Towards Mitigating Misinformation in Times of Pandemic: An Exploratory Work in Progress

Ankit Aich and Natalie Parde
Department of Computer Science
University of Illinois at Chicago
Chicago, IL, USA
\{aaich2, parde\}@uic.edu

Abstract

The modern world is clearly vulnerable to pandemics. It is also rife with unchecked, open news sources, many of which spread much faster than traditional print media. In light of the new pandemic, COVID-19 or coronavirus, a considerable amount of unhelpful and/or harmful news and advice has been propagated through multiple channels. We provide details of ongoing research and upcoming work to study its impact and mitigate the proliferation of health-related misinformation in the future. Importantly, our work in progress includes the development of a comprehensive media dataset spanning both the current pandemic and other recent disease outbreaks, complete with measures of trustworthiness.

1 Introduction

Print media in the form of newspapers has existed in the United States since 1690. With the age of the internet, the demand for open source, digital news has overtaken that of traditional print media. In today’s world people turn to social media for accessible, bite-sized updates on every major event, with some companies even guaranteeing content that can be digested in less than sixty words.\(^1\) People flock to websites such as Facebook,\(^2\) Twitter,\(^3\) or other social media platforms like Instagram\(^4\) and WhatsApp\(^5\) to stay up to date with the latest world news, and as of April 2nd, 2020, there were 417,657,638 tweets alone on coronavirus.\(^6\)

Unscripted, interpersonal interactions can also serve as news sources. Recently the prime minister of one country urged people to show solidarity by clanging utensils, clapping from windows, and lighting lamps. What followed was people making loud processions on streets, bursting firecrackers, or dancing with loud instruments on the road—all at a time when social distancing was recommended time and again by the World Health Organization (WHO). Similarly, world leaders have called COVID-19 a rumor, only to fall sick and get admitted to intensive care units.

We propose to study this phenomenon at a large scale, across not only the current pandemic but also other global disease outbreaks, and describe our ongoing and future work towards doing so. Primarily we wish to examine the effects of health misinformation on everyday media consumers. Our first step towards doing so entails the development of a comprehensive dataset that will be useful not only for our needs, but for others researching health-related fact-checking and trustworthiness. Our downstream goal is to create a system capable of identifying high-impact, low-trustworthiness information sources and recommending topic-relevant alternatives with a higher trust quotient.

2 Related Work

Prior work in our area of interest has rarely been generalized across multiple pandemics, with most previous literature either directed towards a specific disease, or towards misinformation across all news articles. Ahmed et al. (2018) worked to examine moral panic and stress during the H1N1 and Ebola pandemics. They analyzed tweets to find discussion patterns associated with Ebola and H1N1 and identify themes stretching across multiple tweets. While the work performs a thematic analysis of panic, it also maintains that social media is a reliable source of information based on prior non-health work (Reyneke et al., 2011). The thematic analysis is interesting and can be used to identify
panic themes in Twitter. In the case of COVID-19, this may include misguided behaviors (e.g., inadvertently encouraging the hoarding of household goods) or, much more dangerously, increases in hate crimes against specific racial groups (e.g., “Chinese virus” messaging).

Chew and Eysenbach (2010) explored misinformation search terms associated with H1N1, utilizing terms such as “Bacon Flu,” as well as emojis, to categorize tweets. Their tweets also included links to news articles. Although the work did not contribute a dataset, it offers useful guidance in terms of the categorization of tweets and linkage to various news articles.

Goodall et al. (2012) examined printed media in the first five months of the H1N1 swine flu, using the Centers for Disease Control and Prevention (CDC) as a metric to find parity between officially-sanctioned and everyday media messaging. However, the work randomly selects 200 articles and then analyzes a percentage of cases mentioned by them. We hypothesize that these random selections may not yield optimal results, since although this ensures a broad spectrum, it neglects to emphasize the time when people are most affected (the disease’s peak, rather than later when its spread has stopped and vaccines have been made available). We believe that focusing on peak times for print and digital media selections will better suit our needs, and thus make use of the CDC’s pandemic and epidemic timelines when building our dataset.

Previous work has also studied misinformation at a non-health level (Almaliki, 2019). Vlachos and Riedel (2014) create a publicly available fact-checking dataset, utilizing a five-point scale (True, Mostly True, Half True, Mostly False, and False). They use news articles fact-checked by journalists, and assign labels using a K-Nearest Neighbors approach. However, as noted by the authors, the approach does not generalize well to articles that have not been fact-checked. Still, the work provides a good example of how news articles need to be quantified on the basis of trust.

In recent days automated fact-checking approaches have generated substantial interest, with some works using carefully-created knowledge networks to better classify and understand misinformation (Ciampaglia et al., 2015). However, the truth of the situation remains that misinformation reaches people and impacts people faster than does actual news. For example, videos uploaded by the CDC on their YouTube channel accrue fifty-five thousand views per day (based on view count on the YouTube channel as of April 7th, 2020), whereas claims and tweets from well-known politicians reach many more people in the same amount of time.

In the upcoming sections we describe our work in progress, as well as planned future steps. We anticipate that the work will lead to numerous contributions to the research community, including a comprehensive pandemic-related misinformation dataset, analyses of the relationships between information volume, misinformation, and self-reported mental health, and techniques for improving automated fact-checking methods.

3 Work in Progress and Future Plans

The coronavirus outbreak prompted lockdowns and enforced social distancing in most countries across the globe, and has become a constant news focus. However, there are many layers between the origin and impact of news. In the age of online social media, everyone has a voice, but this is not necessarily a good thing. News sources have shown many world leaders referring to the virus as the “Chinese virus,” which has resulted in an extreme escalation of racist attacks—L1ght (2020) reported an increase in hate speech against Chinese people and China by 900%, a 200% increase in hate speech against people of Asian ethnicity in general, and a 70% increase of hate in online chats. Hate speech increases the levels of stress and anxiety in people against whom it is directed, as well as others both within and outside their social circles.

More generally, in times of pandemic individual comments have the potential to serve as catalysts for mental stress and anxiety among millions of people. In fact, 36% of Americans feel that their mental health has been seriously impacted by the coronavirus outbreak (Porterfield, 2020). A simple hashtag search on Twitter shows a large number of tweets combining the hashtags #mentalhealth and #coronavirus.

Our work seeks to tackle this issue using a three-stage process. The first stage, currently in progress, involves creating a comprehensive dataset across multiple disease outbreaks. The second part will leverage this dataset to assess increases in mental health issues and panic during these outbreaks, with the hypothesis being that individuals exposed to a larger ratio of untrustworthy news will also
express higher levels of stress. For the third and final step, we plan to create a system capable of checking impact and trustworthiness for new trending articles, and suggest more suitable alternatives to those identified as being untrustworthy. It is our hope that the resulting system will reduce global stress levels in future disease outbreaks.

3.1 Dataset

Misinformation in today’s open world internet is huge. Consider the two following screenshots taken from the internet in Figures 1 and 2. Clearly, neither is scientifically valid. However, both were rapidly shared to huge networks of people using the popular messaging platform, WhatsApp, shortly after a leader issued an appeal to citizens in the country to switch off all electric lights and instead light candles and earthen lamps to show solidarity with essential services workers.

What followed was utter chaos, with people marching in the streets, bursting fire crackers, and dancing around. In short, people risked the health of a large portion of the community based on one piece of miscommunicated information. WhatsApp recently announced strict new regulations to curb the forwarding of harmful messages after being termed as a “petri dish of coronavirus misinformation” (Rangarajan, 2020). However, the app is just one of many channels through which harmful misinformation may travel. Our goal in creating a health misinformation dataset spanning multiple disease outbreaks is to facilitate the creation of novel methodologies to automatically curtail the spread of untrustworthy news in the future.


We selected these time ranges using timelines provided by the CDC indicating the lifespan of an epidemic, from spread to eradication. From within these timelines, we identified the specific ranges when the following conditions were satisfied:  

- The spread or mortality rate was at its highest.
- The disease was at a peak or near eradication.

We identified both print and open source media sources (including, but not limited to, social media platforms Twitter and Facebook) for these diseases. Open source media is only available for recent outbreaks.

Our data collection process is still underway. Once complete, we plan to release the dataset publicly to stimulate additional work by others. In the meantime, we will make current snapshots available to interested parties upon request to facilitate more rapid adoption when needed.
3.2 Misinformation and Mental Health

As previously stated, the current pandemic has led to a spike in mental health cases across the United States and the world. Intuitively, this makes sense; in addition to posing a major public health concern, the virus has wreaked havoc upon individuals’ social lives, as well as the economy, at a massive scale. Recent reports indicate that as many as 47 million Americans could lose their jobs due to the pandemic, with unemployment skyrocketing to 32% (Cox, 2020). Upticks in suicide rates and domestic violence have also been reported, as well as rampant uncertainties about many facets of daily life often taken for granted (e.g., education availability) (Pell and Lesser, 2020). We support these findings with the transcript from our own one-on-one interview with a psychologist in Table 2.

Clearly, information content does not increase linearly with volume; however, to the best of our knowledge the relationship between these variables during times of pandemic has not been formally studied. We hypothesize that with an increase in news sources and options, reliable information content will plateau after a certain point. Our belief is that this can be plotted out as a function of number of options vs. number of trustworthy sources during the COVID-19 and other outbreaks. We will also analyze the relationship between self-reported mental health and exposure to misinformation based on relevant hashtags and information sharing patterns among social media users within our dataset.

3.3 Downstream Application

Ultimately, we plan to develop new models to predict content trustworthiness and recommend suitable, topic-relevant alternatives to identified misinformation. We will do so by leveraging our novel dataset and building upon empirical findings from earlier stages of our work. We will experiment with a wide range of modeling techniques, and anticipate that approaches harnessing stylistic cues (Rashkin et al., 2017; Potthast et al., 2018) and discourse-level structure (Karimi and Tang, 2019) may be particularly valuable.

4 Conclusion

In this paper, we discussed preliminary work towards three key scientific contributions: (a) the creation of a new, publicly-available health misinformation dataset spanning multiple disease outbreaks, (b) empirical analyses of the relationship between information volume and trustworthy content, as well as the interactions of those factors with self-reported mental health, and (c) downstream methods for identifying untrustworthy content and suggesting trustworthy, topic-relevant alternatives. It is our hope that this work will mitigate stress and panic in future disease outbreaks, and spur additional interest in this topic by others in the research community.
References


Cynthia Chew and Gunther Eysenbach. 2010. Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak.


Jeff Cox. 2020. Coronavirus job losses could total 47 million, unemployment rate may hit 32%, fed estimates. CNBC.


L1ght. 2020. Rising levels of hate speech & online toxicity during this time of crisis. Technical report.

M. B. Pell and Benjamin Lesser. 2020. Researchers warn the covid-19 lockdown will take its own toll on health. Reuters Investigates.