

The Chilling Effect and Internet Surveillance

Yuka Hatori, Santa Clara University

Advisor: Glencora Borradaile, Oregon State University

Abstract

While technological advancements have facilitated the ability of the people of the United States (U.S.) to exercise their First Amendment rights and to disseminate their opinions to a global audience, those same tools can be used to prevent them from fully exercising these rights. In countries with fewer formal civil liberties, these technological advancements may lead to overt censorship by the state. Despite there being protection from censorship in the U.S., legal scholars have argued that speech can be limited through a phenomenon known as the chilling effect. We aim to suggest research methods capable of finding evidence of chilling effects, as well as to design our own study.

1 Introduction

The U.S. Constitution’s First Amendment guarantees freedoms concerning religion, expression, assembly, and the right to petition [1]. It also guarantees the right of citizens to assemble peaceably and to petition their government. Despite these guaranteed freedoms, history has shown that governmental action have infringed on these rights. In a legal context, a chilling effect is “the concept of deterring free speech and association rights protected by the First Amendment as a result of government laws or actions that appear to target expression” [2].

There are multiple contexts in which people can experience the chilling effect and according to Penney [3] there are five main types of chilling: (1) legal/regulatory/statutory chill, (2) personalized law/threats/enforcement chill, (3) surveillance/data collection chill, (4) social media chill, and (5) disinformation/online abuse. The various types of chilling have been broken down into slightly different categories in order to provide more context. The redefined categories include the following: (1) legal chill, (2) personalized law chill, (3) surveillance chill, (4) social media chill, and (5) social chilling.

A legal chill can occur due to moments of situational uncertainty about the law’s requirements, in which people may be unsure about whether their conducts are legal or not and thus cause them to avoid that action. Personalized law chilling can occur when there is a specific law that will effect an individual based on their actions. For example, if an individual were to post a video on a social media platform that has copyrighted material, they would be subject to copyright infringement laws. As such, personalized law chill can be thought of as a subset of legal chill. Surveillance chill can occur when those who are uncertain about what to say or do in a given context are more affected by surveillance and therefore more likely to conform. Social media chill occurs when people are unable to determine what the socially accepted values are and default to self-censorship and disengagement while communicating on social media platforms or other communication channels. Finally, social chilling occurs when people are unable to understand socially accepted values while interacting with others, and thus resort to conforming to group majority values and actions [3].

Even though people have the right to express themselves within certain limitations [4], whether on social media or in public, government entities or large commercial entities have the ability to discourage the exercise of constitutional rights to their fullest extent [5, 6]. In 2013, Edward Snowden released highly classified information about how the National Security Agency (NSA) monitored users online [7], methods including tracking those who looked up certain keywords such as “Department of Homeland Security” and “National Guard” [23]. In more recent news, police departments monitor Twitter and other social media platforms to surveil Black Lives Matter Protest activity using tools like Dataminr, a service that uses artificial intelligence (AI) in order to provide real-time alert data [8, 9]. During the 2015 Freddie Gray protests, Baltimore police used Twitter, a social media platform, and Geofeedia, a social media location tracker tool,

to gain knowledge of a local high school walkout and subsequently prevented the protest from taking place [10].

As it is difficult to show evidence on a large-scale of people purposely avoiding behavior and not performing an action at all, prior work has focused on specific instances of when chilling effects can be detected. Furthermore, the interdisciplinary nature of this topic as well as the complexity associated with determining what events could cause a chilling effect makes it a difficult topic to study. This research aims to investigate the following topics:

RQ. 1 What are the research methods and topics that could be explored to find evidence of different types of chilling effects?

RQ. 2 What potential contributions can computer scientists make related to chilling effects?

2 Related Work

There are several U.S. court cases dealing with First Amendment rights that have helped to set a legal precedent regarding the chilling effect, including *Doe v. Reed* 561 U.S. 186 (2010), *Miami Herald Publishing Co. v. Tornillo* 418 U.S. 241 (1974), and *Buckley v. Valeo* 424 U.S. 1 (1976) [11].

Doe v. Reed was a case in which the Public Records Act (PRA), which made referendum petitions available publicly, was appealed because the plaintiffs argued that the PRA violated the First Amendment. However, the U.S. Court of Appeals held that the disclosure does not violate the First Amendment, but there could be a chilling effect on people who sign petitions because disclosure could expose them to harm [13]. The *Miami Herald Publishing Co. v. Tornillo* case was in regards to political candidates having the power to have their responses to criticism from any newspaper to be published. The Supreme Court of Florida’s ruling overturned a Florida state law that would have chilled editorial speech [14]. Finally, in *Buckley v. Valeo* in which the limits placed on electoral expenditures were questioned, the court held that precision of language regarding election or defeat of candidates is required to avoid chilling speech involving public discussion of political issues [15].

Canes-Wrone et al. [12] did a study on abortion law. They tested whether the enactment of a state law restricting late-term abortions affected the early-term, mid-term, late-term, and overall abortion rate in that state. They found that “late-term abortion restrictions reduce not only late-term abortions, but also ‘near-late-term’ abortions, i.e., abortions in the roughly one-month period be-

fore the period in which abortions are forbidden.” [12] Meaning, the number of abortions performed at the near-late-term period likely decreased in fear that they would be too close to the deadline, so some may have elected to have them earlier, suggesting that these people experienced a legal chill.

Penney has described the four different types of chilling effects mentioned in the introduction of this paper and performed an empirical study of surveillance chilling [3]. After news about Edward Snowden and the NSA surveillance leak, Penney used Wikipedia data to illustrate how people were experiencing various types of the chilling effect. There were certain key terms that the government was disclosed to be monitoring, and Penney’s study [23] gathered data of online behavior before and after the release of this list to see whether the number of lookups of these particular keywords and their correlated Wikipedia page had decreased. The results showed that traffic to these sensitive keyword pages had statistically significant decreases and suggested immediate and long-term chilling effects resulting from the general populous gaining knowledge of the government surveillance [23].

This prior work that has relied on specific events that would potentially cause a chilling effect and gather data on behavior change before and after, which informs our proposed study methods and our research.

3 Proposed Studies

Potential study ideas and research methods for uncovering evidence of different types of chilling are discussed to address RQ 1. These methods are not verified and are not guaranteed to provide evidence of chilling effects in different contexts, but may be a starting point for further exploration. We do not suggest projects in the social chilling context because we do not have knowledge of the field.

3.1 Legal Chill

A proposed project idea for studying a legal chill is to compare how different jurisdictions approach laws that may cause a chilling effect. Some suggested laws that could be studied are anti-strategic lawsuit against public participation (SLAPP) laws [16], anti-protest laws [17], and animal enterprise terrorist acts (AETA) laws [19] to see how pronounced the legal chill is across different states. Further research could use data from the Institute of Free Speech in order to contextualize their data. Currently, there is a study that ranks states by political giving freedom, but future studies hope to

publish ratings of state laws based on other types of political speech restrictions [18].

3.2 Personalized Law Chill

A proposed project idea for studying a personalized chill is to survey those who have experienced it. We expect to find evidence of people changing their behavior in some way after receiving notification of violating a specific law. Some instances of personalized chill would be censorship by copyright infringement bots on social media platforms [20]. Thus, a study could be done by gathering data from social media users such as Youtube users and content creators, Twitter users, and Facebook users alike who have had their posts removed due to copyright issues and survey if or how they have changed their behavior in response.

3.3 Enforcement Chill

A proposed project idea for studying an enforcement chill is to gather data from immigration and visas processes. During the process, applicants must submit their social media accounts in order to be vetted, and this may cause a chilling effect in that they are afraid that certain content on their social media accounts may result in their visa being denied [21]. Thus, those who are in the process of obtaining a visa may experience a chilling effect and could be surveyed for their experiences to understand if they change their behavior, and why or why not. Further, information on people’s privacy-preserving strategies could be surveyed.

3.4 Surveillance Chill

A proposed project idea for studying a surveillance chill is to survey those who have experienced the threat of surveillance or have been actively surveilled by such software tools. A potential topic to explore would be the Israeli spyware, Pegasus, and to survey journalist’s experiencing a chilling effect as a result. Marczak et al. [22] has performed studies similar to this proposed idea.

4 Project Focus

To address RQ 2, a topic related to enforcement chill has been reviewed in depth. While there is an official organization by the same name Black Lives Matter (BLM), BLM is a political and social movement whose mission is to eradicate white supremacy and combating misinformation and acts of violence inflicted on Black communities. It gained traction in 2013 in response to the acquittal of Trayvon Martin’s murderer [24]. On May 25,

2020, George Floyd’s death instigated widespread protests covered by media in response to police brutality and racism [25].

For this project, evidence of the chilling effect is gathered from various social media platforms, specifically on Twitter, regarding these protests. Federal agents, as well as local police departments, have been utilizing social media monitoring tools in order to gather intelligence on BLM protests [5, 6]. This surveillance could cause a chilling effect on protesters because protesters may opt to refrain from using social media or other insecure communication channels in order to avoid detection.

4.1 Methodology

This section reviews and discusses the various data collection methods and tools used for this research.

4.1.1 Twitter Data

Given that local police have been using social media to monitor the BLM protests, the social media platform, Twitter, may prove to support the conjecture of people attempting to avoid surveillance by reducing their tweeting about BLM. The publicly available data is to be collected from Twitter from November 2019 to January 2021 using the Twitter API. May and June 2020 have been selected to constitute as the main protest period based on data from Penney’s Wikipedia study [26]. Additionally, data will be collected from six months prior and after the selected time frame to determine whether there is a statistically significant change in behavior after the protests. Data from multiple geographical locations impacted by increased arrests and government response due to the BLM protests are collected as well. Specific locations to look into may include those in which reports show to have social media monitoring. This information is contextualized by collecting data from locations with limited protest activity in the same state as the locations with significant protest activity.

4.1.2 Application Program Interface Data

Data is to be collected using website traffic from popular sites during the BLM protests and data on secure communication application downloads using Application Program Interfaces (API). An increase in downloads of secure communication applications such as Signal and Keybase may indicate that users felt the need to switch to other channels in order to communicate without fear of surveillance. Furthermore, an increase in traffic to popular sites during the BLM protests related to safe protesting practices may indicate that protesters are concerned

about their safety related to protesting. These data will provide context for the Twitter data.

4.1.3 The Onion Router Data

The Onion Router (Tor) is free software that enables secure communication based on a decentralized network design [27]. It was launched in 2002 and the original technology was developed by the United States (US) Navy [28]. This tool may be used by legitimate users such as journalists, activists, and more. We are hoping to survey Tor users as well as to gather Tor statistic data in order to determine why people started to use Tor as well as determine its usability. Since Tor enables secure communication, data about download increases during our timeline could indicate that protesters used this software in order to avoid surveillance.

5 Future Work

Future work could include using similar data collection methods on other social media platforms like Facebook, Youtube, Instagram, and Reddit to determine how pronounced the social media chill is across platforms. In order to gain a more robust understanding of chilling effects, future projects could be collaborations with interdisciplinary members. Furthermore, studies could be done on peer-enforced chilling to understand whether a power-imbalance may result in different intensities of chilling. User sentiment from sentiment analysis about actions that could cause chilling may elucidate what the general public's understanding and feelings toward chilling effects are.

6 Acknowledgments

This work has been made possible by the support of the Computing Research Association's Distributed Research Experience for Undergraduates program, Ph.D. students Nafisa Nowshin and Alexandria LeClerc at Oregon State University, Dr. Glencora Borradaile at Oregon State University, and grant #1915768 from the National Science Foundation (NSF).

References

- [1] "First Amendment," Legal Information Institute. [Online]. Available: https://www.law.cornell.edu/constitution/first_amendment. [Accessed: 26-Aug-2021].
- [2] F. Askin, "Chilling Effect," The First Amendment Encyclopedia, 2009. [Online]. Available: <https://www.mtsu.edu/first-amendment/article/897/chilling-effect>. [Accessed: 26-Aug-2021].
- [3] J. Penney, "Understanding Chilling Effects," Minnesota Law Review, no. 106, pp. 1–78, Jun. 2021.
- [4] S. Livingston, "New York state senator Introduces 'social media hate Speech Accountability Act'", Harvard Journal of Law and Technology, 12-Feb-2020. [Online]. Available: <https://jolt.law.harvard.edu/digest/new-york-state-senator-introduces-social-media-hate-speech-accountability-act>. [Accessed: 26-Aug-2021].
- [5] N. Ozer, "Police use of social media surveillance software is escalating, and activists are in the digital crosshairs," ACLU of Northern CA, 22-Sep-2016. [Online]. Available: https://medium.com/@ACLU_NorCal/police-use-of-social-media-surveillance-software-is-escalating-and-activists-are-in-the-digital-d29d8f89c48#.fowkro6dy. [Accessed: 26-Aug-2021].
- [6] "Map: Social Media Monitoring by Police Departments, Cities, and Counties," Brennan Center for Justice, 10-Jul-2019. [Online]. Available: <https://www.brennancenter.org/our-work/research-reports/map-social-media-monitoring-police-departments-cities-and-counties>. [Accessed: 26-Aug-2021].
- [7] G. Greenwald, E. MacAskill, and L. Poitras, "Edward Snowden: The whistleblower behind the NSA surveillance revelations," The Guardian, 11-Jun-2013. [Online]. Available: <https://www.theguardian.com/world/2013/jun/09/edward-snowden-nsa-whistleblower-surveillance>. [Accessed: 26-Aug-2021].
- [8] S. Biddle, "Police Surveilled George Floyd Protests with help From Twitter-Affiliated Startup Dataminr," The Intercept, 09-Jul-2020. [Online]. Available: <https://theintercept.com/2020/07/09/twitter-dataminr-police-spy-surveillance-black-lives-matter-protests/>. [Accessed: 26-Aug-2021].
- [9] A. Cornish, "Police monitoring of social media sparks concerns in black and brown communities," NPR, 21-Aug-2020. [Online]. Available: <https://www.npr.org/2020/08/21/904646038/police-monitoring-of-social-media-sparks->

- concerns-in-black-and-brown-communities. [Accessed: 26-Aug-2021].
- [10] “Baltimore County Police Department and Geofeedia Partner to Protect the Public During Freddie Gray Riots,” ACLU, Oct-2016. [Online]. Available: http://www.aclunc.org/docs/20161011_geofeedia_baltimore_case_study.pdf. [Accessed: 26-Aug-2021].
- [11] F. Schauer, “Fear, Risk and the First Amendment: Unraveling the Chilling Effect,” William and Mary Law School Scholarship Repository, 1978.
- [12] B. Canes-Wrone and M. C. Dorf, “Measuring the chilling effect,” New York University Law Review, Oct. 2015.
- [13] “Doe #1 v. Reed,” Oyez. [Online]. Available: <https://www.oyez.org/cases/2009/09-559>. [Accessed: 26-Aug-2021].
- [14] “Reaffirming the freedom of the Press: Another look at Miami Herald Publishing Co. v. Tornillo,” Michigan Law Review, vol. 73, no. 1, p. 186, 1974.
- [15] “Buckley v. Valeo,” Oyez. [Online]. Available: <https://www.oyez.org/cases/1975/75-436>. [Accessed: 26-Aug-2021].
- [16] “State anti-slapp laws,” Public Participation Project. [Online]. Available: https://anti-slap.org/your-states-free-speech-protection/#s_corecard. [Accessed: 26-Aug-2021].
- [17] S. Quinton, “Eight States enact anti-protest laws,” The Pew Charitable Trusts, 21-Jun-2021. [Online]. Available: <https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2021/06/21/eight-states-enact-anti-protest-laws>. [Accessed: 26-Aug-2021].
- [18] S. Blackburn, “Free speech rankings – grading 50 states on political giving freedom,” Institute For Free Speech, 02-Mar-2018. [Online]. Available: <https://www.ifs.org/research/freespeechindex/>. [Accessed: 01-Sep-2021].
- [19] “The animal Enterprise Terrorism Act (AETA),” Center for Constitutional Rights, 19-Nov-2007. [Online]. Available: <https://ccrjustice.org/home/get-involved/tools-resources/fact-sheets-and-faqs/animal-enterprise-terrorism-act-aeta>. [Accessed: 26-Aug-2021].
- [20] K. R. Casey, “How to Avoid Copyright Infringement; Bots Are Watching!,” Pennsylvania Bar Association Intellectual Property Law Section, 2017.
- [21] U.S. department of state, 04-Jun-2019. [Online]. Available: <https://travel.state.gov/content/travel/en/News/visas-news/20190604-collection-of-social-media-identifiers-from-U.-S.-visa-applicants.html>. [Accessed: 27-Aug-2021].
- [22] B. Marczak, J. Scott-Railton, S. McKune, B. A. Razzak, and R. Deibert, “Hide and seek: Tracking NSO group’s Pegasus spyware to operations in 45 countries,” The Citizen Lab, 08-May-2020. [Online]. Available: <https://citizenlab.ca/2018/09/hide-and-seek-tracking-nso-groups-pegasus-spyware-to-operations-in-45-countries/>. [Accessed: 26-Aug-2021].
- [23] J. Penney, “Chilling Effects: Online Surveillance and Wikipedia Use,” Berkeley Technology Law Journal, vol. 31, no. 1, Apr. 2016.
- [24] “About,” Black Lives Matter, 16-Oct-2020. [Online]. Available: <https://blacklivesmatter.com/about/>. [Accessed: 27-Aug-2021].
- [25] E. Hill, A. Tiefenthäler, C. Triebert, D. Jordan, H. Willis, and R. Stein, “How George Floyd was killed in police custody,” The New York Times, 01-Jun-2020. [Online]. Available: <https://www.nytimes.com/2020/05/31/us/george-floyd-investigation.html>. [Accessed: 27-Aug-2021].
- [26] R. Kishi, “Demonstrations and political violence in America: New data for summer 2020,” ACLED, 26-May-2021. [Online]. Available: <https://acledata.com/2020/09/03/demonstrations-political-violence-in-america-new-data-for-summer-2020/>. [Accessed: 27-Aug-2021].
- [27] “The Tor Project: History,” Tor Project. [Online]. Available: <https://www.torproject.org/about/history/>. [Accessed: 01-Sep-2021].
- [28] S. Dredge, “What is Tor? A beginner’s guide to the privacy tool,” The Guardian, 05-Nov-2013. [Online]. Available: <https://www.theguardian.com/technology/2013/nov/05/tor-beginners-guide-nsa-browser>. [Accessed: 01-Sep-2021].