Automated measurement of attitudes towards social distancing using social media: A COVID-19 case study

by A.S.M. Kayes, Md. Saiful Islam, Paul A. Watters, Alex Ng, and Humayun Kayesh

Abstract
The COVID-19 outbreak has focused attention on the use of social distancing as the primary defence against community infection. Forcing social animals to maintain physical distance has presented significant challenges for health authorities and law enforcement. Anecdotal media reports suggest widespread dissatisfaction with social distancing as a policy, yet there is little prior work aimed at measuring community acceptance of social distancing. In this paper, we propose a new approach to measuring attitudes towards social distancing by using social media and sentiment analysis. Over a four-month period, we found that 82.5 percent of tweets were in favour of social distancing. The results indicate a widespread acceptance of social distancing in a selected community. We examine options for estimating the optimal (minimal) social distance required at scale, and the implications for securing widespread community support and for appropriate crisis management during emergency health events.

Contents

Introduction
Coronavirus infections and consequences across Australia
Proposed model for measuring attitudes towards social distancing
Discussion
Conclusion

Introduction
The outbreak of COVID-19 has had a significant impact on the core dimensions of daily life for the global population (Australian Department of Health, 2020b). As an advanced economy in the Asia-Pacific, Australia has been no exception. Media reports suggest that the first COVID-19 case was confirmed on 25 January 2020 in Australia and confirmed cases appeared to be doubling every three days. However, formal advice from the state Department of Health and Human Services (DHHS) on social distancing (Glass, et al., 2006) was only announced on 12 March 2020 (Godfrey, 2020). Based on DHHS guidelines, a requirement to ensure a physical distance of 1.5 meters between people was mandated as the primary social distancing measure (Victoria Health and Human Services, 2020). According to a message from the Australian government and Chief Medical Officer (Victoria Government, 2020), achieving widespread compliance with social distancing measures over the course of the next six months would “flatten the curve” (Rhodes, et al., 2020), reducing the burden on over-stretched public health facilities.

However, there is a direct conflict between the public health requirement to “socially distance” and a natural human desire for social interaction (Cohen, 2004). Classic attachment theory suggests that humans (and other primates) have an innate and natural desire for social and physical contact, that begins in early life (Bowlby, 1979). Child primates...
denied physical contact with their mothers will go to extreme lengths to achieve that contact, even engaging in self-injurious behaviour. As Bowlby (1958) himself noted, “separation experiences are pathogenic”. Anecdotal reports from the current crisis describe a range of pathological and stress-based responses to enforced social isolation, from aggression in supermarkets over toilet tissue and spitting with the intent to infect, through breaking laws to breach social distancing limits. Furthermore, the mental health consequences of social isolation (Harlow, et al., 1965) cannot be ignored in the quest to suppress infectious diseases: there is significant co-morbidity of anxiety and depression in Australia, with elderly people being chronically affected (Almeida, et al., 2012) — ironically, social distancing is meant to protect elderly people who are also at a higher risk of contracting COVID-19 (Porcheddu, et al., 2020).

There are many ways that individuals can be affected by coronavirus. Based on a COVID-19 survey (Australian Department of Health, 2020a), people who have been admitted to hospitals for emergency treatments have been infected by coronavirus patients, with severe consequences. Though the primary goal of social distancing measures is to slow the curve against coronavirus infections, another important issue surrounds delays in COVID-19 testing. To achieve a greater good of social distancing, a non-infected population has been compelled to behave in a way which defies a natural desire to socialise (Stewart, et al., 2015). On the other hand, non-infected individuals benefit potentially from the overall reduction in the spread of the disease, since eradication or suppression will reduce their own chances of being infected. The selfish (and selfless) responses to infectious diseases have been examined previously in relation to HIV (Blechner, 1993) and Ebola (Falade and Coults, 2017), yet there has been no previous research specifically investigating attitudes towards social distancing. Yet the success of social distancing as an infection prevention strategy requires widespread social acceptance and adoption. How do we resolve this circularity?

According to the Australian Government Department of Health (2020c), COVID-19 is most likely spreading in three possible ways: (1) direct close contact with a confirmed case; (2) close contact with droplets of a carrier of the virus who coughs or sneezes; or (3) touching objects or surfaces contaminated with COVID-19 virus dispersed by a carrier of the virus, and then touching the mouth or rubbing the nose. Social distancing (Victoria Health and Human Services, 2020) is one of the crucial ways to reduce the COVID-19 pandemic, where different physical distancing restrictions are recommended such as no gatherings of more than two persons, stay at home and only go out if it is absolutely essential, keep 1.5 metres away from others, and avoid physical greetings such as handshaking, hugs and kisses.

When we speak to the main root causes, we refer to social distancing parameters as a preventive measure (Gillis, et al., 2000). However, the role of social media against spreading coronavirus is another important parameter to be considered. This is really another great challenge to identify the passive ways to protect people. There is a great need for using as much data from social media platforms such as Twitter about people’s attitudes towards COVID-19, including test delays and influencers’ hashtags related to government messages, and panic related to a rising curve of virus cases.

This research proposes a solution of ensuring an appropriate social distancing measure and protecting against spreading coronavirus infections. We provide some preliminary explanations for why people may be so willing to subvert critical controls in social distancing, intentionally or by a lack of awareness. In attempts to reduce illness through messaging, such as tobacco control, awareness and messaging have been effective (Chapman, 2007). Without awareness, the vast majority have no idea about preventive parameters that can contribute to measuring social distance, notwithstanding the role of the media in promoting government hegemony in relation to public health (Wallack and Dornman, 1992). In this paper, we use social media data about people’s sentiments and experts’ predictions, especially from Twitter (Layton, et al., 2010). We also use real COVID-19 datasets about infected people along with total cases by states and territories across Australia, for instance to know how frequently symptoms are trending in different states.

The first step in exploring behaviours is to explore attitudes that give rise to a given behaviour. One way to do this would be to use self-reporting measures such as polls or questionnaires. However, such measures have innate biases, where only those individuals with a particular interest are likely to respond. In order to address these biases, observational techniques to measure attitudes have been proposed. Previous research has suggested using social media sentiment analysis (SMSA) to measure attitudes at scale; Prichard, et al. (2015) developed an approach to measuring sentiments expressed in tweets and other types of social media to answer complex questions about public attitudes towards crime and criminals. The results — matching findings from attitudinal surveys (Prichard, et al., 2016) — suggest that a given sample of the population has quite sophisticated and complex attitudes towards seemingly simple questions about crime and criminals.

We propose to use SMSA to explore public attitudes towards social distancing measures, as the first step in developing a more sophisticated model of optimal social distancing. Twitter has a long history of being used for surveillance of
Automated measurement of attitudes towards social distancing using social media: A COVID-19 case study

infectious disease outbreaks (e.g., Signorini, et al. 2011). Can we use tweets and other social media data to better understand public sentiment towards social control measures that are intended to reduce infection? Can we predict what social attitudes to new and stricter measures may be, in the same way that tweets can be used to predict disease outbreaks (St. Louis and Zorlu, 2012)? A sophisticated model that can describe and predict public sentiment would be useful to lawmakers and public policy officials in the coming months and years.

The main research problem we explore in this paper is to identify, model and measure relevant parameters that we should consider to solve the social distancing problem, by understanding the characteristics of opinions and attitudes towards COVID-19. Consequently, we can maximise social distancing to the extent that is medically necessary to suppress viral transmission and for appropriate crisis management during emergency health events, while minimising psychological distress and harm caused by social distancing and isolation.

Coronavirus infections and consequences across Australia

Based on a report of 16 March 2020 (Godfrey, 2020), Australian COVID-19 cases appeared to be increasing daily, though restrictions on social distancing were already taking place as of 12 March 2020. The Australian prime minister announced a new social distancing measure on 29 March (New South Wales Health, 2020) along with stage-3 restrictions on gatherings, like no more than two persons allowed to gather physically unless they were household residents, even in outdoor spaces, like backyards.

From a purely descriptive perspective, we considered COVID-19 cases across Australia from 20 February to 27 March 2020. Figure 1 demonstrates the timeline of COVID-19 cases across Australia: (a) the time duration until when social distancing restrictions appeared on 12 March 2020 (Victoria Health and Human Services, 2020), and (b) another time duration declaring a clear guideline on non-essential closures on 27 March 2020 (Victoria Government, 2020).

![Figure 1: Australian COVID-19 cases from 20 February 2020 – 27 March 2020: (a) Number of cases before social distancing measures taken into account, (b) Number of cases after social distancing measures taken into account.](image)
Automated measurement of attitudes towards social distancing using social media: A COVID-19 case study

Figure 1 illustrates that Australian COVID-19 cases were on a similar trajectory to other locations around the world. Both timelines share a significant number of virus cases, both before and after social distancing restrictions were in place. The numbers reveal the dilemma of achieving social distancing given that humans are innately social creatures. We observed the number of confirmed cases increased from 159 on 12 March (date declared social distancing) to over 3,500 on 27 March, after 14 days of efforts to enforce “social distancing”.

In the literature, there are studies of practices to promote social distancing in K–12 schools (Uscher-Pines, et al., 2018) and advice on other measures in response to an influenza pandemic (Rashid, et al., 2015). These studies illustrate that effective social distancing requires serious restrictions within a critical period of 10 days before a peak of a pandemic. Other studies (e.g., Maharaj and Kleczkowski, 2012) have focused on modelling the purely economic outcomes associated with different social distancing strategies, completely ignoring a social impact (including compliance) driven by social attitudes. We argue that those studies are inconclusive and vital factors are missing.

Numerous recent studies utilised social media data in order to extract insights into the current COVID-19 pandemic outbreak. Aguilar-Gallegos, et al. (2020) analysed over 8.98 million Twitter posts from 21 January to 12 February 2020. The data collection was based on hashtags related to the coronavirus or COVID-19 mentions, categorising each post into (i) posts without mentions; (ii) posts with mentions; (iii) retweets; and (iv) replies. Rosenberg, et al. (2020) quantified and collected misinformation and disinformation during COVID-19 events. Wong and Jensen (2020) examined different Twitter activities on public trust and the development of trust in government in Singapore from January 2020 (when the first coronavirus case was reported in Singapore) to April 2020. In this paper, we model relationships between “social distancing against COVID-19” and “Twitter posts related to health concerns”, and measure attitudes towards health concepts like social isolation and working from home.

In the literature, different IoT and artificial intelligence mechanisms have been proposed to combat COVID-19. Gallotti, et al. (2020) and Cinelli, et al. (2020) assessed the risks of ‘infodemics’ in response to COVID-19. Gallotti, et al. (2020) analysed over 10 million Twitter messages and classified the reliability of news. Cinelli, et al. (2020) analysed massive COVID-19 data from different online platforms like Twitter, Instagram, YouTube, Reddit and Gab. The results found that the unreliable and misinformation (Pennycook, et al., 2020) from questionable sources encouraged irrational social behavior and created serious threats to public health. Gao, et al. (2020) analysed and assessed the prevalence of mental health problems and examined their association with social media exposure during COVID-19. The findings demonstrated that governments should pay attention to mental health issues like anxiety and depression as a result of an ‘infodemic’.

Shahsavari, et al. (2020) analysed COVID-19 news and notions of conspiracies behind the news, including rumors and misinformation. Singh, et al. (2020) identified information versus misinformation about COVID-19 based on Twitter datasets. The authors analysed how much of certain COVID-19 discussion topics were connected to information and misinformation. The findings revealed that misinformation was less dominant than COVID-19 crisis information. Yang, et al. (2020) analysed low-credibility information on Twitter during COVID-19. The findings revealed that social bots were associated with information of low credibility, spread via retweets. Different data mining and machine learning mechanisms have been proposed in order to measure emotions (Kleinberg, et al., 2020), to analyse coronavirus-related content (Li, et al., 2020) and to identify anti-Chinese behavior (Schild, et al., 2020) during the pandemic.

Proposed model for measuring attitudes towards social distancing

One of the biggest challenges in implementing a defense mechanism against a pandemic is to ensure public participation and acceptance of any mechanism proposed by authorities (Young, 2013). Therefore, understanding how people perceive a pandemic and their sentiments towards any mechanism is important in relation to successful implementation of a defense against a pandemic. Social media is a great source of public opinion, which has been explored to understand reactions and propagation of Ebola related information (Tran and Lee, 2016). In this paper, we propose a data driven automated measurement approach to measuring social attitudes towards defence mechanisms such as “social distancing” in the COVID-19 outbreak in Australia.
The proposed measurement model is a transfer learning based deep Twitter sentiment detection model. The model is built on a bidirectional encoder representations from transformers (BERT) (Devlin, et al., 2019) model. Central to our model is social media data, such as tweets. To test our proposed model, we collected a total of 100K tweets with the hashtag #coronavirus within Australia posted from 1 December 2019 to 30 March 2020. We use the Python module GetOldTweets3 [1] to download the tweets. We used ‘Australia’ as the location parameter in the method setNear() of the module to confine the target location of collected tweets. Out of these 100K tweets, 3,076 tweets contained the keyword “social distancing” and/or hashtag #socialdistancing. Figure 2 illustrates a wordcloud prepared from this dataset. We found that the most prominent word in the wordcloud was “social distancing”, which suggests that this was the most discussed topic in the sample.

To further understand the topics that were mentioned most in tweets during the sampled period, we extracted the top 15 uni-grams, bi-grams and tri-grams. The uni-grams, bi-grams and tri-grams correspond to one-word, two-word, and three-word phrases used in tweets. Figures 3 (a), (b) and (c) illustrates that the most used phrases: ‘coronavirus’, ‘social distancing’ and ‘practicing social distancing’. Additionally, ‘covid19’, ‘quarantine’ and ‘pandemic’ were among the top 15 single-word phrases during the period. Figure 3 (d) focuses on the top 15 hashtags used during the period [2]. The most used hashtags during this period appeared to be ‘#covid19’, mentioned in more than 25K tweets.
out of 100K tweets in our analysis. Frequently used phrases like ‘practice social distancing’ and common hashtags such as ‘#stayhome’ and ‘#stayhomestaysafe’ suggest were was support for the notion of social distancing.

Figure 3: The most frequent word phrases: (a) one-word, (b) two-word, and (c) three-word phrases, respectively; and (d) the most used hashtags.
In the preprocessing step, we tokenized each filtered tweet in wordpieces and converted them into vectors. We used the BERT tokenizer [3] to prepare tweets to finetune a pretrained BERT. We used the ‘bert-base-uncased’ version of the pretrained model, available in the ‘huggingface’ Python module [4]. Figure 4 illustrates a schematic diagram of our proposed model. We used a publicly available sentiment analysis training corpus [5] to train our model and then applied the trained model on the vectors of the filtered tweets to assess sentiments. The model labeled each tweet’s sentiment as either “Positive” or “Negative” towards “social distancing” as a defence mechanism for COVID19.

To analyse perceptions on “social distancing”, we fine-tuned and trained our sentiment detection model (as illustrated in Figure 4) on the collected tweets. We use 8K randomly chosen tweets for training and validation and another 2K randomly chosen tweets for testing. Our model achieved 83.70 percent accuracy and 81.62 percent f1-score on the test dataset. We then applied the trained model on the 3,076 tweets that mentioned “social distancing” as a hashtag or in simple text in a given tweet. We found that more than 80 percent of the tweets that talked about “social distancing” had a positive sentiment as illustrated in Figure 5. Social distancing as a measure to protect the community against COVID-19 appears to have been well accepted in Australia, in spite of social isolation.
Discussion

At first glance, COVID-19 crisis management and communication strategies have varied wildly between countries. We believe this paper might open up new avenues of future symptomatic coronavirus research for different communities, especially in predicting public attitudes towards future social controls that may become necessary. In future work, we propose to use different artificial intelligence techniques to understand the impact of social media on the COVID-19 panic cycle of individuals in a country, and to predict its trajectory. Furthermore, we will also investigate the role of messaging and chatbots in engaging concerned users on a very large scale — our current corpus of 100K represents a very small sample of all global discourse concerning COVID-19.

According to statistics from the Australian Department of Health (2020b, 2020c), there have been 3,166 confirmed cases in Australia, including 367 severe and mild cases of COVID-19 as on 27 March 2020. As the situation with COVID-19 continued to change at a rapid pace (Australian Department of Health, 2020b), we noted:

- Either social distancing was not adopted by all residents.
- Or, the parameters/metrics that were associated with such distancing were not adequate to combat the spread of COVID-19.
Automated measurement of attitudes towards social distancing using social media: A COVID-19 case study

Given that we have shown widespread support for social distancing, we believe the second observation carries more weight. We now present a two-layer social distancing model for effective COVID-19 crisis management to protect against the spread of coronavirus cases. In order to prepare an appropriate crisis communication plan, we can cluster two layers of social distancing parameters (“imposed” versus “derived” parameters).

- The parameters that are directly imposed by the government.
- The parameters that can be identified through data analytics on sentiments.

As demonstrated earlier, coronavirus cases are still growing rapidly though governments are trying to impose different restrictions on social gatherings. As such, in this study we propose an approach to measure attitudes towards social distancing so that we can separate government-imposed physical distancing restrictions from social gathering restrictions that can be identified from COVID-19 related social media data. In our future research, we will extend this model and propose a technical/mathematical model along with more insights from both datasets: (i) coronavirus symptomatic data from officials, and (ii) sentimental data from social platforms.

The coronavirus mobile app is an initiative by Australian government that currently offers up-to-date advice and information on COVID-19 (Mubin, et al., 2020). It provides eight distinct features for both iOS and Android platforms, including a ‘symptom checker’ along with a diagnostic tool from the Australian Department of Health, ‘register isolation’ for users in self-isolation, and other features. The coronavirus app could be extended through incorporating the social distancing model proposed in this paper — using a wide range of AI methods (including machine learning algorithms) so that it can provide relevant alert features, for instance.

The results from this research can be used to secure widespread community support and for crisis management during adverse health events like the COVID-19 outbreak. We can quantify relationships between pandemic characteristics (like social distancing measures) and Twitter activity, such as reactions and sentiments surrounding social distancing.

There is a concern over politicizing ideologies or agendas on social media using automatic accounts or bots (Zelenkauksaitė and Niezgoda, 2017; Ferrara, 2020a, 2020b). An automatic bot analysis could be used along with machine learning methodology to confirm that collected tweets were not dominated by auto-generated tweets. In our sample, we have not performed any bot analysis nor measured the effect of an automated influence induced by bots on the dataset.

In this work, we have applied an English language format and hashtags with most frequent word phrases such as top 15 uni-grams, bi-grams, tri-grams and hashtags other than #coronavirus. We have assessed the dataset relative to an acceptance of social distancing in a community. We have applied the BERT model for data preprocessing and filtering and eventually removed controversial hashtags from the dataset.

In the future, we will measure social distancing constructs with respect to different scale items such as posts without mentions, posts with mentions, retweets and replies over the course of different timelines (e.g., before and after social restrictions have been imposed). However, we believe that a more sophisticated approach should be developed to check the accuracy of data before any firm conclusion can be drawn.

Conclusion

A new social distancing model for COVID-19 has been introduced based on measuring the natural human desire for social and physical interactions by social media. We have identified, modeled and measured relevant parameters for “social distancing”: (i) the parameters that are directly imposed by governments, like a 1.5-meter physical distancing restriction; and (ii) other parameters that are identified through data analytics, such as from sentiments.

Although the proposed model is applicable to the management of the COVID-19 crisis, it has been formulated primarily for those infected cases whose occurrences are more likely because of the lack of applying appropriate social distancing measures. We considered not only the statistics about COVID-19 from official announcements of state and federal governments across Australia, but also examined social media data about “social distancing”.

We believe that our model provides new insights into the COVID-19 outbreak by considering public reactions and sentiments towards defense mechanisms. There are opportunities to collect COVID-19 pandemic datasets from
Automated measurement of attitudes towards social distancing using social media: A COVID-19 case study

different communities and social online platforms like Twitter (Chen, et al., 2020; Alshaabi, et al., 2020). We have provided a foundation for data driven automated analysis of public understandings and responses that could be exercised as a useful tool in future public health crises along with collecting coronavirus misinformation (Pennycook, et al., 2020) and unverified rumors.

About the authors

Dr. A.S.M. Kayes is a Lecturer in Cybersecurity at La Trobe University, Australia. He currently holds a research grant from the Oceania Cybersecurity Centre (OCSC). His research interests include information modelling, security and access control, cloud and fog computing, IoT security, ransomware detection and defence mechanism.
E-mail: a [dot] kayes [at] latrobe [dot] edu [dot] au

Dr. Md. Saiful Islam is a Lecturer in Big Data Analytics at Griffith University, Australia. His research interests include database usability, advanced data analytics, graph data management, machine learning and security analytics.
E-mail: saiful [dot] islam [at] griffith [dot] edu [dot] au

Paul A. Watters is Professor of Cybersecurity at La Trobe University, and leader of the Cybersecurity and Networking Research Group. Professor Watters currently holds an Australian Discovery Council grant and three grants from the Oceania Cybersecurity Centre (OCSC). His research interests lie in the human factors which underpin (or undermine) cybersecurity. He was previously a consultant to the Medical Research Council in the U.K., developing groundbreaking approaches to enable secure access to public health datasets.
E-mail: p [dot] watters [at] latrobe [dot] edu [dot] au

Dr. Alex Ng is a Lecturer in Cybersecurity at La Trobe University, Australia. His research interests span the fields of blockchain security, blockchain-based B2B collaboration, AI-based security defence for intelligent city, malware detection and prevention mechanisms.
E-mail: alex [dot] ng [at] latrobe [dot] edu [dot] au

Humayun Kayesh is a PhD student in the Griffith School of Information and Communication Technology. His research interests include natural language processing, data analytics and deep learning.
E-mail: humayun [dot] kayesh [at] griffithuni [dot] edu [dot] au

Notes

2. Here, we exclude the hashtag #coronavirus as we use this hashtag as the search key for collecting the tweets.

References


Automated measurement of attitudes towards social distancing using social media: A COVID-19 case study


Automated measurement of attitudes towards social distancing using social media: A COVID-19 case study


Automated measurement of attitudes towards social distancing using social media: A COVID-19 case study


doi: [https://doi.org/10.4324/9781315064475](https://doi.org/10.4324/9781315064475), accessed 20 October 2020.

doi: [http://dx.doi.org/10.5210/fm.v22i15.7795](http://dx.doi.org/10.5210/fm.v22i15.7795), accessed 19 October 2020.

---

**Editorial history**

Received 4 April 2020; revised 27 May 2020; revised 18 July 2020; revised 26 August 2020; accepted 14 October 2020.

Copyright © 2020, A.S.M. Kayes, Md. Saiful Islam, Paul A. Watters, Alex Ng, and Humayun Kayesh. All Rights Reserved.

doi: [https://dx.doi.org/10.5210/fm.v25i11.10599](https://dx.doi.org/10.5210/fm.v25i11.10599)