Comparative Study of Twitter Usage in Canada and UK During COVID-19's First Wave

Rebecca Di Francesco[a], **Christian Ellehammer Andersen**[b], **and Huayuan Song**[c]

[a]s221424; [b]s194255; [c]s183897

During the outbreak of the Coronavirus, lockdowns forced people to stay home, and as such, they spent more time socializing online (1). Twitter was, since the beginning, an important tool for expressing concerns and raising awareness (2)(3). After March 2020, hashtags related to Covid-19 became stable in the ranking of the top trending hashtags for the months to come. As of January 2022, 18.4 million Twitter users are from the UK and 7.9 million users from Canada (4). This study compares a sample of tweets related to Covid-19 hashtags (#corona, #wuhan, #nCov and #covid) in the period between 2. March 2020 to 22. July 2020 for the two anglophone countries. The scope of the project is to understand the usage of Twitter in light of the COVID-19 pandemic in a comparative study of these two countries. Our results demonstrate that in the UK, a country whose COVID response has been considered flawed and inadequate (5), the sentiment directed toward their Prime Minister is significantly more negative than the sentiment embedded in other tweets, whereas we found no difference between the sentiment of tweets directed towards Justin Trudeau and other Canadian tweets. In our framework, we gave much importance to the visualization of important words using TF-IDF statistics and we used this tool to represent the communities detected with the Louvain algorithm (6). We found a difference in the tones when comparing the word clouds through different stages of the initial COVID wave. Lastly, when comparing the general sentiment of the tweets, no significant difference in the average sentiment score between the two countries was observed.

network science | NLP | Twitter | COVID

In early 2020, after seeing the rapid spread of a novel coronavirus across borders and continents, countries across the globe went into hard lockdown to mitigate the contagion. The initial COVID response from the UK government led by Prime Minister Boris Johnson has been labeled as inadequate (5) and is a contrast to the early Canadian response, which has received praise for its strong performance and persistence, resulting in fewer deaths per capita (7). With reduced mobility and physical connections, people turned to their mobile devices in order to connect virtually (2). While Facebook is the dominant social medium for most of the world (8), Twitter has been characterised as a unique discussion hub with immense influence on media and politics (3)(1)(9). Its diverse set of users ranges from ordinary citizens to celebrities and official government accounts (10). This paper examines the usage of Twitter in Canada and the UK during the initial COVID wave in 2020 using methods from network science and natural language processing (NLP). We build social network graphs of users that mention each other in their tweets and identify communities using the Louvain algorithm. We carry out an explorative analysis with the goal of understanding the public sentiment and shedding light on the following open questions: What were the most important discussion themes in the tweets? Who were the most central users?

Results

Community Detection. We partitioned the Canada and UK networks in communities using the Louvain algorithm (6). The modularity found was high (0.91 and 0.90, respectively) for both. Since modularity scores (defined as Q(G, C)) are greater than 0.3, both networks present good community structures (11). This means that these networks present high strength of divisions into communities with dense connections within them and sparsely connected with nodes outside of them (See Fig. 5). As expected, the number of communities found is very large, given the networks also have numerous connections. By looking at the histograms of community sizes (A.7.3.0.1 & B.7.3.0.1 in the explainer notebook), it is also evident that the vast majority of communities are very small, with only a few having thousands of nodes. For visualization purposes and to gain an understanding of the communities, we displayed word clouds only for the 10 largest communities, which account for around 23% and 19% of total network nodes for Canada and the UK, respectively. For both networks, there are certain words such as 'help', 'please', 'need' that show concerns for the outlook of the pandemic. For UK, the word 'death' is present in all the 10 word clouds, while only in two for Canada. Clearly, UK tweets were more focused on death numbers given that during the peak in April, they reached over 1000 per day (12) compared to around 100 daily deaths (13) in Canada during their peak time at the beginning of May. Moreover, we found that for both countries, the largest community seems to be centered around Donald Trump. As we will see in the following paragraph, this makes sense given that he is the most influential user of the network.

Centrality Measures. We used centrality measures to investigate who are the most influential users of the networks. Interestingly, when we look at in-degree centrality for Canada, Donald Trump, figuring at the first place, is even more mentioned than their Prime Minister Trudeau, who follows right after. Then, in order, there is the official YouTube account, Ford Nation (an opinion news program hosted by Rob Ford)

Significance Statement

The purpose of this study is to better understand Twitter usage in Canada and the UK during the initial COVID-19 wave in 2020. Our research shows that the choice of words in COVID-19 tweets and the sentiment directed towards Prime ministers can be good indicators of a country's COVID severity in terms of mortality.

Contributions

The contributions of each author can be seen in the Table of contents at the beginning of the explainer notebook linked in the Materials and Methods section of this paper.

and CBC news (news account). Ford Nation and CBC News are news sources and YouTube was also used as a news source since it was observed that the tweets that mentioned this account were sharing YouTube videos about COVID news. These accounts are also the most central according to betweenness centrality. However, the order is different: YouTube goes to fifth place instead of third, while Ford Nation and CBC are before that account. Out-degree centrality is not meaningful in this context since it is related to people that mention others a lot. If we look at the UK network, the most central users according to in-degree centrality are in order: Boris Johnson (politician), Piers Morgan (broadcaster), Change.org UK (petition platform), Matt Hancock (politician) and then Donald Trump (politician). Interestingly, among the most mentioned users, there is a petition website Change.org and by printing some of the tweets that mentioned this account, it was seen that most of the linked petitions were for supporting the NHS, the National Health Service of the UK, which was put under heavy stress during the initial outbreak (14). The first five central users according to in-degree betweenness centrality are also very similar to the first in-degree centrality users, except that after the first three there is @realDonaldTrump and then @BBCNews. The assortativity coefficient being close to 0 for both networks mean that they are non-assortative, which makes sense since we do not expect that users mentioning a lot of other users also connect more with users that follow the same pattern. However, it may also be meaningful to expect that users that are mentioned a lot by others (famous accounts) are largely connected with users that are also mentioned a lot (also famous). However, it is outside the scope of this project to test for that hypothesis since not all the famous accounts are in our dataset of tweets and, consequently, we cannot find whom they mentioned in that period.

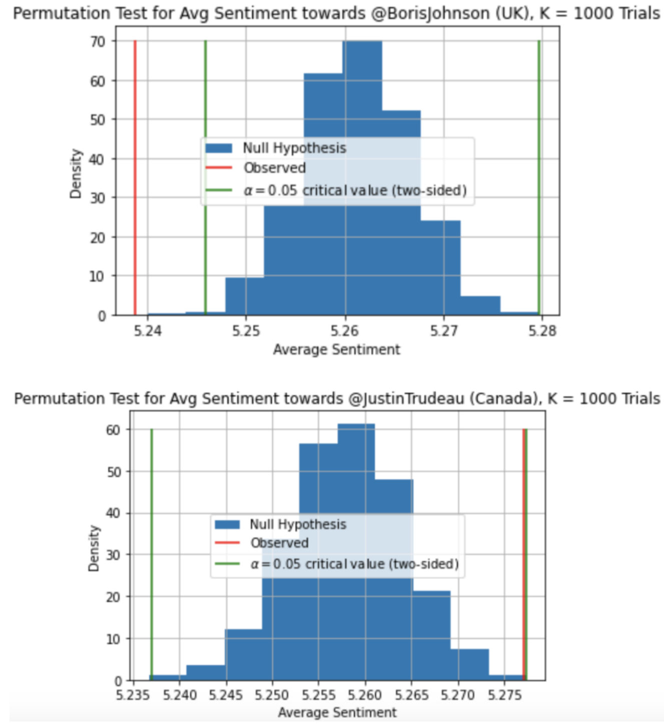

Fig. 1. Permutation tests for average sentiment of tweets directed towards @BorisJohnson and @JustinTrudeau visualized.

Time-dependent Text Analysis. We used the same framework of word clouds used for the community visualization to show the most important words during different stages of the initial COVID wave: initial phase, peak and final phase. In Figure 3 a fluctuating picture emerges. For Canada, at the initial phase, we see words like 'flu' that were minimizing the magnitude of Covid-19, but also the word 'outbreak' that clearly signal the start of the virus contagion and we see the word 'China' that is the origin of the virus. At the peak, words such as 'support' and 'spread' appear. Whereas, at the final stage, we see the word 'mask', which corresponds with the time numerous Western countries first started discussing surgical masks as a tool in mitigating the spread of virus (15). For UK, at the initial phase we also spot the words 'flu' and 'China', however, what is more noticeable is the word 'staff' and 'petition' as the health system in UK was strained (14). Then at the peak, we see the word 'death' appearing with 'help'. Even looking at the final stage word cloud, the tone of words still appears more negative when compared to the Canadian one. For UK, there are the words 'please', 'death', and the call for help, 'help', 'support', 'need', 'government', while Canada seems more neutral. We also produced some dispersion plots (see A.7.5.0.2 & B.7.5.0.2 in explainer notebook) with certain words that we thought could be used to find some interesting patterns. The plots do not look very meaningful, however, as we can see for the word '5g', there is definitely higher frequency for the UK network than for the Canada network and the dispersion decreases for the peak period. This word was checked since it was a central word of a viral conspiracy theory related to COVID-19 (16). This may support the fact that in the UK, this theory was more spread out and that it gained more popularity right when the UK was under higher pressure. It is, however, outside the scope of this project to test whether this hypothesis is true. We also produced plots that show a comparison of word frequency over time. We can notice from both plots that the word 'mask' shows an increasing behavior in frequency since it was a new measure adopted. We can also see that while for Canada, the peak of frequency of the word 'death' is aligned with the true peak of daily deaths, this is not true for the UK, in which the frequency of that word is a bit delayed with respect to the 'true' peak.

Sentiment Analyses. Figure 2 shows the sentiment score of the tweets for the two different countries. The plots include a 95% confidence interval. Based on the plots, it seems that there is a time dependency regarding the sentiment score. There is not a clear peak or a clear valley. Instead, the scores seem to be oscillating up and down. It is not known whether the oscillations are due to real-life events or something else entirely.

It is interesting to see that the sentiment scores for both countries start very low (sad sentiment) and then increase. Perhaps at the beginning of the pandemic, the countries were in a state of shock. This was unprecedented and people were probably really scared about the prospect of the future. Later on, maybe people get more used to the new situation, and maybe the news flow changes. That *could* explains why the sentiment score starts low and then rises, but it's just speculation.

It is also intriguing that for Canada, the sentiment seems

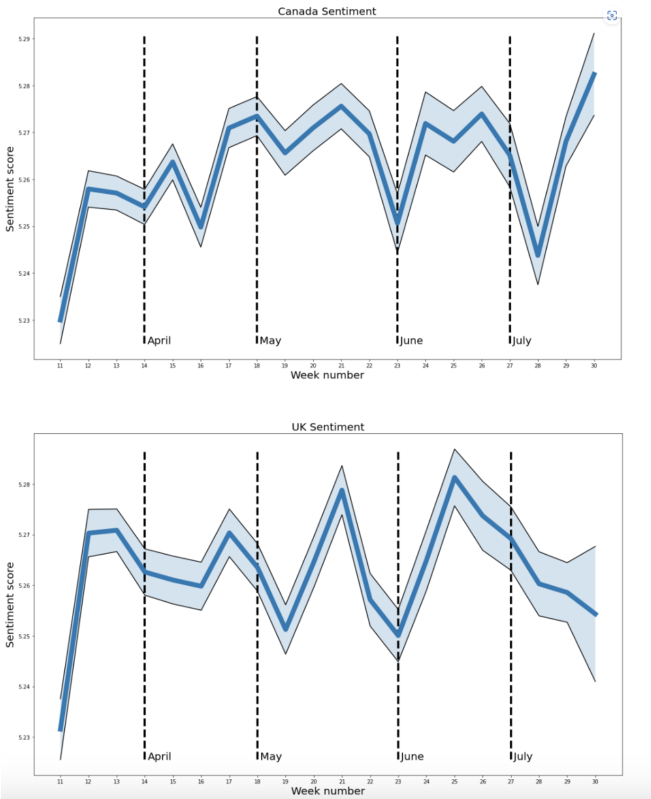

Fig. 2. Evolution of sentiment over time for both countries.

to be rising in the end, whereas for the UK, the opposite. *Maybe* this reflects the fact that the UK had higher mortality during the first COVID wave (5). Further work is required to determine this for certain, though.

Permutation Tests. To gain an understanding of the public sentiment towards the two countries' leaders, we performed a permutation test with label shuffling for the tweets directed towards Boris Johnson and Justin Trudeau in the UK and Canada datasets, respectively, with 1000 trials. It is seen in Figure 1 that the observed value for average sentiment directed towards @BorisJohnson is lower than the critical value corresponding to an $\alpha = 5$ % significance level. Accordingly, we reject the null hypothesis of the average sentiment towards @BorisJohnson being equal to the average sentiment of tweets in the UK dataset. For Canada, however, the observed value is within the critical value range and thus, the average sentiment towards @JustinTrudeau is not significantly different from the average sentiment of tweets in the Canada dataset.

Discussion

By building the social network graph and using in-degree and in-betweenness centrality measures, central users were identified as influential political figures, news media and organizations. For both the Canada and UK networks, there was at least one detected community where the main discussion theme was Donald Trump. For the found UK communities, the word 'death' appears more frequently in the top 10 TF-IDF word clouds compared to Canada. This observation corresponds with the UK having experienced higher mortality per capita due to COVID during the initial wave compared to Canada. However, since our analysis was conducted on a relatively limited amount of tweets (229550 for Canada and 269760 for the UK) in a limited time window from 02-03-2020 to 22-07-2020, more work is required in verifying this finding and applying it to the broader discourse on Twitter beyond the keywords that determined the scope of the raw data and beyond the time frame chosen by this study. Nonetheless, we reach pertinent conclusions with permutation tests indicating that while there is no significant difference in the average sentiment score of tweets between Canada and the UK, the sentiment directed towards the UK prime minister is more negative than the sentiment embedded in other tweets whereas no such difference between sentiment directed towards the Canadian prime minister and other Canadian tweets were found. Technical learnings from this project consist not only of a multitude of course methods within network science, such as community detection and NLP, including creating word clouds from the TF-IDF vector and computing sentiment scores, but also implementation of two-sided permutation tests on average sentiment towards famous users. The main technical challenge was, however, working with a large amount of data. The data acquisition proved challenging due to limitations built into the free version of the Twitter API, which reduced the number of API calls permitted per 15 min. time intervals. Hence, scraping a total of 499310 tweets took approximately 16 hours. Moreover, scraping user handles for those tweets took approximately an additional 4 hours combined.

Materials and Methods

For a detailed walkthrough of the methods, see the explainer notebook: https://github.com/HuayuanSong/2805_covid_tweets/blob/main/main.ipynb

The Dataset. The dataset used in our project comes from the paper "COVID-19 Twitter Dataset with Latent Topics, Sentiments and Emotions Attributes" (17). The full dataset contains tens of millions of rows. We took a subset of it by considering a time window that goes from the 2nd of March 2020 and the 22nd of July, and we considered only the data for UK and Canada. This dataset provides the tweet IDs, the user IDs, the precise date and time of the tweets, and some columns related to computed sentiment scores such as fear intensity, anger intensity, sadness intensity and happiness intensity. However, the original dataset did not provide the tweets' text; accordingly, this was the first computational challenge we had to face. We developed some functions based on the Tweepy API to find the tweet text from the tweet IDs and the usernames from the user IDs. Given the time limitations and the Tweepy functionalities, we had to further reduce the dataset by taking every third row of the dataset for the Canada dataset and every sixth row for the UK one as it was bigger. This was done in order to end up with a similar amount of tweets for both countries as we had 229550 and 269760 rows of data for Canada and the UK, respectively. We built the directed networks by connecting user A and user B if, in the tweet of user A, B is mentioned. Both networks, as expected from social network theory, presented a power-law distribution for both in-degree and out-degree distributions, with the majority of users that are low connected to others and very few that have large connections. The UK network has 256041 edges and 255126 nodes, whereas the Canada network has 143538 edges and 144510 nodes. Finally, we created a new column for the tokenized tweets, and we removed stop words as well as tokens representing URLs and usernames. For the tokenization, the RegExpTokenizer from the NLTK library (18) that automatically removes punctuation was used.

Community Detection Algorithm. To identify communities within the network, we grabbed the undirected version of the graph. The un-

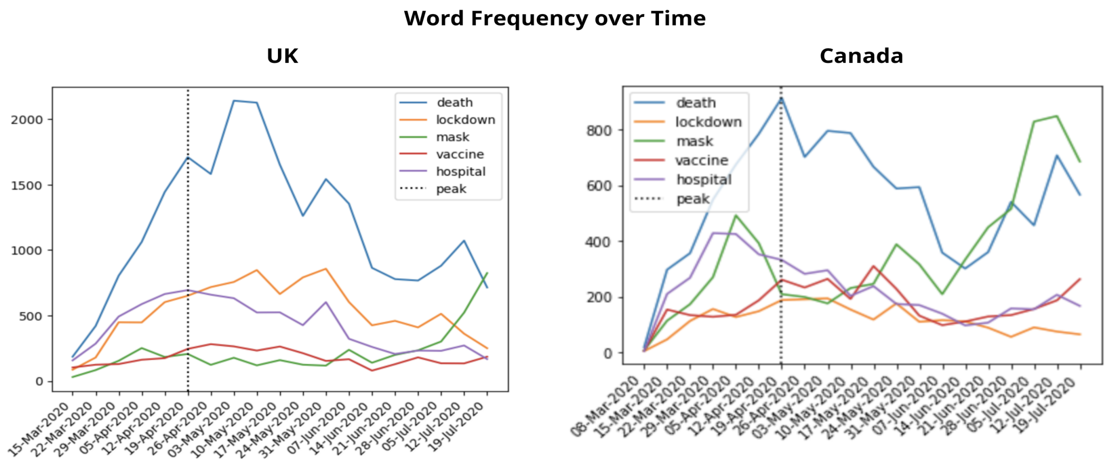

Fig. 3. It is seen that the word frequency of "death" does not peak at the same time as the peak in weekly average daily mortality rate (the dotted vertical line) for the UK. There seems to be a 'lag' in the frequency of the word 'death', whereas for Canada the two peak at the same time. Note: The peak of the weekly average daily mortality rate was found with official UK and Canadian data (12)(13)

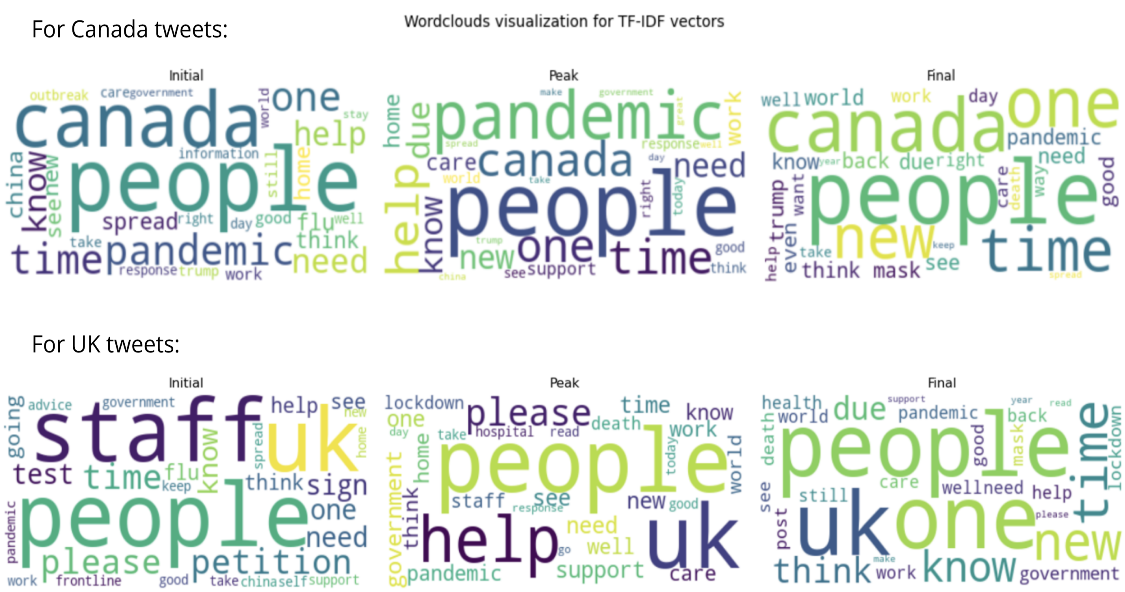

Fig. 4. Wordclouds of the top 30 TF-IDF words for the initial, peak and bottom phase of the period between 03-02-2020 and 07-22-2020 for UK and Canadian tweets. The following time spans were used; "Initial": 03-02-2020 to 03-22-2020; "Peak": 03-31-2020 to 04-13-2020; "Final": 07-02-2020 to 07-21-2020.

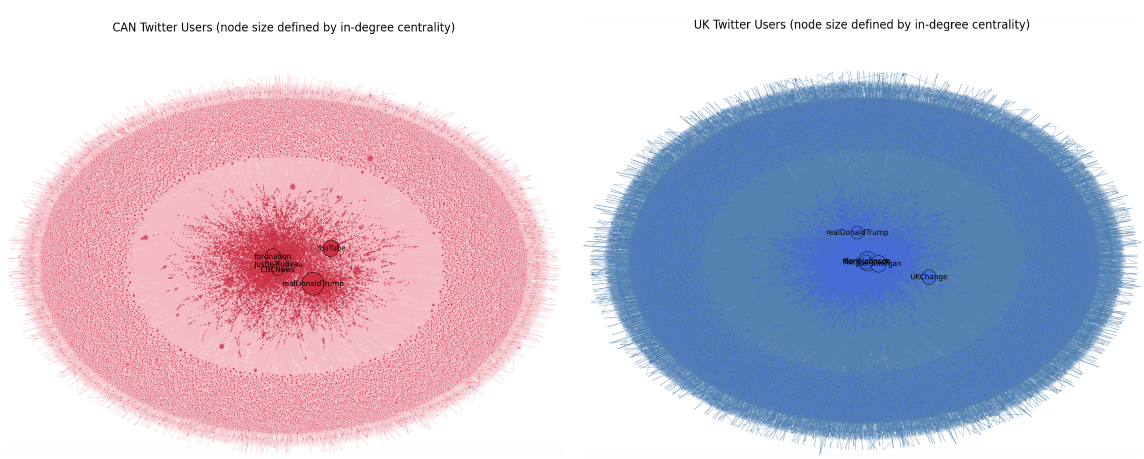

Fig. 5. The directed network of Canadian (left) and UK (right) COVID tweets visualized with the top 5 most central users according to in-degree centrality labeled. Users (nodes) are connected (links) when they mention each other in their tweets. Node size is determined by in-degree centrality.

supervised Louvain algorithm is used for detecting the communities. It is a greedy algorithm for extracting communities from large networks by Blondel et. al. (6). The algorithm seeks to obtain the maximum modularity for each iteration. We selected the Louvain algorithm to partition the network into communities instead of using the Girvan-Newman algorithm due to Louvain having a more efficient runtime of $O(n \cdot logn)$ compared to a runtime between $O(m^2 n)$ to $O(n^5)$ for Girvan-Newman (19). Given a graph G with $|E(G)| = m$ and clustering of communities $\mathcal{C} = (C_1, \ldots, C_k)$ of $V(G)$ into communities C_i. Modularity $Q(G, C)$ can be defined as:

$$Q(G, \mathcal{C}) = \frac{1}{2m} \cdot \sum_{C_i \in \mathcal{C}} \sum_{\substack{v,w \in V(C_i) \\ v \neq w}} \left[A_{vw} - \frac{d_G(v) \cdot d_G(w)}{2m} \right]$$

A_{vw} is the indicator function: $A_{vw} = \begin{cases} 1, & \text{if } vw \in E(C_i) \\ 0, & \text{if } vw \notin E(C_i) \end{cases}$ The sum is scaled by $\frac{1}{2m}$ to obtain range for $Q(G, \mathcal{C})$ within $[-1, 1]$ (20).

TF-IDF. TF-IDF is short for Term Frequency - Inverse Document Frequency. It is a statistic that describes the importance of a word for a document while accounting for the relation to the rest of the documents from the corpus. Practically, this is done by counting the frequency of a word in a document while paying attention to the frequency of that same word in other documents in the same corpus.

- **TF**: We calculate TF by

$$TF(t, d) = \log(1 + f(w, d))$$

- **IDF**: For IDF, we use:

$$IDF(w, D) = \log(\frac{N}{f(w, D)})$$

Finally, TF-IDF is computed by multiplying TF with IDF:

$$TFIDF = TF(t, d) \times IDF(w, D)$$

(21)

Betweenness Centrality for Undirected Graph. In betweenness centrality, the more times a node is in the shortest path between two other nodes, the more central it is (22). Label each edge $e = \{u, v\} \in E(G)$ with a score $b(e)$. Define:

$$b(e) = \sum_{\substack{x,y \in V(G) \\ x \neq y}} \frac{\#\text{ shortest } x - y \text{ paths that use e}}{\#\text{ all shortest } x - y \text{ paths}}$$

Due to limited time and compute, we only considered betweenness centrality for the greatest connected components of Canada and UK networks.

Sentiment Analyses. We made an analysis of the sentiment in the tweets. This is done using labMT 1.0 (23), which is a dataset containing an estimated sentiment score for all common words. The word which is supposed to be positive has a higher sentiment score than negative words. For example, a word like "laughter" has a sentiment score of 8.50, whereas "murder" has only 1.48. Calculating the sentiment score of an entire tweet is a simple matter of averaging the scores of all the words in it. This a nice and simple method, but the context is not taken into account. The sentence "I don't laugh" would maybe not be classified as negative because the context is not being considered. We are going to be looking at how the tweets' sentiment scores change over time.

References.

1. A Pandya, P Lodha, Social connectedness, excessive screen time during covid-19 and mental health: A review of current evidence. *Front. Hum. Dyn.* **3** (2021).
2. D Pandey, S Wairyaa, B Pradhan, Wangmo, Understanding covid-19 response by twitter users: A text analysis approach. *Heliyon* **8** (2022).
3. JH Kietzmann, K Hermkens, IP McCarthy, BS Silvestre, Social media? get serious! understanding the functional building blocks of social media. *Bus. Horizons* **54**, 241–251 (2011).
4. Statista, Leading countries based on number of twitter users as of january 2022 (https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/) (2022).
5. G Scally, B Jacobson, K Abbasi, The uk's public health response to covid-19. *BMJ* (2020).
6. VD Blondel, JL Guillaume, R Lambiotte1, E Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* (2008).
7. F Razak, S Shin, CD Naylor, AS Slutsky, Canada's response to the initial 2 years of the covid-19 pandemic: a comparison with peer countries. *cmaj* **194** (2022).
8. Statista, Most popular social networks worldwide as of january 2022, ranked by number of monthly active users (https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/) (2022).
9. J Nguyen, Politics and the twitter revolution: A brief literature review and implications for future research. *Soc. Netw.* **7** (2018).
10. A Jungherr, Twitter in politics: A comprehensive literature review. *SSRN Electron. J.* (2014).
11. S Fortunato, M Barthélemy, Resolution limit in community detection. *PNAS* **104**, 36–41 (2006).
12. UHS Agency, Deaths in united kingdom (https://coronavirus.data.gov.uk/details/deaths?areaType=overview%26areaName=United%20Kingdom#card-deaths_within_28_days_of_positive_test_by_date_of_death) (2022).
13. ea E. Mathieu, Coronavirus pandemic (covid-19) (https://ourworldindata.org/coronavirus) (2022).
14. H Alderwick, Is the nhs overwhelmed? *BMJ* (2022).
15. YSD Zhang, KA Noels, H Young-Leslie, NM Lou, "responsible" or "strange?" differences in face mask attitudes and use between chinese and non-east asian canadians during covid-19's first wave. *Front. Psychol.* (2022).
16. E Flaherty, T Sturm, E Farriesc, The conspiracy of covid-19 and 5g: Spatial analysis fallacies in the age of data democratization. *Elizabeth* (2022).
17. R Gupta, A Vishwanath, Y Yang, Covid-19 twitter dataset with latent topics, sentiments and emotions attributes. *OPENICPSR* (2022).
18. S Bird, E Klein, E Loper, *Natural language processing with Python: analyzing text with the natural language toolkit.* (" O'Reilly Media, Inc."), (2009).
19. V Kumar, A Sisodia, UM Pankaj, A Anand, Comparing algorithms of community structure in networks. *Indian J. Sci. Technol.* **9** (2016).
20. MEJ Newman, Modularity and community structure in networks. *PNAS* **103** (2008).
21. S Qaiser, R Ali, Text mining: Use of tf-idf to examine the relevance of words to documents. *nternational J. Comput. Appl.* **181** (2018).
22. LC Freeman, A set of measures of centrality based on betweenness. *Sociometry* **40** (1977).
23. P Sheridan Dodds, KD Harris, IM Kloumann, CA Bliss, CM Danforth, Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PLOS ONE* (2011).