from public figures to users who have few followers that post regularly, and verified users. We created a range of fake users with different views, characteristics and personalities, a mix of profile pictures or no profile pictures, randomly generated names and a variation of likes and retweets to make the accounts as genuine as possible. A bot account and a troll account were created to observe how participants perceived these users and if they are more sceptical of them, or even acknowledge that they might be a troll or bot.

The participants were shown 5 random tweets by each fake user then filled out a Likert scale, answering questions based on if they trusted the user or would retweet any of the tweets from that user, questions used can be seen in Figure 1. At the end of the questionnaire a free text option was provided for participants to discuss any thoughts they had while reading the Tweets and if there were any factors that impacted their responses. This free text option is designed to provide more information as to how the selected Twitter users were perceived and any specific features that may make them stand out. After gathering this information from the responses, we have gathered a wide range of information that can be analysed using various text processing techniques such as removing punctuation and stop words in order to finalise the term frequency.

3.2 Participants

The participants in this experiment were recruited by using a Snowball sampling procedure[7] and posting about the questionnaire on social media. A total of 75 participants filled out the questionnaire, including 44 responses to the first free text option and 55 responses to the second free text option. This experiment was not aimed at a particular demographic, we aimed for a mixture of Twitter and non-Twitter users to ensure we could gain a range of responses. A brief description of the experiment was given to the users before they were shown any tweets, giving them a basis of what to expect from the questionnaire ensuring they understood what they were consenting to. We requested the participants email address in case they wished to withdraw their responses from the experiment at any time. No other personal or demographic information was gathered to ensure participant anonymity. Participants were not informed that the users

² <https://www.tweetgen.com/create/tweet.html>

presented in this experiment were fake until the end of the questionnaire during the debrief, this deception was intended so that the participants started the experiment believing they were real users and provide genuine responses to the content. This experiment received the universities ethical approval, following the appropriate guidelines and ensuring no personal data was gathered from the participants unless absolutely necessary.

4 Preliminary results

4.1 Quantitative results

As part of the data collection, participants filled out a Likert scale answering 8 different questions after they had read through 5 tweets from each different user, 10 users in total. Some of the questions for this section include asking participants if they would retweet this user, if they trust the information discussed and if they think the users are a social media bot or troll, the participant then selects their responses between ‘Strongly disagree’ to ‘Strongly agree’ as shown in figure 1. All 75 participants responded to each of these questions.

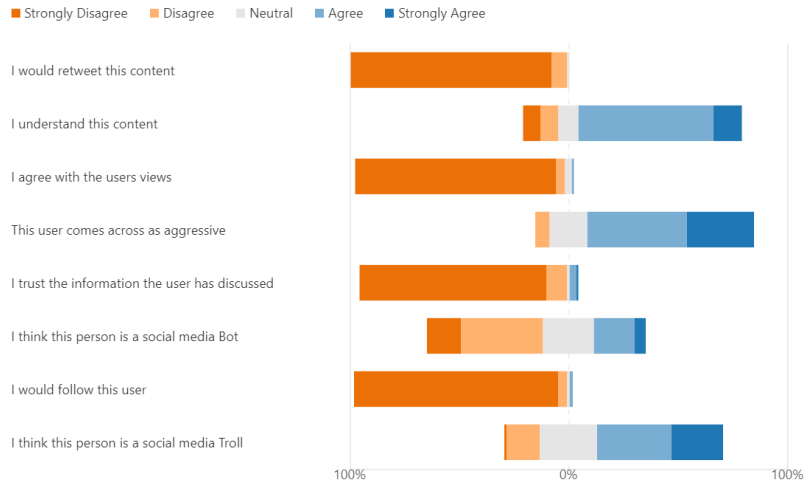


Fig. 1: Result from likert scale

From the results it was found that participants generally had the same responses to not liking, trusting or wanting to retweet tweets of the the Twitter accounts that expressed very strong views. For example the results shown in Figure 1 is from a user who expressed very strong controversial political views, showing that there was a majority response to each questions implying that most participants viewed them in the same way. The Twitter users that shared more

personal information had more of a neutral responses compared to users who were sharing information or expressing political views. The Twitter user that was created to resemble a bot was overall mostly viewed as a bot by the participants (65.3%), though not everyone agreed. There was more of a positive or neutral responses to the cat account (an account made to look like it is the cat tweeting rather than a human) and the users that appeared to be younger. Though there was nothing to specify the exact ages of any of these Twitter users, participants assumed that if they looked younger in their profile pictures then it was actually them and a few participants made comments comparing how they viewed the ‘younger’ users compared to the ‘older’ users.

Overall most users had a mixed results, but generally there was a majority response for each user if the participants found them to be trustworthy or not and if they would share any of their tweets. The cat account, example in figure 2a was most trustworthy (28% agree - strongly agree), the account that came across as very right wing, Sully in figure 2b came across as the least trustworthy (94.5% disagree - strongly disagree).

Fig. 2: Examples of tweets used in the experiment



4.2 Qualitative results

For the analysis of results gathered from the free text option of the questionnaire, we looked at term frequency for each response and looked for any common factors that participants mentioned which could have impacted their response. Out of the 75 participants that filled out the questionnaire, 44 responded to the free text option giving them the opportunity to discuss any further information they could share about the users that impacted their responses to the questions. The second free text option was to gain information from participants about if there were any factors such as profile picture or number of likes that affected their responses. The results of the term frequency analysis can be seen in tables 1a for the first responses and 1b for the factors.

5 Conclusion and future work

Overall the preliminary results of our study show us that there are specific factors that impact how a user is portrayed based on what and how they share posts on social media. Our research identified that a user's profile picture and the actual content of their tweets are the two most significant factors in determining whether a user will share a post on social media. The language, spelling and grammar will play a role in the user's decision making process. This information is key when understanding what makes a user more attractive, engaging and why other social media users would interact with them.

Based on the preliminary results from this experiment, we have gained an insight into some of the key factors that play a role in how a Twitter user might be perceived based on a small sample of what they might discuss on Twitter. After a more in-depth analysis is carried out on these results, it could lead to further experiments around deception; if participants have a certain expectation before viewing users tweets does this alter their perception of the person, leading into a discussion on the halo effect[8] in social media. An experiment like this could include creating a similar questionnaire to the one generated for this experiment, but populated only with bots, can we deceive participants into thinking bots are actual users and vice versa? This has the potential to lead into more work covering how bots are perceived on social media and if people can find a bot trustworthy. Future experiments similar to this also open up new paths to further text and sentiment analysis, looking at how language used can alter the sentiment of a social media post.

References

1. Wang, P., Angarita, R., and Renna, I. 2018. Is this the Era of Misinformation yet: Combining Social Bots and Fake News to Deceive the Masses. In Companion Proceedings of the The Web Conference 2018. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1557–1561.
2. Rosenberg, H., Syed, S., & Rezaie, S. (2020). The Twitter pandemic: The critical role of Twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. *CJEM*, 22(4), 418-421. doi:10.1017/cem.2020.361
3. Qiu, L., Lin, Q., Ramsay, J., Yang, F.: You are what you tweet: Personality expression and perception on Twitter. *Journal of Research in Personality* V 46, Issue 6 (2012)
4. Azucar, D., Marengo, D., Settanni, M.: Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences* Volume 124 (2018)
5. Sterrett, D., Malato, D., Benz, J., Kantor, L., Tompson, T., Rosenstiel, T., Sonderman, J. & Loker, K.: Who Shared It?: Deciding What News to Trust on Social Media. *Digital Journalism*, 7:6, 783-801 (2019)
6. Golbeck, J., Robles, C., and Turner, K.: Predicting personality with social media. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (2011)
7. Goodman, Leo A. "Snowball Sampling." *The Annals of Mathematical Statistics*, vol. 32, no. 1, 1961, pp. 148–170.
8. Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35(4), 250–25

Table 1: Term frequency for how many times a word was used in the free text responses

Thoughts	
Word	Word count
tweet	23
users	16
twitter	12
people	11
information	11
account	10
bot	8
user	7
political	7
follow	7
views	6
felt	6
content	6
none	5
social	5
media	5
across	5
agree	5
news	5
interesting	4
posting	4
politics	4
aggressive	4
things	4
come	4
opinions	4
5g	4
cat	4
though	4
different	4

(a) Thoughts

Factors	
Word	Word count
like	34
user	27
tweet	23
bot	22
profile	16
picture	15
would	14
language	14
didnt	14
think	13
content	12
account	11
used	11
blue	10
also	10
definitely	9
aggressive	9
twitter	9
retweets	8
really	8
grammar	7
number	7
look	7
likely	7
views	6
spelling	6
verified	6
dont	6
people	5
posts	5

(b) Factors