

Using Twitter for developing insights into the 2009 Swine Flu and 2014 Ebola outbreaks

Wasim Ahmed¹, Peter A. Bath¹, Laura Sbaffi¹,
Gianluca Demartini²

¹ Information School, University of Sheffield ² School of Information Technology and Electrical Engineering, University of Queensland

Abstract.

Infectious disease outbreaks are a global public health risk that have the potential to take many lives in a short amount of time. It is important to understand the views and thought processes of the general public to have a better understanding of their perceptions of infectious diseases and how they spread. Social media platforms, originally intended for personal use, have recently been used in academic research for analysing public views and opinions as well as for disease mapping and tracking. Twitter, a widely-used microblogging platform, provides a unique opportunity to study the instant reactions of the public during disease outbreaks. This is because news of such epidemics on Twitter typically generate bursts of tweets. This abstract describes a study that is investigating user views during the peak of the 2009 Swine Flu and the 2014 Ebola outbreaks. Based on Google Trends data, tweets were retrieved from Twitter during a peak in Web search queries. Data were retrieved from a two-day period corresponding to the 2009 Swine Flu and 2014 Ebola outbreaks. A total of 214,784 tweets were retrieved from the two-day period of April 28th and April 29th 2009 for Swine Flu and 181,110 tweets were retrieved from 29th and 30th September 2014 for Ebola. The study then utilised thematic analysis in order to uncover potential similarities and differences between the cases. The results of this study will allow for the creation of guidance that can be disseminated by health authorities during an outbreak.

Keywords: Twitter, Swine Flu, Ebola, Social Media

1 Introduction

Infectious diseases outbreaks are a severe public health concern: the World Health Organisation and the World Bank noted that they account for at least 29 out of 96 causes of major human mortality [1, 2]. The 2009 pandemic of Swine Flu infected between 43 and 89 million people worldwide leading to between 8,870 and 18,300 lives [3] (CDC, 2010). During the occurrence of the Swine Flu outbreak people drew comparisons to the Spanish influenza [4, 5]. The Spanish influenza epidemic (caused by the A/H1N1 virus) was a catastrophic outbreak occurring between 1918 and 1920: it infected an estimated 500 million people and led to almost 100 million deaths worldwide, representing 5% of the world's population at the time [6]. Going back further, the Black Death, which took place from 1346 to 1353 is said to have claimed between one-fourth and three-fourths of the world population at that time [7]. Those past pandemics occurred without the modern communication devices that are available now and Twitter is one of the channels that enable people affected by such outbreaks to communicate and share information. It also offers researchers the unique ability to gather insights into how people respond during disease outbreaks. Until 2017, Twitter allowed its users to send 140-character messages known as 'tweets' and these can contain thoughts, feelings, activities, and opinions [7]: in October 2017 this limit was increased to 280 characters. Tweets are a useful source of data that can be analysed to provide insights into public views and opinions.

1.1 Research Aims and Objectives

This study is based on the retrieval of Twitter data on the Swine Flu outbreak from 2009, and the Ebola virus from 2014. The aim of the study was to examine the types of content that are shared during the peak of the two outbreaks, and then to compare them in order to identify similarities and differences between them. The rationale for the study was that both of these two cases were very significant, high-profile health scares, as both of the outbreaks had the highest media coverage in the 21st century so far [9]. By comparing the results of the two cases, important similarities may emerge which will aid the development of guidance on what information to disseminate to Twitter users. In the online world, Google had ranked Swine Flu as a term which was among the fastest rising Web Search query to appear in Google News [10], and Ebola was a term that was most searched for in 2014 [11].

2 Methodology

The dataset on Swine Flu contained 214,784 tweets which were retrieved corresponding to the two-day period of April 28th and April 29th 2009. The keywords utilised were: 'Swine Flu', 'SwineFlu', and 'H1N1'. This specific time period was selected for data retrieval because there was a heightened interest in the outbreak and an increase in Google Web Search queries around that time. The same approach was followed when selecting data for the Ebola outbreak. The dataset obtained for this consisted of 181,110 tweets which were sent and received from 29-30th September 2014, which corresponded to the period of heightened interest in the outbreak and an increase in Google Web Search queries around that time. The data were retrieved using the keyword 'Ebola'.

The rationale for using Google Trends to identify peaks was because it is not possible to search Twitter for peaks in tweets. The two datasets which were retrieved, as described above, were filtered (duplicates and near duplicates were removed at a 60% threshold) and a 10% sample of these tweets were entered into NVivo in order to carry out qualitative textual analysis method known as thematic analysis. A total of 7,678 tweets were analysed for Swine Flu, and 5,695 were analysed for Ebola. Filtering was important as previous research undertaken by the same project analyzing tweets on Ebola [12] found that the majority of content revolved around sharing news and/or updates, and it was important to cluster and remove duplicates in order to uncover personal opinions shared by Twitter users. The study adopted a pragmatic approach, employing a case study approach and will apply the Health Belief Model and concepts from Information Theory to the results.

3 Initial Results

Preliminary results have uncovered a number of themes and sub-themes for the respective outbreaks. Table 1 below highlights the themes that were similar to the 2009 Swine Flu outbreak and the 2014 Ebola epidemic. Table 2 provides an overview of the differences between the cases. These themes were identified by using the in-depth methodology of thematic analysis [13].

Table 1. Similarities in themes that emerged from the in-depth thematic analysis

Themes to Emerge from Swine Flu and Ebola
Fear, Anger, Fear of travel, Transmission, Prevalence monitoring, Speculative diagnosis, Prevention, Symptoms, Medications e.g. vaccines, Economic impact of disease, Information seeking, Voice of reason or downplaying the outbreak, General discussions, References to official organisations, References to Obama, References to areas affected, Humour, Sarcasm, References to popular culture

Table 2. Differences in themes that emerged from Swine Flu and Ebola

Themes to Emerge from Swine Flu only	Themes to Emerge for Ebola only
Worry, Prevention Products, References to other infection or disease, Unfollowing users, Frightening Scenarios, Images used in tweets, Name Discussion, Humor related to pigs	Dead rising generated fear, Praying, Prayer or call to God, Quarantine, Link to Instagram, Links to YouTube, Links to other tweets, Conspiracy Theories

There were similarities in both cases in how Twitter users responded, which could indicate that there is a feature of infectious diseases that may evoke a similar response due to the nature of the diseases. A number of differences emerged which were specific to the outbreaks such as Twitter users referring to pigs and pork produce for the Swine Flu outbreak. In contrast, discussion around Ebola centered on discussions of popular news stories at that time.

4 Conclusion

This study has provided an initial overview of PhD research in progress and which seeks to examine the information that was shared on Twitter during the 2014 Ebola epidemic and the 2009 Swine Flu pandemic. The results of this study may help to inform policy makers, media, and health organisations on the types of information that is required by the public during infectious disease outbreaks. The results of this study, once completed, will be of interest to the World Health Organisation (WHO), the United Nations (UN), the National Health Service (NHS), and the Department for Environment, Food & Rural Affairs (DEFRA).

References

- [1] Taylor, L. H., Latham, S. M., & Mark, E. J. (2001). Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 356(1411), 983-989.
- [2] Murray, C. J. L. & Lopez, A. D. (1996) *The global burden of disease: a comprehensive assessment of mortality and disability from diseases*. Geneva, Switzerland: World Health Organization.
- [3] CDC (2010). Updated CDC estimates of 2009 H1N1 influenza cases, hospitalizations and deaths in the United States, April 2009 – April 10, 2010. *CDC*. Retrieved from http://www.cdc.gov/h1n1flu/estimates_2009_h1n1.htm.
- [4] Nicolson, J (2009). The war was over - but Spanish Flu would kill millions more. [ONLINE] Available at: <http://www.telegraph.co.uk/health/6542203/The-war-was-over-but-Spanish-Flu-would-kill-millions-more.html>. [Last Accessed 11/11/2013].
- [5] Honigsbaum, M (2009). Swine flu and the lessons of 1918. [ONLINE] Available at: <http://www.theguardian.com/global/2009/apr/30/swine-flu-historical-lessons>. [Last Accessed 11/11/2013].
- [6] Taubenberger JK, Morens DM (2006). 1918 influenza: the mother of all pandemics. *Emerging Infectious Diseases*. [ONLINE] Available at: http://wwwnc.cdc.gov/eid/article/12/1/05-0979_article.htm#suggestedcitation [Last Accessed 11/11/2013].
- [7] Haensch, S., Bianucci, R., Signoli, M., Rajerison, M., Schultz, M., Kacki, S., & Carniel, E. (2010). Distinct clones of *Yersinia pestis* caused the black death. *PLoS pathogens*, 6(10), e1001134.
- [8] Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one*, 5(11), e14118.
- [9] McCandless, D. (2009). *Information is beautiful* (pp. 1-255). London: Collins.
- [10] Google Zeitgeist, (2009). Google. (Online) Available at: <https://archive.google.com/intl/en/press/zeitgeist2009/> [Last Accessed 09/07/17].
- [11] Google. (2017). Google Trends. [Online]. Available at: <https://trends.google.co.uk/trends/> [Last Accessed 14/05/17]
- [12] Ahmed, W., Demartini, G., and Bath, PA. (2017) Topics Discussed on Twitter at the Beginning of the 2014 Ebola Epidemic in United States. *iConference 2017*. Wuhan, China, March 2017.
- [13] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.