Challenges in Climate Change Communication on Social Media

Submitted in partial fulfillment of the requirements for
the degree of
Doctor of Philosophy
in
Engineering and Public Policy

Aman Tyagi

B.Tech., E.C.E, IIIT Hyderabad
M.S., E&TIM, Carnegie Mellon University
M.S., EPP, Carnegie Mellon University

Carnegie Mellon University
Pittsburgh, PA

May, 2021

For all those who still care

Acknowledgments

First and in particular, I would like to thank my advisor Prof. Kathleen M. Carley. I am grateful and lucky to be advised by her; I will never forget her empathy and support, without which this thesis would not have been possible. Next, I would like to thank my committee members, without their help this thesis would not have been anywhere near the current quality. Prof. Douglas C. Sicker for helping me and guiding me during those tough times. Prof. Alex Davis for motivating me to take up research and helping me throughout my Ph.D. student years. Prof. Dongwon Lee for his help and consideration during the process. I would also like to thank Vicki Finney, Debbie Kuntz, and Adam Loucks for their constant help.

I like to thank my lab friends in CASOS and IDEAS. Matthew Babcock helped me get at pace with research and think about the host of different ideas discussed in this thesis. Sumeet Kumar, David Beskow, Thomas Magelinski, and Ramon Villa Cox answered my questions and led the way, which I just have to follow. Lastly, I would like to thank Joshua Uyheng for helping me in making this thesis possible.

I am grateful to my friends for supporting me throughout my Ph.D. journey. Mohit and Tanya for being loving roommates, especially during this pandemic. I want to thank my CMU peers - Priyank, Sai Gopal, Prithvi, Christophe, Rohit, Jorge, Juan, Nyla, Turner, Priya, Kristen, Matt, Nicole, and many more. I also want to thank my childhood friends - Darsh, Manasvi, Ankita, Jasmeet, Rohit, Abhishek (*2), Abhilash, Ayush, Govind, Piyush, Harshit, Jayant, Ashish, Siddharth (*2), Nitish, among others.

Of course, none of this would have been possible without my family. Papa, Mummy, Neha, and Abhishek, thank you for your comfort and constant support.

The research reported in this thesis has been supported in part by the Department of Engineering and Public Policy at Carnegie Mellon University, the Knight Foundation, the Center for Computational Analysis of Social and Organization Systems (CASOS) and the Center for Informed Democracy & Social - cybersecurity (IDeaS). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the funding agencies.

Abstract

In today's fast-paced lifestyle, internet users depend on social media platforms to obtain and debate essential socio-political and economic topics. However, this same vital source suffers from various challenges. On social media platforms, such as Twitter, users do not necessarily face a lack of information; instead, they are overwhelmed with diverse information sources. These myriad sources of information on social media can make users unknowingly confined to or associated with other users or groups. Moreover, facts or news can be reported in ways that create confusion and affect public sentiment on scientific actualities. Such social media challenges can cause a long-lasting impact in reshaping our society, slowing down scientific progress, and dampen regulatory endeavors. Thus, social media's impact on socio-political and economic topics must be analyzed. In this thesis, I analyze each of these problems using conversations and news articles about one of the most significant challenges our society faces today, i.e., climate change.

In my first study, I analyze climate change discussions on Twitter to study users confined to competing belief groups. I classify Twitter account users into: (a) users who believe in the anthropogenic cause of climate change (Believers); and (b) users who don't (Disbelievers). I study the differences in communication topics and network structure in Disbelievers and Believers. I find that both Disbelievers and Believers talk within their group more than with the other group; this is more so the case for Disbelievers than for Believers. In my second study, I develop a framework to quantify hostile communication between Believers and Disbelievers. I show that Disbelievers of climate change are more hostile towards Believers than vice versa. I examined the framing bias of climate change news articles shared on Twitter as part of my third study. I find that climate change news articles are predominantly framed as related to policy issues in the context of a social group's traditions, customs, or values. Finally, I explore the spread of conspiracy theories in climate change conversations on Twitter. Results suggest that Disbelievers are primarily responsible for sharing messages that contain keywords related to conspiracy theories. Overall, my work in this thesis develops frameworks to analyze social media challenges and contributes to climate change communication research.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Information media, such as print media and television, has always been influential in shaping public opinion. Before the advent of the World Wide Web, the number of distribution channels in information media and breadth of content was limited. With the dawn of new media technologies, the Web democratized the information distribution channels by giving voice to each individual, which was then bolstered by various social media platforms. These social media platforms, such as Twitter and Facebook, gave unrestricted access to information to the masses, decreasing the information gap between the elite and the newly strengthened middle-class [105].

Although there are many benefits to social media platforms, they also suffer from multiple challenges. Firstly, the current social media platforms have several sources of information about topics, which goes beyond human cognition to process and draw meaningful inferences about this information [130]. In this information overflow landscape, malicious entities can use alternative facts to mold the truth and portray their own opinion. Thus, creating a planned confusion that benefits certain people or a group of people with similar agendas. Moreover, the proliferation of these alternative facts (or lies) on social media platforms is higher compared to the actual fact (or truth) [157].[1] One vast subfield of alternative facts or lies is conspiracy

[1] In §5.6.1 we discuss different types of false information with examples from climate change discussion.

theories, which aim to create confusion on scientific facts. These conspiracy theories are likely to be shared by a significant number of people, potentially driving conversations about specific topics.

Secondly, users on these forums could be confined to certain topic groups, such as users supporting one view would not be exposed to another viewpoint. This leads to the formation of echo-chambers [145], or groups of people with certain viewpoints and beliefs. These "echo-chambers" create bias in information consumed by social media users. This problem only becomes more acute with the use of user-specific recommendations [1, 46], which potentially segregate users by recommending information that is in line with their views. In social media literature, this phenomenon is called positional polarization. Moreover, a group of people with similar beliefs could be hostile to people of opposite beliefs. This leads to affective polarization. Affective polarization could further widen the divide between the groups.

Lastly, news shared on social media platforms could be framed in ways such that specific points are emphasized or de-emphasized to create confusion or further a political agenda. For example, research shows that authoritarian governments can benefit from altering/framing news on social media to further their aim [61, 149]. Moreover, the use of these subtle language framing can even create confusion on scientific facts leading people to disregard scientific consensus [134].

Given these challenges described above and their impact on the economy and society, researchers should explore technical and policy tools to mitigate such problems. Today, social media platforms are essential sources for sharing and debating important socio-political and economic topics [138]. In this thesis, I focus on one of the most important challenges humanity faces today, i.e., climate change.

I develop methods to analyze the above mentioned challenges in climate change discussion on social media. Previous research on climate change communication has relied on survey based and manual methods. Traditional methods such as surveys could cost $70,000 for typical 1,000 participants [52]. Apart from being expensive, traditional methods suffer from other limitations. Surveys are limited in finding nuanced beliefs and in studying extensive social network structures. These methods also suffer from well-known biases such as "social desirability bias" [52], which could be detrimental in studying socially sensitive and partisan topics such as climate change. In other words, asking people directly might not be as accurate as using social media data. On the other hand, in my thesis, I rely on rich social media data to uncover the climate change communication challenges. Social media provides real-time, extensive, but fine-grained, and unfiltered data that could be used to find nuanced beliefs and analyze extensive social networks. Comprehensive social media data represent population across age, gender, income level, and other demographics. The complexity and size of social media data warrant the use of new computational methods. As part of this thesis, I develop computational methods that could be used on big data feasibly. Thus, this thesis uses social media and computational methods to provide new avenues to progress climate change communication research.

According to United Nation's ex-general secretary Ban Ki-Moon climate change is potentially the "challenge of our generation" [114]. Although there is virtually 100% consensus among scientists about the anthropogenic cause of climate change [126], multiple studies suggest that between 20%-40% of the U.S. public believe that climate change is a hoax [151]. Climate change scientists have alerted the policymakers about the dangers of current carbon-related policies, but a significant set of people resist any change in policies. Out of the people who resist any change in policies, arguably the most alarming subset of people are those who do not believe in scientific facts. Climate change is an interesting topic as although there is near consensus among experts, there is polarization in belief among the general public.

As more and more people get their news from social media, the role social media plays in shaping our society's perspective on the climate change debate must be analyzed. Understanding these conversations is warranted to help recognize people's beliefs on climate change and the underlying constructs by which more people are either attracted or repelled by the topic.

All of this can help tailor appropriate future messaging. Although the methods discussed in this thesis can be applied to other social-economic topic, we consider climate change as a case study because of its socio-economic importance, urgency, and lack of research in this area.

1.1 Dataset Contribution

1.1.1 Twitter Dataset

I collected tweets using Twitter's standard API[2] with keywords "Climate Change", "#ActOnClimate", "#ClimateChange". The dataset was collected between August 26th, 2017 to September 14th, 2019. Due to server errors, the collection was paused from April 7th, 2018 to May 21st, 2018, and again from May 12th, 2019 to May 16th, 2019. I ignore these periods in my analysis. The collected dataset consists of 38M unique tweets and retweets from 7M unique users after deduplicating tweets [3]. In Table 1.1, I describe the statistics of the dataset. I use these tweets in Chapters 3,4 and 5 of the thesis.

1.1.2 Climate Change News Articles

I collected news articles shared via the Tweets collected above. I scrape all the articles shared by news agencies on Twitter using the collected Tweets. To find out whether an account is from a news agency, I use a pre-trained model as described in [80]. The model uses a long-short-term memory neural network [77] with an attention mechanism [15] trained on over 10,000 users. The test accuracy reported on a held-out dataset is 91.6%. I found ∼4% percent of users as news agency account with ∼1.1M unique tweets and retweets. I scraped the article shared via URL for each tweet by using python's *requests* library. I collected 900k files shared via URL. Out of these 900k files, I removed the files which were non-text files and all the files with the error

[2] `https://developer.Twitter.com/en/docs/tweets/search/overview/standard`

[3] The dataset can be found at `/storage/sumeetsstuff/debatetopics/climatetweets jan24.zip`

Table 1.1: Statistics of the Tweets used in Chapters 3,4, and 5 and the articles collected to analyze media framing in Chapter 4.

	Tweets	**Articles**
Total Number	38M	810k
Mean per day	48,860.5	1,157.5
Min per day	2	0
Max per day	243,574	6,513

message returned from scraping the news outlet's website. After removing the unwanted files, I was left with 810k articles spread across the same timeframe as the Tweets dataset above [4]. In Table 1.1, I describe the statistics of the dataset. I will refer to these articles as *news articles*, I use this dataset in Chapter 4.

1.2 Summary of Thesis

This thesis is composed of the following chapters:

1. Chapter 1: Introduction
2. Chapter 2: Positional Polarization (published at SBP-BRiMS 2020 [148])
3. Chapter 3: Affective Polarization (published at ASONAM 2020 [150])
4. Chapter 4: Media Framing (under review)
5. Chapter 5: Climate Change Conspiracy Theories
6. Chapter 6: Conclusion

Chapter 2 presents a framework to find and analyze positional polarization among two com-

[4]The dataset can be found at `/storage3/amant/articles`, `/storage3/amant/articles2/articles`, `/storage3/amant/articles3`. Each article is saved in format 'UserID'_'StatusID'.txt where 'UserID' and 'StatusID' are unique Twitter's user account identification and Twitter's status identifier respectively.

peting groups of Twitter users. In this chapter, I present a case study that analyzes the conversation between two competing groups of Twitter users, one who believe in the anthropogenic causes of climate change (Believers) and second who are skeptical (Disbelievers). These discussions occurred during the United Nation's (UN) Climate Change Conference – COP24 (2018), Katowice, Poland. I use hashtags from tweets and retweets associated with Twitter accounts to classify users into Disbelievers, Believers, or neutral groups. I find about seven times the number of Believer accounts compared to Disbeliever accounts. I classified accounts exhibiting bot-like behavior and news agencies using state-of-the-art methods and explore their activities in Believer and Disbeliever networks. I find that both Disbelievers and Believers talk within their group more than with the other group; this is more so the case for Disbelievers than for Believers. The Disbeliever messages focused more on attacking those personalities that believe in the anthropogenic causes of climate change. On the other hand, Believer messages focused on calls to combat climate change. I find that in both Disbelievers and Believers bot-like accounts were equally active. Unlike Believers, Disbelievers get their news from a concentrated number of news sources. Lastly, I find multiple tweets that spread a variety of conspiracies in climate change conversations.

Continuing my work on polarization, next I look at the affective aspect of polarization. In Chapter 2, I use a case study to discuss how people are divided into echo-chamber-like groups. Chapter 3 presents my work related to quantifying hostile communication or *affective polarization* between two competing groups. I propose a systematic, network-based methodology for examining affective polarization in online conversations. Further, I apply my framework to 100 weeks of Twitter discourse about climate change. I find that deniers of climate change (Disbelievers) are more hostile towards people who believe (Believers) in the anthropogenic cause of climate change than vice versa. Moreover, Disbelievers use more words and hashtags related to natural disasters during more hostile weeks as compared to Believers. These findings bear implications for studying affective polarization in online discourse, especially concerning the subject of climate change. Lastly, I discuss my findings in the context of climate change communication research.

In Chapter 4, I present the work on framing analysis of news articles on climate change. Information presented by news media channels could be manipulated in ways to emphasize or de-emphasize a particular topic. In this chapter, I present ways to find framing bias in news media articles. I use the Media Frames Corpus (MFC) [35] and develop a method to find the framing bias on 810k news articles shared on Twitter about climate change. Further, I connect the framing analysis to Affect Control Theory (ACT) to find each type of framing's emotional value. I find that the *cultural identity* frame dominates in climate change news articles. Moreover, I find that climate change news articles' frames are low in emotional value, and the emotional value does not change over the 100 week period of our dataset. We also conclude that frames are not reshared based on their *affect*. To the best of my knowledge, this is the first attempt to connect the computational frames to ACT. I expect my work as an important stepping stone for social scientists to build better communication analysis tools for future climate change communication messages.

One of the most essential emerging challenges in climate change communication is the prevalence of conspiracy theories. In chapter 5, I discuss some of the major conspiracy theories related to climate change. I use state-of-the-art stance detection method (also used in chapter 3 of this thesis) to find whether conspiracy theories are more popular among Disbelievers or Believers. I find that Disbelievers are overwhelmingly responsible for sharing messages with conspiracy theory related keywords. Lastly, I report the conspiracy theories that are more popular than others and how their sharing pattern is related. I end the chapter with a discussion on other climate change conspiracy theory related work.

Chapter 6 summarises this thesis.

Chapter 2

Positional Polarization

2.1 Introduction

Social media platforms such as Twitter have become an important medium for debating and organizing around complex social issues [138]. One such complex issue with significant socio-economic and political implications is climate change. [162] studied how Twitter is used as a medium for debating climate change and their work found segregated polarized attitude towards the causes of climate change.

Debates over climate change involve different groups with different inherent motivations and beliefs. For instance, among the people who are skeptical of climate science findings are people who outright reject the data that climate change is occurring, and others who argue that climate change is occurring due to non-anthropogenic causes. In a similar manner, there is significant difference in beliefs among people who believe in anthropogenic causes of climate change. Work by Matthews [108] argue that there are groups who believe that impact of climate change is exaggerated (so-called "luke-warmers"), others who argue that we need an across-the-board technological change in energy production [140] or even end of capitalism [8, 99]. Furthermore, there are groups who argue that it is already too late to avoid climate catastrophe [137]. In this paper, we analyze conversations between two broad competing groups of Twitter users, one who believes in anthropogenic causes of climate change (Believers) and a second who are skeptical or outright deny climate change is occurring (Disbelievers).

Previous studies suggest that somewhere between 20% and 40% of the U.S. public believe that climate change is a hoax [151]. The beliefs around climate change tend to depend upon location [79], political inclination, and education attainment [56]. People tend to "persuade themselves to change attitude and behavior" [124] and hence communicators should tailor climate change messaging based on the beliefs of audience groups [100]. An understanding of the conversation between different groups is needed to help recognize the beliefs of these groups and the underlying constructs by which more people could be attracted or repelled by different messaging, all of which is helpful for tailoring appropriate future messaging.

Fault lines between groups could also be attractive to entities seeking to manipulate consensus. Exposure to debates on anthropogenic cause of climate change may lead people to believe that there is no scientific consensus [138]. Additionally, social media is frequently used to spread disinformation [3]. Disinformation about climate change could be promulgated by bots - automated user accounts - in addition to human actors. Previous studies suggest that bots seek to create false amplification of contentious issues with the intention to create discord [59]. There has been previous research that studied false stories such as conspiracy theories in the context of climate change [151]. That work looked at the how climate change "denialism" is at least partially driven by underlying mindset of people believing in conspiracy theories. In this work, we validate the arguments made in that study by looking at false stories and conspiracy theories shared on social network in climate change debates.

In this paper, we present a case study by analyzing conversations between two competing groups of Twitter users - Believers and Disbelievers, during United Nation's (UN) Climate Change Conference – COP24 (2018), Katowice, Poland. Previous studies about climate change discussions on social media, such as [162] and [42], lacked the context of a significant event. They also didn't take into account the behavior of bots and false information during such an event. We examine what role, if any, that bots and disinformation stories play within Disbeliever and Believer competing groups. By restricting the study to a particular event, we were able to manually inspect large fractions of stories in the competing groups. This case study should be helpful to inform future studies regarding climate change conversations on social media that covering longer time span. Our research questions are as follows:

1. What are the conversational subtopics within the Believer and Disbeliever groups, and what does specific hashtag and common word use by these competing groups highlight? Do individuals of one group interact with individuals of the other group? What are the popular sources of information within these groups?
2. Are bots more active in one particular group over another? Are bots driving or amplifying the conversations within these groups?
3. What are the common disinformation stories within these groups?

We analyze these research questions using Twitter conversations on climate change during COP24. We describe our data collection method in §2.2.2. We use hashtag based method to classify users into Disbelievers and Believers described in §2.2.3. In §2.3 and §2.4 we present our results and their implications. Through this research study we provide a framework to analyze polarizing networks and the implications for climate change discussion.

2.2 Background and Method

In this section, we first give a brief background of Twitter messaging and previous research on Twitter conversations. We then describe our data collection procedure. Lastly, we describe the methods used to identify groups within the twitter user base, to isolate fake stories, conspiracy theories and exaggerations, and to detect bots and news agencies.

2.2.1 Background

We look at four main types of communication on Twitter in this paper: 1) Tweet-ing – a user account sharing a message with its followers and the public, 2) Retweet-ing – a user account forwarding a Tweet so as to share the message, 3) Replying – a user account replying to a Tweet by another user account, 4) Mentioning – a user account explicitly mentioning another user account in its tweet. We call the sum of the three types of communication as "all communication". In this paper, we look at these communications as networks and find network measures to compare and con-trast communication from and between Disbelievers and Believers.

Twitter has been an important social media platform to study conversations about natural disasters, medical decisions, race relations and numerous other important is-sues. Dredze et al. [53] studied conversations of vaccine opponents during 2016 US presidential elections. Dredze and colleagues [51] additionally showed how misconceptions were spread about Zika vaccine by the vaccine skeptic community on Twitter. Broniatowski, et al. [31] showed how Russian trolls and bot-like accounts promoted discord in vaccine-relevant messages on Twitter. Babcock et al. [11] studied diffusion of different disinformation stories on Twitter related to the movie Black Panther. This work found fake stories maligning one particular community and established how satirical posts were helpful in calling out the disinformation stories. In a different work using the same dataset, Babcock and colleagues characterized dis-information and found that all disinformation is not the same [12]. Twitter has also been used to study conversations between social groups having contrasting beliefs. Yardi & Boyd [165] studied conversation between pro-life and pro-choice groups and how messaging within a group helped in strengthening identity of that group members. Studying social media marketing to different groups helps marketers design more engaging messaging as shown in a case study by Burton et al. [34]. This work looked at Twitter accounts advocating safe drinking or abstinence vs accounts maintained by alcohol companies to find difference in reach and influence between those two groups. Specifically, with regard to climate change conversations [122] divided twitter accounts into supportive, unsupportive or neutral to the 2013 Intergovernmental Panel on Climate Change's (IPCC) Fifth Assessment Report (AR5) and studied popular topics within these communities. Kirilenko et al. [98] showed how Twitter conversations about climate change are driven by extreme climate anomalies. Jang & Hart [86] studied the conversations on Twitter and how framing of these conversations could be different depending on user's location. In this paper, we compare and contrast messaging of Believers and Disbelievers on Twitter to help in understanding the differences in underlying reasoning which is creating and nourishing the fault lines between the groups.

2.2.2 Data Collection

UN Framework Convention on Climate Change's (UNFCCC) Conference of Par-ties (COP) is an annual meeting of different states represented at the UN and acts as a venue to discuss the progress and establish obligations with regards to responding to climate change [118]. This event provided an opportunity to look at the Disbeliever and Believer climate change messaging on Twitter in context of a significant event.

We collected tweets with hashtags and certain keywords from November 27th to December 20th, 2018 using Twitter's API. We decided on collection hashtags based on hashtags related to #climatechange found on best-hashtags.com. We added more keywords based on these hashtags and news articles found after searching for keyword "COP24" on Google [1]. The combined data set contains a total of 1,379,584 distinct tweets (including retweets).

[1]Hashtags and keywords used for collection:#COP24, #ClimateChange, #ClimateHoax, #ParisAgreement,#IPCC, #InsideCOP24,#Climate, #ClimateChangeisReal, #ClimateAction, #GlobalWarming, COP24, Climate Change, Paris Agreement, Climate Hoax, IPCC, Climate, Global Warming

2.2.3 Method to Identify Groups

We identified competing groups of Believers and Disbelievers by hashtags used by these groups. Hashtags have been shown to be a realistic substitute to identify stances among different groups on social media [58]. For example, previous studies suggest that climate Disbelievers use terms such as hoax and scam [132]. We analyzed common hashtags used in our dataset and found that "ClimateHoax" and "ClimateChangeIsReal" hashtags are used mostly by Disbelievers and Believers respectively. There are 528 distinct tweets with keyword "#ClimateHoax" and 9,008 tweets with keyword "#ClimateChangeIsReal" in our data set. We manually checked all tweets with hashtag "ClimateHoax" and randomly sampled 1,000 tweets from data subset with hashtag "ClimateChangeIsReal". We identified 96% of tweets with "#ClimateHoax" as climate change Disbeliever tweets. For "#ClimateChangeIsReal" out of the 1,000 randomly selected tweets, we identified about 99% as climate Believer tweets. We therefore conclude that hashtag "ClimateHoax" and hashtag "ClimateChangeIsReal" can be used as proxies for tweets broadcasted by Disbelievers and Believers respectively in our data set.

To identify more hashtags used by Believers and Disbelievers, we use the method described in [149]. We choose hashtags which are most used with hashtag "ClimateHoax" and hashtag "ClimateChangeIsReal" and are associated with conspiracy in case of Disbelievers or have similar meaning to "ClimateChangeIsReal" in case of Believers [2]. We give an initial weight of -1 to Disbeliever hashtags and +1 to Believer hashtags. We use these labels in a weighted hashtag x hashtag co-occurance network, to find an average label from -1 to 1 for other hashtags. The method used for propagating labels to other hashtags is reported in Algorithm 1. We aggregate hashtags used by each user and found a weighted average of all hashtags used by a particular user. We label a user as Disbeliever, Believer or unclassified if the weighted average was negative, positive or zero respectively. We assume that within our collection period

[2] Disbeliever hashtags: ClimateHoax, YellowVests and Qanon. Believer hashtags: ClimateChangeIsReal,ClimateActionNow, FactsMatter, ScienceMatters, ScienceIsReal

Disbelivers or Believers do not change their stance and hence unlike in [149] we only look at aggregate polarized hashtags over entire dataset. The algorithm is similar to methods used to infer user-level polarities, in which a small seed of users is hand-annotated and a graph-based algorithm propagates labels to other users by assuming that users who retweet each other share the same views [45, 159]. For example, [65] quantify polarity based on a graph structure by assuming that the controversial topics induce clusters of discussions, commonly referred to as "echo-chambers". However, we conduct propagation at a hashtag level, by assuming that hashtags that frequently occur in the same tweets indicate similar polarities. Also, our approach does not assume homophily in retweet network nor that user polarities are constant over time. Graph-based approaches have also been used to examine sentiment or for mixed tweet/hashtag/user-level analyses [43, 125].

Overall, we found a set of 8,413 tweets from 2,170 Disbelievers and 120,497 tweets from 15,640 Believers. We randomly sampled 100 users from both groups of users and manually checked their timeline to find approximately 91 percent of Disbelievers as showing activity akin to a Disbeliever and about 96 percent of Believers showing activity akin to a Believer.

2.2.4 Isolating Fake News, Conspiracy Theories and Exaggeration

We searched for tweets within our data with fake and conspiracy stories as follows: We broke each tweet into unigrams (single item) and bigrams (sequence of two items) and removed the stop words as done in [48]. We then searched for keywords in our set of unigrams and bigrams for keywords related to fake and conspiracy stories. We identified fake news shared in social media related to climate change from FactCheck.org, Politico, truthorfiction and hoax-slayer as listed in [157]. We further collected keywords used in each conspiracy theory from the list of conspiracy theories on Wikipedia [161]. We then searched for keywords collected from conspiracy theories and fake news articles in our unigrams and bigrams made from tweets. We further added tokens from our reading of tweets to get a more exhaustive set of tweets with fake news, exaggerations and conspiracies. We report list of tokens in §2.5.1. We found a total of 21,688 tweets containing all tokens. We manually went through each tweet to select stories that were fake, conspiracy related and/or clear exaggerations of effects from climate change.

2.2.5 Bot Detection and Account Types

To find bots accounts in our data set, we used CMU's Bot-Hunter [18, 19]. The output of Bot-Hunter is a probability measure of bot-like behavior assigned to each account. Unless otherwise stated, we report our analysis for a probability threshold of 0.5,as done in various machine learning classification methods [123]. In other words,we classified an account as bot-like if output probability from Bot-Hunter was greater than 0.5. At 0.5 threshold level we found 596,282 bot-like accounts out of total 1,035,416 users in our data set.

We used a classification model trained on the users' tweets and personal descriptions to find news agency accounts associated with our list of user accounts. The model is similar to the state-of-the-art model used in [80]. The paper describes the model as a long-short term memory neural network [77] with an attention mechanism [15]. In total, we find 2.2% of Believer tweets as classified to be from news agencies and 6.2% of Disbeliever tweets as classified to be from news agencies.

2.3 Results

We begin by discussing popular hashtags and topics of interest expressed by Believers and Disbelievers. Then we look at how these users in these groups are interact-ing within and between each group and what are the most popular news agencies within these groups. Next, we look at the bot-like account's behavior in these two groups. Lastly, we present our findings about fake news and conspiracy theories.

2.3.1 Topics of Discussion

To understand the conversations of Disbelievers and Believers we first found the most frequent words used by these competing groups. The results are presented in Table 4. Note that in the construction of unigrams, we exclude common stop words (Refer §2.2.4) From Table 2.1, we see that Believers use words such as "need", "action", "leaders" and "future" more often, potentially indicating tweets calling for action to combat climate change. On the other hand, Disbelievers use words such as "private", "sanders", "end" and "warming", potentially indicating attacks on pro-climate change personalities and their messaging.

To further our understanding of the conversations and to find topics of opinions within these group, we performed topic modelling of tweets by Believers and Disbelievers using Latent Dirichlet Allocation (LDA) [21]. We ran our model to find top ten topics on the unigrams of tweets generated after removing common stop words. Among the top ten topics we report the top 3 list of words we were able to infer topics about in table 2.2. In the first topic Disbelievers use words such as "scam" and "fakenews" with words associated with "climate change", potentially calling out climate change as scam or fake. In the second topic Disbelievers are calling out personalities believing in human caused climate change. In the third topic Disbelievers are talking about yellowvests movement which relates to the French movement against raising fuel taxes based on climate policy [41]. On the other hand, for the first topic Believers use words related to using renewables and giving up fossil fuel. This can be inferred from the use of word "keepitintheground", as the word is used on social media to ban any new use of fossil fuel [69]. The second topic for Believers is about the climate change politics in Australia with words such as "auspol" (short for Australian politics) and "stopadani". Specifically, "stopadani" is used in social media to protest against Adani group of companies digging Carmichael coal mine in Queensland, Australia [166]. Lastly, the third topic for Believers relates to COP24 with word "takeyourseat" used in COP24 to signify the People's seat initiative launched by the UN [118].

Next, we provide further evidence of the political nature of conversations by looking at the trending hashtags in each group. In figure 2.1, we report top trending hashtags before and during the COP 24 conference for Disbelievers and Believers. We remove the hashtags we used to collect our data such as "#climatechange" while reporting our results. Hashtag associated with Australian politics #auspol, is popular among both of the groups. For the Disbeliever group, the #Yellowvests, is one of the most popular hashtags. For Believers, #nca4, which is the hashtag associated with the fourth US climate assessment report, is one of the most popular. From our analysis we can conclude that conversation in these groups are political in nature, this extends McCright and Dunlap [56] conclusions to social networks, where the study found that beliefs around climate change tend to depend on political inclination.

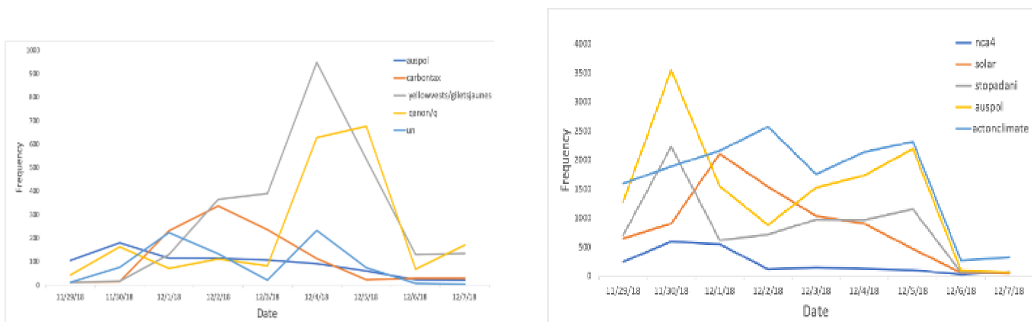


Figure 2.1: Trending hashtags in users classified as Disbelievers (left) and Believers (right).

2.3.2 User Accounts and Conversations

We explored how Believers and Disbelievers are connected to each other and within themselves. We divided the communication between users into: 1) a retweet network – Each node is a user and an edge between two nodes describes whether or not one of the users has retweeted any of the other's tweet, 2) a mentions network – Each node is a user and an edge between nodes describes whether one user mentioned another user in any tweet, 3) a reciprocal network – Each node is a user and an edge between two nodes describes whether both users mentioned each other in our data set, and 4) a reply network - Each node is a user and an edge between nodes describes whether one user replied to the other user. In this section, we first report our results for different network level measures and then we report our results for individual node level measures.

Network Level Measures We first look at different network level measures to contrast Believers and Disbeliever networks. We report various measures for the Believer and Disbeliever communication networks in Table 2.3. Network density is defined as the ratio of actual connections and potential connections [67] and reciprocity is defined as the ratio of bi-directional edges and the total number of edges [158]. We find that the number of users and distinct tweets and retweets is much more for Believers than for Disbelievers (p ¡ 0.05). The network density for different types of communications is greater for Disbeliever networks than for Believers indicating proportionally more connections within their respective community. In contrast, reciprocity is higher in case of Believers than in case of Disbelievers indicating higher fraction of bi-directional communication amongst Believers.

Table 2.3: Network measures for Disbeliever and Believer networks.

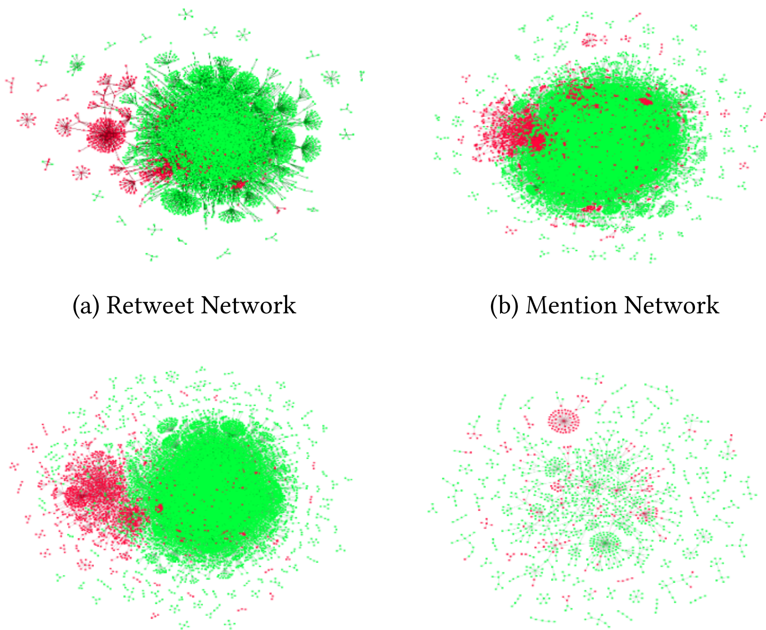
Measure		**Believers**	**Disbelievers**
Distinct Users		15,640	2,170
Distinct Tweets and Retweets		120,497	8,413
Network Density (All Links in parenthesis)	Mention	$1.25 * 10^{-4}$ (49,459)	$5.32 * 10^{-4}$ (4,191)
	Reply	$0.04 * 10^{-4}$ (1,808)	$0.39 * 10^{-4}$ (310)
	Retweet	$0.71 * 10^{-4}$ (28,322)	$2.56 * 10^{-4}$ (2,018)
	All Communication	$1.25 * 10^{-4}$ (49,598)	$5.34 * 10^{-4}$ (4,206)
Reciprocity		$10 * 10^{-4}$	$7.2 * 10^{-4}$

We first look at different Twitter networks formed from various communications to contrast Believers and Disbeliever. In figure 2.2, we report figures for all four networks. In the retweet network, we can see a clear distinction between Believers and Disbelievers; Disbelievers retweet other Disbelievers more than they retweet Believers, and vice versa for Believers. The mentions network of Disbelievers and Believers shows more links between these groups meaning that Believers and Disbelievers do mention users from other groups on tweets. The reciprocal network has less activity between the groups than the mentions network, suggesting that although users from one group mention people from other group, they tend to have reciprocal relationships with their own group. The reply network has a much lower number of nodes compared to other networks which suggests that users in both groups prefer mentioning or retweeting rather than replying to tweets. The stark contrast in mentions and reciprocal activity confirms that users from one group do not engage in conversations with users from another group. After establishing differences in different type of behavior on Twitter, next we look at the combined communication of these groups to check how much these groups talk within themselves, i.e. how much "echo-chambery" these groups are.

To compare echo-chamber effects in these two groups, we combine all the above networks to make a network of all communications to find echo-chamberness(e)[3] for each group with and without unclassified accounts. We find that for Disbelievers $e = 0.007$ and for Disbelievers with unclassified accounts $e = 0.003$. On the other hand, for Believers $e = 0.006$ and for Believers with unclassified accounts $e = 0.003$. The values of e is small compared to a denser symmetric graph because the communications network does not represent the actual follower's network of the users. The e of both groups decreases on adding unclassified accounts, which indicates that each group is talking more to themselves, this is marginally truer for Disbelievers compared to Believers.

Node Level Measures Next, we look at the difference in fraction of most crucial and influential spreaders of information in both the networks. This helps us determine whether or not

[3]For a unimodal network G, the e of G is $(r * d)^{(1/3)}$,where r is the reciprocity of graph G, that is the fraction of edges in the graph that are reciprocal (a symmetric graph has $r = 1$), and d is the density of graph G.



(a) Retweet Network
(b) Mention Network

(c) Reciprocal Network
(d) Reply Network

Figure 2.2: Communication networks between Twitter accounts classified as Disbelievers (red) and Believers (green). The graphs were made using ORA-PRO [4, 37]

.

these groups are influenced by multiple influencers or via a central actor. ORA-PRO twitter report labels users as "super spreader" as the most influential users in spreading information and "super friends" as most crucial users in bi-directional communication on twitter [4, 37]. Super spreaders in ORA-PRO are defined as user accounts in sum of mentioned-by and retweeted-by network which are in top 3 of following measures: 1) Often mentioned/retweeted by others, 2) Iteratively mentioned/retweeted by others, and 3) Often mentioned/retweeted by groups of others.To give an example, in figure 2.3 we visualize the ego network [4] of a super spreader node for Disbeliever network. The super spreader node is connected to many single nodes, meaning that the node is often mentioned/retweeted by others. The super spreader node is also connected to many other nodes which are connected to each other, meaning that the super spreader is often mentioned/retweeted by groups of other nodes. These properties make the super spreader significantly more influential in spreading conversations compared to other user accounts. Similarly, super friends in ORA-PRO are defined as user accounts in sum of reciprocal mentioned-by and re-tweeted-by networks which are in top 3 of the following measures: 1) Often mentioned/retweeted by others, 2) Often mentioned/retweeted with many others, 3) Mentions/retweets in network cliques, and 4) Mentions/retweets in groups.

To compare the two groups, we look at the fraction of user accounts labelled as super spreaders and super friends. We find that Disbelievers (0.48%) have fractionally higher percentage of super spreaders than Believers (0.37%). Disbelievers also have higher fraction of users classified as super friends than compared to Believers (0.38% vs 0.28%). We quantify the modular structure of the Disbelievers and Believers networks by finding the modularity values of different networks. In Table 2.4, we report Louvain modularity [22] value for different networks.

[4] An ego network is a network of a main/focal node with its ties and any connection in between those ties.

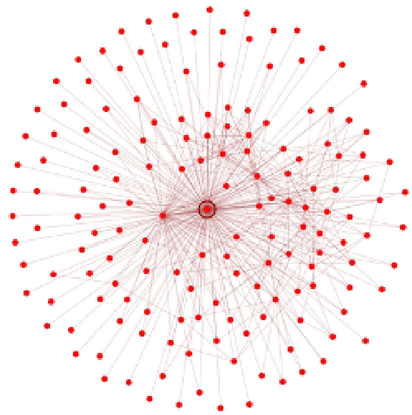


Figure 2.3: An ego network for a super spreader (highlighted at the center) in Disbeliever network of sum of mentioned-by and retweeted-by networks.

Table 2.4: Modularity values of different networks for user accounts classified as Disbelievers, Believers, or neutral.

	Disbeliever		**Believer**		**Neutral**	
	Louvain Modularity	Nodes	Louvain Modularity	Nodes	Louvain Modularity	Nodes
Reply Network	0.942	411	0.962	2264	0.932	48920
Retweet Network	0.811	1702	0.620	13802	0.747	368146
Mention Network	0.783	2802	0.622	19842	0.704	431920
Reciprocal Network	0.759	778	0.618	6290	0.687	321613

For node level analysis of Disbelievers and Believers network, we look at the influencers in each network by looking at the different centrality measures for mentioned-by and retweeted-by networks. We report our findings in Table 2.5. We calculate the total degree centrality for user accounts as done in [63] and report the top 5 normalized centrality values. We can infer from the higher values for Believers that the Believers top 5 accounts are more mentioned and retweeted than Disbelievers top 5 accounts in their respective networks.

Algorithm 1: Label Propagation Algorithm

Input: Graph G; Nodes = *n*; Edges = *e*; Edge Weight = e_{ij}, $i \in n$ and $j \in n$

initialize $\gamma = 100$ and *i*=0;

for *each n* **do**

  define l = integer(i/γ); i+=1;

  for *each n* **do**

    if *n not labeled* **then**

      compute t = neighbors of n;

      compute t_l = labeled neighbors of n;

      if $|t_l| + l \geq t$ **then**

        initialize *score*, c

        for *each* $t_i \in t$ **do**

          score += label $t_i * e_{nt_i}$

          c += e_{nt_i}

        end

        update label $n = score/c$

      end

    end

  end

end

Table 2.1: Table of top 10 words (excluding hashtags) used by Disbelievers and Believers.

Believers	**Disbelievers**
climate	climate
change	change
world	global
US	private
need	US
action	UN
UN	Sanders
global	world
leaders	end
future	warming

Table 2.2: Table of top 10 words (excluding hashtags) used by Disbelievers and Believers.

Disbelievers			**Believers**		
climate	climate	yellowvests	science	climatechangeisreal	cop
scam	bernie	maga	climate	climatechange	climateaction
change	sanders	trump	year	auspol	climate
nuclear	travel	carbontax	keepitintheground	climatestrike	katowice
industry	potus	france	climateemergency	climateactionnow	world
fakenews	month	macron	record	greennewdeal	takeyourseat
crisis	change	policy	renewableenergy	climate	leader
record	planet	french	bcpoli	cdnpoli	solar
global	great	people	end	stopadani	change
agw	face	hoax	fact	globalwarming	poland

The weighted in-degree centrality gives us a measure of the reach of individual user accounts, either due to their direct following or due to followership of those who retweet or mention that account. The normalized values as calculated in [158] for the top 5 accounts are report-ed. Disbelievers top 4 accounts have much higher normalized value, meaning that these Disbelievers accounts reach a higher fraction of other Disbeliever accounts than the top 4 accounts of Believers reaching other Believers. We also calculated PageRank centrality [120] for both the networks. PageRank centrality ranks a node, a user account in our case, higher based on the importance of incoming nodes. In this centrality measure, we do not see much difference between the two networks.

Table 2.5: Normalized centrality measure values for top 5 Believers and Disbelievers user accounts by the respective measure for sum of mentioned-by and retweeted-by network.

Weighted Total Degree Centrality		**Weighted In-Degree Centrality**		**Page Rank Centrality**	
Believers	**Disbelievers**	**Believers**	**Disbelievers**	**Believers**	**Disbelievers**
0.003	0.001	1	1	0.132	0.136
0.001	8.40*e^{-4}	0.563	0.978	0.076	0.042
0.001	3.12*e^{-4}	0.536	0.929	0.038	0.029
0.001	2.82*e^{-4}	0.480	0.731	0.028	0.028
0.001	2.73*e^{-4}	0.284	0.279	0.024	0.024

Next, to understand most crucial users in bi-directional communication, we look at the weighted and unweighted total degree centrality of the sum of reciprocal mentioned-by and retweeted-by networks. For calculating total degree centrality, we use a similar method as used in calculating total degree centrality for the mentioned-by and retweeted-by networks. In Table 2.6, we report our results for the top 5 accounts in respective category for Disbelievers and Believers. The unweighted total degree centrality is more than the weighted degree centrality suggesting that higher fraction of users do reciprocate to top 5 accounts by mentioning or retweeting, but a lower fraction of users reciprocate more than once to top 5 users. All top 5 Disbeliever users have higher unweighted total degree centrality than Believers top 5 accounts; however, 4 out of top 5 Believers have higher weighted total degree centrality. This indicates that a higher fraction of Disbelievers reciprocate to top 5 users just once com-pared to the fraction of Believers reciprocating to top 5 Believers, but a lower fraction of Disbelievers reciprocate more than once compared to Believers. We conclude that Disbelievers have higher fraction of influential users in the network compared to Believers.

Table 2.6: Normalized centrality measure values for top 5 Believers and Disbelievers user accounts by the respective measure for sum of reciprocal mentioned-by and retweeted-by network.

Un-Weighted Total Degree Centrality		**Weighted Total Degree Centrality**	
Believers	Disbelievers	Believers	Disbelievers
0.003	0.009	6.76e-04	0.001
0.003	0.006	5.78e-04	3.88e-04
0.003	0.005	3.91e-04	3.82e-04
0.002	0.004	3.58e-04	1.79e-04
0.002	0.004	3.42e-04	1.49e-04

News Agency and Bot-like Accounts Behavior We look at the popular news sources within our different groups. In figure 2.4, we present a word cloud of the names of accounts classified as news agency (§2.2.3) by the number of tweets in the competing groups. "Patriot News" dominates Disbelievers' tweets (including retweets), but for Believers there is no one account which dominates.

Bot Activity Next, we compare bot-like activity in the two groups of Believers and Disbelievers. In figure 2.5, we report the bot-like account's activity at different probability thresholds

Figure 2.4: Word cloud of tweets by news agencies classified as Disbeliever (left) and Believer (right).

for an account to be classified to be bot-like for the Believers and Disbelievers. We find that the fraction of tweets and user accounts classified as bots are similar for both the groups at all threshold levels. This indicates that bots are similarly active in both the groups.

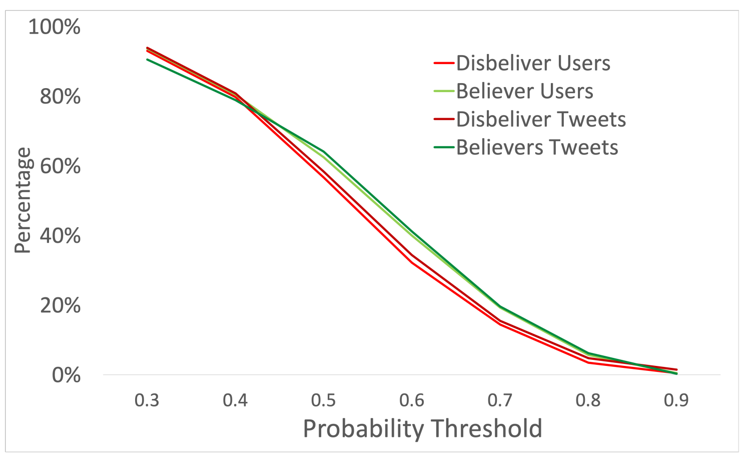


Figure 2.5: Percentage of bots and tweets at different probability threshold for an account to be classified to be bot-like as predicted by Bot-Hunter [18] for climate Disbelievers and Believers group.

Fake News, Conspiracies and Exaggerations Using the corpus of stories from our unigram and bigram search in §2.2.4, we manually checked each story to find fake news, conspiracies and exaggerations about climate change effects. We did not find fake news related to climate change listed in FactCheck.org, Politico, truthorfiction and hoax-slayer. We did find multiple fake stories related to conspiracies and exaggerations about climate change effects.

With regard to climate change, politicians and others have been vocal about their criticism of science, even using conspiracy theories as possible explanations [151]. This makes the study of conspiracy theories in climate change context even more relevant. To this end, we look at the most talked about conspiracy theories in our data set. In figure 2.6, we report the number of unigrams and bigrams related to conspiracies found in tweets and retweets. User accounts tweet more stories containing conspiracy theory phases than they re-tweet. Conspiracy theory regarding QAnon, which is a deep state conspiracy theory originating from 4chan [160], is the most popular conspiracy theory in our data set.

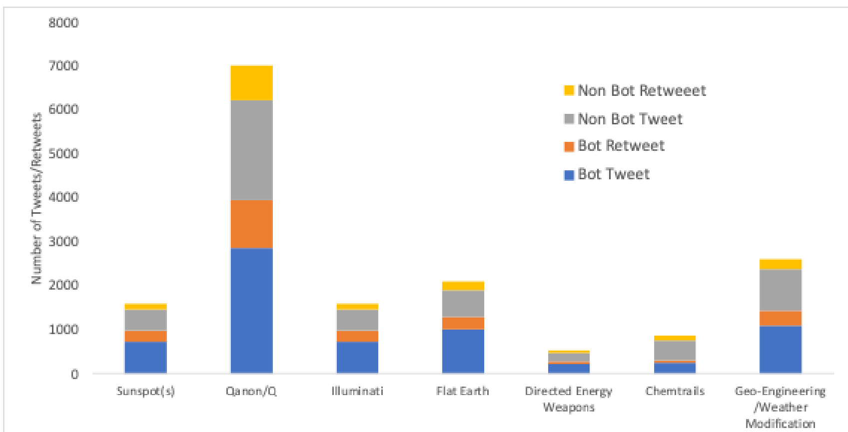


Figure 2.6: Number of tweets and retweets by accounts classified as bots and non-bots containing unigrams and bigrams related to conspiracy theories.

2.4 Discussion

Understanding the conversations and underlying beliefs of segmented groups helps in understanding the constructs by which more people could be attracted or repelled by different messaging. We find that Disbelievers words usage focus more on attacking personalities believing in anthropogenic origin of climate change and their messaging; on the other hand, Believers words usage focuses on callings to combat climate change. Climate change messaging by personalities believing in anthropogenic origin of climate change appears unlikely to change the opinion of Disbelievers but rather be used by Disbelievers as a conversation topic to target Believers. Our results indicate that unlike conversations on personalities in opposing groups, messages about social movements are dominated by discussion on movements aligned towards group's beliefs rather than calling out movements driven by contrasting beliefs. Moreover, an important finding of this paper is that different communities primarily use different sets of hashtags. To bridge communities it is important to use hashtags with neutral literal meaning.

We looked at the network structure of Believers and Disbelievers for these twitter interactions. We found greater homophily in retweet, reply and reciprocal networks compared to the mention network. This is consistent with the fact that typically users retweet, equivalent to resharing, if they endorse that message and are hence more likely to endorse a message from users aligned to their own perspective. On the other hand, mentioning activity could be a way to call out or malign members of other community. For all the types of networks, we found that both Disbelievers and Believers talk within their group more than with the other group; this is more so the case for Disbelievers than for Believers. Our results confirm findings from [162], which concluded that there are segregated communities in climate change conversations on Twitter. We conclude that users in their respective groups are involved in discussions mostly within their group, and previous studies have suggested that such "echo-chambers" could lead to extreme viewpoints [144]. Much of the earlier work in climate change communication such as [10] focusses on the drivers of climate change discourse online. In this paper we focused on how influential ac-counts frame discourse online. We looked at the central figures in these networks for both the groups. Valente [154] suggested that node centrality based on follower's network can be used to identify opinion leaders. However, centrality by follower net-work may not be useful in measuring whether or not the followers of a central figure ascribe to the views of the central figure in the case of a specific event. Hence, in this paper, we look at centrality measures for the retweeted-by and mentioned-by networks as a proxy for influence. Disbelievers communication activity is influenced by higher fraction of users in their group compared to Believers. This coupled with the fact that Disbelievers are more "echo-chambery" suggests that higher fraction of conversations within Disbelievers happen with the influencers compared to Believers' network. We conclude that Disbelievers are more organized around certain influencers in their net-work compared to Believers. Further research is needed on whether reciprocity leads to denser follower's network over time in these groups as shown generally for Twitter users in [9].

We found that "Patriot News" is most active in tweeting stories in Disbelievers' group, and for Believers, there is no one news agency that dominates. Unlike Believers, Disbelievers get their news from a concentrated number of news sources and hence may be more vulnerable to manipulation. We also classified accounts into users showing bot-like behavior and users not showing such behavior. We found that in both Disbelievers and Believers bot-like accounts were equally active. This is in similar vein with previous findings that bot-like accounts tend to stir conversations in differently politically aligned groups rather than concentrating on conversations in one group [20]. We conclude that bot activity is further creating and nourishing the divide between Believers and Disbelievers. Furthermore, the collection period of our dataset is short, it would be interesting to study how bot-like activity is changing over time and which particular misinformation stories are originating from bot-like accounts.

We searched our data set to find fake news, conspiracies, and exaggerations about climate change. We found multiple tweets with conspiracies mostly from the Disbeliever group. People believing in conspiracy theories are more likely to believe that a conspiracy theory is a possible explanation of climate change [151]. Hence, conspiracy theories could be used as a potential re-cruitment tool by Disbeliever lobbyists. We also found exaggerations made by Believers which were targeted by Disbelievers to strengthen their argument. This shows the effect of excessive attention of media reporting on outliers in climate change discussion as claimed in study done by Baykoff [27]. It would be interesting to find whether these exaggerations are used by Disbelievers to recruit more members and create confusion within Believers over time. In Disbelievers and Believers, a considerable number of original tweets were used to negatively attack members of the opposite group. This is consistent with network theory that suggests that a way of identification for group members is by negatively engaging with another group members [146]. Furthermore, this behavior leads to more inward communication within the groups and reduction in communication outside the group, as users tend to ignore consistent negative criticism from users who don't share their beliefs. In such scenario, climate change messaging should not attack Disbelievers especially the influencers within the Disbeliever group.

In this paper, we look at the conversations happening on Twitter during COP 24. We classify Twitter account users into Believers and Disbelievers by the hashtags of the user tweets or retweets. We find about seven times the number of Believers compared to Disbelievers. We manually checked each Disbeliever and randomly sampled a large number of Believers to confirm our findings. Since the aim of this research was to find online communities with different views on climate change and characterize the conversations within these communities, we used a range of hashtags and keywords related to climate change to collect our data. Nevertheless, our collection method gives us a large corpus of tweets but is not able to capture debates on climate change that does not use these keywords or hashtags. Moreover, collecting tweets using keywords or hashtags tend to miss out replies, as replies may not contain same keywords or hashtags as the original tweet.

2.5 Appendix

2.5.1 Unigrams and Bigrams used to Identify Fake News and Conspiracies

We report the set of tokens used in identifying fake news and conspiracies in Table 2.7. We searched for these tokens in unigrams and bigrams made for each tweet. Tokens such as "World End" were selected to find exaggerated claims about doomsday scenario of climate change.

Table 2.7: Unigram and Bigram search terms used for searching exaggerations and fake stories. Capitalization is done for better readability.

Unigrams	**Bigrams**
Fake, #Fake, FakeNews, #FakeNews	Fake News, #Fake News
WorldEnd, SocietyEnd, EndSociety	World End, Society End, End of, #Society
#SocietyEnd, #EndSociety	End, #End Society
Alarmist(s), Alarmism, #Alarmist	Liberal Society, #Liberal Society
LiberalSociety, #LiberalSociety	
GlobalCooling	Global Cooling, #Global Cooling
Tree(s)Killed, #Tree(s)Killed	Tree(s) Killed, Animal(s) Killed
Animal(s)Killed, #Animal(s)Killed	#Animal(s) Killed, Population End
Rapidly, urgently, quickly	Very Soon, judgement day
judgement-day, Biblical	Politicizing Science, Climate Emergency
ClimateEmergency	Climate Totalitarianism
Sunspot(s), #Sunspot(s), Qanon, #Q	Sunspot(s) Activity, Deep State
DeepState, Soros,#Soros	Directed Energy, Energy Weapons
Pizzagate,#Pizzagate, Rothschild	Club Rome, Weather Modification
Pertodollar, DEW, #DEW	#Weather Modification, Geo Engineering
ClubofRome, Chemtrails,#ChemTrails	#Geo Engineering, #Flat Earth
WeatherModification, #WeatherModification	Flat Earth, Planet X, Planet Niburu
GeoEngineering, #GeoEngineering	
Illuminati, #Illuminati	
FlatEarth, #FlatEarth, PlanetX, Niburu	

Chapter 3

Affective Polarization

3.1 Introduction

Online social networks represent a powerful space for public discourse. Through large-scale, interconnected platforms like social media, diverse communities may potentially participate in open exchanges of views and information about a vast range of issues. However, research has increasingly demonstrated the dangers of *polarization* in online communication [17, 147, 149]. Attributed to various psychological, social, and technological factors, intergroup communication on cyberspace has displayed tendencies to feature pathological dynamics especially concerning contentious issues [66, 164]. Opposed groups may communicate in a highly balkanized fashion, such that members of an in-group are only minimally exposed to out-group members and their beliefs [64, 94]. This phenomenon has been termed *interactional polarization*. Polarization can also pertain to highly negative sentiments toward out-groups in the form of *affective polarization* [6, 54]. Social scientific research examines how these phenomena are interconnected across a variety of contexts, such that online groups that disagree on a given topic are also more likely to be hostile toward each other [164]. In this paper, we focus on quantifying affective polarization between two groups with opposing beliefs using Twitter discourse on a significant social issue.

One significant issue which has received heated attention in online public discourse is climate change [56, 62, 148]. We focus on those who cognitively accept anthropogenic causes of climate change (*Believers*) and those who reject the same (*Disbelievers*). Previous work demonstrates not only sharp divergences in climate change beliefs but also the emergence interactionally polarized groups [70, 112, 148]. In other words, online discussions about climate change are *interactionally polarized*, implying the persistence of echo chambers between Believers and Disbelievers [86, 156, 162]. Much less work, however, engages the question of *affective polarization* in online climate change discourse. A crucial limitation in prior work lies in the methodological options available to past researchers. Relying consistently on manually annotated corpora and datasets of limited size, existing scholarship has faced barriers to measuring the emotional component of climate change discussions in a generalizable fashion [6, 86, 156]. Drawing on recent advances in computational stance detection, targeted sentiment analysis, and network science measures, we present an integrated methodological pipeline for addressing this gap in the literature. [81, 101].

This work leverages computational methods to generate (a) automated stance labels for climate change Believers and Disbelievers, (b) individual measurements of the interaction valence between in-group and out-group members, and (c) broader assessments of group-level affective polarization. We demonstrate the utility of our framework by applying our methodology to a large-scale dataset of 100 weeks of online climate change discussion on Twitter. Furthermore, we link our findings to natural disasters words to explain important climate change belief constructs.

In sum, we probe the following research questions:

1. How can affective polarization be measured on a large-scale online conversation about climate change?
2. Do climate change Believers or Disbelievers feature greater levels of affective polarization?

3. What is the relationship of affective polarization with use of natural disaster related words?[1]

The subsequent sections of this paper are organized as follows. First, we provide an overview of related work in this area, illustrating computational analysis of polarization in general terms and then in the case of climate change specifically. We zero in on the dearth of principled empirical work on affective polarization specifically in relation to online climate change discourse. Second, we present our proposed methodological pipeline which integrates machine learning models and network science techniques to facilitate a novel and effective framework for assessing affective polarization. Third, we share our findings on our large-scale, long-term Twitter dataset. Last, we discuss implications for understanding the state of climate change discourse on digital platforms as well as related empirical investigation of affective polarization on online social networks.

3.2 Related Work

3.2.1 Computational analysis of polarization

Recognizing the ubiquity of online conflicts, rigorous scholarship in the computational and social sciences has tackled the problem of polarization. More traditional approaches in offline settings have relied on survey measures to empirically assess divergence in beliefs between groups [16, 54]. But with burgeoning developments in computational methods - especially with respect to natural language processing and machine learning - automated methods have also arisen to leverage the vast digital traces linked to online activity [81, 102].

General approaches to studying polarization infer individual attitudes from user information, such as the texts associated with an account on social media (e.g., Facebook comments,

[1]We provide the list of natural disaster related words in our project repository: `https://github.com/amantyag/affectiveclimatechange`

tweets). Group membership as well as group communication are similarly incorporated into analyses of polarization, by examining the beliefs of individuals in conjunction with their traceable patterns of digital interaction with other individuals. Given various conceptualizations of polarization, different frameworks have been developed to quantify pathological patterns of communication across groups holding similar or opposed stances on a given issue [45, 48, 115, 159]. Representing online conversations as graphical structures, Social network approaches typically measure polarization in terms of a function of homophily in local community structures [107, 143]. In other words, the extent to which the likelihood that those holding similar views interact with each other - in contrast to those with whom they disagree. For example, one may quantify the probability of a random walk starting from a node belonging to a given stance group ending up in a node belonging to the same or a different stance group [64, 149, 164]. More recent scholarship, however, emphasizes the importance of examining not just pathologically isolated communication, or interactional polarization; but also pathologically hostile communication, or affective polarization. Burgeoning evidence suggests that echo chambers represent an incomplete picture of polarization [17]. People holding opposed views, in fact, do interact with each other - but this does not necessarily mitigate polarization [94]. Instead, research finds that intergroup exposures trigger further incivility [6]. Hence, reliable measures for affective polarization are needed, although the computational literature in this area remains in its nascent stages [164].

3.2.2 Climate change and polarization

In the specific case of climate change discourse, analysis of polarization has also represented a major research topic. Numerous Numerous studies link polarized beliefs about climate change to partisan divides, with more conservative individuals less likely to cognitively accept anthropogenic climate change than liberals [56, 70]. Past work specifically demonstrates that although higher levels of education and information access may increase the likelihood of climate change belief, these effects remain much lower among conservatives [70, 110]. Such effects have been explained from the lens of elite signalling - whereby followers emulate the beliefs of their preferred political leaders - uneven exposure to information based on partisan media, as well as a generalized dislike for the members of the opposed ideological group [24, 39, 155]. However, with time, scholars have also noted general trends toward increasing climate change beliefs overall [112]. Even if these do not necessarily translate into concrete support for policy [62], the long-term instability of skepticism points to valuable ways forward for science communication [87]. Collectively, these finding suggests the importance of accounting for the psychological processes surrounding climate change belief and disbelief [93].

These issues take on specific forms in cyberspace, where information flows are inextricably entangled with community dynamics. On social media, studies suggest employing social network analysis have uncovered robust evidence that online climate change discussions tend to exhibit echo chamber-like interactions [148, 162]. Qualitative analysis further showed that in rare instances of intergroup communication, more negative frames prevailed, featuring dismissal of climate change as a hoax, identity-based derailment of conversations, as well as overall higher levels of incivility [6, 156]. Notwithstanding the valuable idiographic insights derived from these studies, Existing studies, however, rely on a minuscule fraction of the larger conversation to facilitate in-depth content analysis. Hence, larger-scale and more generalizable findings on the affective dynamics of online climate change discourse are notably lacking in the literature.

3.2.3 Contributions of this work

Motivated by the foregoing insights, this work seeks to contribute to the literature by offering a methodological pipeline for examining affective polarization. As the succeeding sections demonstrate, our framework combines machine learning and network science methods in a novel, scalable, and generalizable fashion for ready application in a variety of contentious issues. This overcomes methodological barriers present in prior work, including their common reliance on expensive survey or experimental measures, or manually annotated datasets in the context of social media research on climate change discourse [70, 112, 162].

From a theoretical standpoint, we additionally contribute a nuanced operationalization of affective polarization as located on a group level. We unpack how group-level metrics valuably produce asymmetrical views of hostile behavior, thereby facilitating more fine-grained analysis of how different stance groups engage in varied levels of affectively polarized interactions. This conceptually aligns with the asymmetry of psychological factors characterizing climate change Believers and Disbelievers, especially over time [56, 87, 155]. Finally, on an empirical level, our work also extends prevailing scholarship on polarized climate change discourse. While established findings paint a picture of consistent echo chambers between climate change Believers and Disbelievers, we provide evidence for the flipside of these dynamics. We We specifically quantify, over a larger-scale and longer-term dataset than previously examined in prior work, the extent to which intergroup interactions systematically feature hostility. This may inform possible data-driven interventions for policymaking beyond more prevalent frames of intergroup contact and science communication [93].

3.3 Data and Methods

3.3.1 Data collection

We collected tweets using Twitter's standard API[2] with keywords "Climate Change", "#ActOnClimate", "#ClimateChange". Our dataset was collected between August 26th, 2017 to September 14th, 2019. Due to server errors, the collection was paused from April 7th, 2018 to May 21st, 2018, and again from May 12th, 2019 to May 16th, 2019. We ignore these periods in our analysis. After deduplicating tweets, our dataset consisted of 38M unique tweets and retweets

[2] `https://developer.Twitter.com/en/docs/tweets/search/overview/standard`

from 7M unique users. For our analysis, we aggregate tweets from each user for seven day period (1 week) to get a total of 100 weeks.

3.3.2 Stance labels

We use a state-of-the-art stance mining method [102] to label each user as a climate change Disbeliever or Believer. We use a weak supervision based machine learning model to label the users in our dataset. The model uses a co-training approach with label propagation and text-classification. The model requires a set of seed hashtags essentially being used by Believers and Disbelievers. The model then labels seed users based on the hashtags used at the end of the tweet. Using the seed users, the model trains a text classifier and uses a combined user-retweet and user-hashtag network to propagate labels. In an iterative process, the model then labels users who are assigned a label by both methods with high confidence.

We set *ClimateChangeIsReal* and *SavetheEarth* as Believers seed hashtags and *ClimateHoax* and *Qanon* as Disbelievers seed hashtags. These hashtags have been shown to be used mostly by the respective groups[148]. The algorithm labels 3.9M as Believers and 3.1M as Disbelievers. We provide details of manual validation of stance results and the parameters in our project repository `https://github.com/amantyag/affectiveclimatechange`.

3.3.3 Affective polarization metrics

We measure affective polarization in this work by combining outputs from an aspect-level sentiment model, a classic network science measure known as the E/I index [101] and Earth Mover's Distance (EMD) [76].

Aspect-level sentiment

Aspect-level sentiment refers to the emotional valence of a given utterance toward one of the concepts it mentions. Sentiments toward specific entities are vital to consider in polarized discussions such as those we consider here. For instance, climate change Disbelievers might express negative feelings toward notions of greenhouse gases, while in agreement with a fellow Disbelievers with whom they are interacting. We utilize Netmapper to extract entities from each tweet, and predict the aspect-level sentiment of each tweet toward each entity [38]. Word-level sentiment is computed based on the average of known valences for surrounding words within a sliding window. For the purposes of this work, each tweet by a certain agent i which mentions or replies to agent j is assigned an aspect-level sentiment score from -1 (very negative) to $+1$ (very positive) directed toward the concept "@[agent j]". This allowed us to compute affective dimensions to the communication between groups of the same or opposed stance groups.

Affective networks

Let $G^+ = (V, E^+)$ denote a positive interaction network where the set of vertices V contains all Twitter accounts in our dataset and the set of directed edges E^+ contains all positive-valenced mentions and replies between agents in V. Similarly, let $G^- = (V, E^-)$ denote a negative interaction network over the same set of agents V and the set of directed edges E^- representing their negative-valenced mentions and replies. Let S_{ij} denote the set of all aspect-level sentiments in tweets by agent i toward agent j, where $i, j \in V$. Then the weight w_{ij}^+ of edge $e_{ij}^+ \in E^+$ from i to j is given by $\sum_{x \in S_{ij}} \min(0, x)$. Conversely, the weight w_{ij}^- of edge $e_{ij}^- \in E^-$ from i to j is given by $\sum_{x \in S_{ij}} \min(0, -x)$.

E/I indices

We assess group-level differences in positive and negative interactions using Krackhardt's E/I index [101]. For a given affective network, the E/I index intuitively captures the extent to which each stance group k engages in correspondingly valenced interactions with members of the out-group relative to their in-group [152]. Hence, for instance, high values of the E/I index for the negative interaction network would indicate that the given stance group interacts in a more negative way to their opponents relative to those who share their beliefs. To compute the E/I indices, let $V_k \subseteq V$ denote the set of agents belonging to stance k and $V_{k'}$ those who do not hold stance k. The E/I index of stance group k on the positive interaction network is therefore computed as follows:

$$P_k^+ = \frac{E_k^+ - I_k^+}{E_k^+ + I_k^+} \tag{3.1}$$

where $E_k^+ = \sum_{i \in V_k, j \in V_{k'}} w_{ij}^+$ and $I_k^+ = \sum_{i,j \in V_k} w_{ij}^+$. On the other hand, the E/I index of stance group k on the negative interaction network is similarly computed thus:

$$P_k^- = \frac{E_k^- - I_k^-}{E_k^- + I_k^-} \tag{3.2}$$

where $E_k^- = \sum_{i \in V_k, j \in V_{k'}} w_{ij}^-$ and $I_k^- = \sum_{i,j \in V_k} w_{ij}^-$. Given the construction of P_k^+ and P_k^-, we note that both values are bounded between -1 and $+1$.

Polarization valence

We find whether the interactions have negative valence or positive valence by defining polarization P_k as expressed below:

$$P_k = P_k^- - P_k^+. \tag{3.3}$$

In this work, we operationalize our view of affective polarization in terms of high E/I indices on the negative interaction network, and low values on the positive interaction network. P_k assigns positive values for groups that display disproportionately hostile or negative interactions toward the out-group relative to their in-group. Values close to 0, on the other hand, indicate relatively even levels of positive and negative interactions. Finally, negative values indicate that those holding stance k are more negative to their in-group but positive to their out-group. Values of P_k are usefully bounded between -1 and $+1$ for ease of interpretability and comparison over different networks and time periods.

Polarization magnitude

To find the magnitude of affective polarization we use Earth Mover's Distance (EMD) on the distribution of weighted edges for outgroup and ingroup interactions. This is similar to computing first Wasserstein distance between two 1D distributions[127]. Similar to affective networks, we define $G = (V, E)$ as interaction network where the set of vertices V contains all Twitter accounts in our dataset and the set of directed edges E contains all valenced (positive or negative) mentions and replies between agents in V. In this case, we do not separate negative and positive valence graphs and treat weight w_{ij} of edge $e_{ij} \in E$ from i to j as given by $\sum_{x \in S_{ij}} x$. Let u_k be distribution of w_{ij}, where $i \in V_k, j \in V_{k'}$ and let v_k be distribution of w_{ij}, where $i \in V_k, j \in V_k$. For a group holding stance k, we define our novel affective polarization metric as:

$$l_k = \begin{cases} -\int_{-\infty}^{+\infty} |U_k - V_k| & : P_k < 0 \\ \int_{-\infty}^{+\infty} |U_k - V_k| & : P_k \geq 0 \end{cases} \tag{3.4}$$

where U_k and V_k are the respective CDFs of u_k and v_k. Here, EMD is proportional to the minimum amount of work required to covert one distribution to another. [3]. We use P_k to assign positive or negative valence to the EMD. Although there are other techniques to find the difference in distribution such as KS-Test [106]. However, during our experiments, we found that EMD is able to capture more nuanced differences in distributions. More likely because the EMD can capture differences in heavy-tailed distributions better and it does not make any parametric assumptions [127].

Our novel affective polarization metric l_k is positive when $P_k > 0$. As noted in §3.3.3, a positive value would mean more hostility or negative sentiment in intergroup communication compared to intragroup communication. On the other hand, a negative value of l_k is when $P_k < 0$, meaning more positive sentiment in intergroup communication compared to intra-

[3] `http://infolab.stanford.edu/pub/cstr/reports/cs/tr/99/1620/CS-TR-99-1620.ch4.pdf`

group communication.

3.4 Results

Using the metric defined in Equation 3.4, in this section, we first explore how affective polarization between Believers and Disbelievers is changing over the 100 weeks. Then we explore how hostile periods are related to natural disaster-related words.

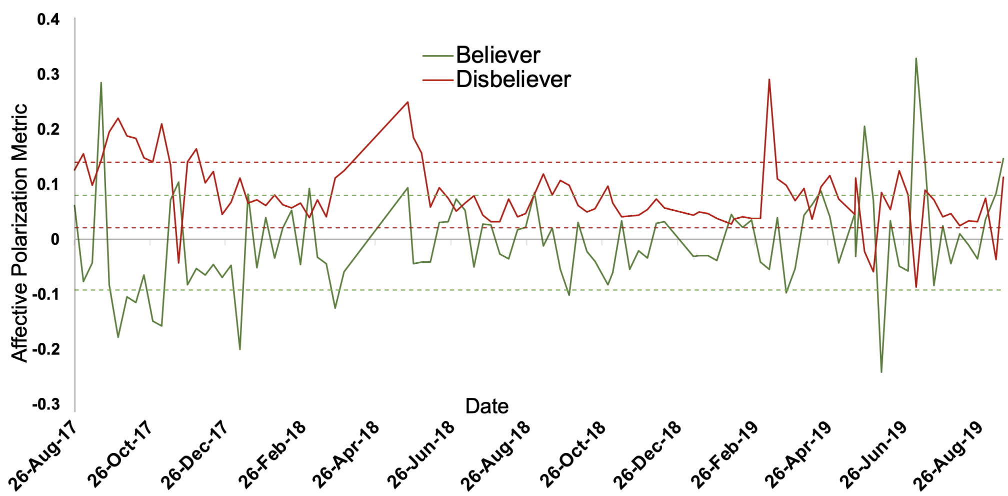


Figure 3.1: Affective polarization metric (l_k) for Believers and Disbelievers of climate change. Higher positive values denote more hostility towards the other group. The dotted lines represent mean ±1 standard deviation, which for Believers is -0.091 and 0.080 and disbelievers is -0.117 and 0.106. The analysis was done on data collected from 26th August 2017 to 14th September 2019 as described in §3.3.1.

We first look at how the affective polarization metric is changing over time in figure 3.1. Overall, our analysis found that climate change Disbelievers tended to exhibit high levels of hostility toward climate change Believers. This finding was relatively consistent throughout the 100-week period under observation, as the time series for climate change Disbelievers only very rarely goes below the threshold of 0, which indicates similarly valenced interactions toward in-group and out-group members. Some weeks displayed exceptionally high levels of hostility toward climate change Believers, greater than one standard deviation from the mean. The standard deviation of l_k is lower for Disbelievers than for Believers. Indicating that Disbelievers act in much more organized manner over the 100 weeks than Beleivers. Climate change Believers, on the other hand, were not generally hostile toward Disbelievers, as the time series for climate change Believers tends to fluctuate over and under the threshold of 0. This indicates that climate change Believers communicate with in-group and out-group members with relatively similar emotional valence. However, on certain weeks, climate change Believers did also feature exceptionally high hostility scores. This suggests that climate change Believers may also behave in a hostile manner toward climate change Disbelievers, even if not over the long term.

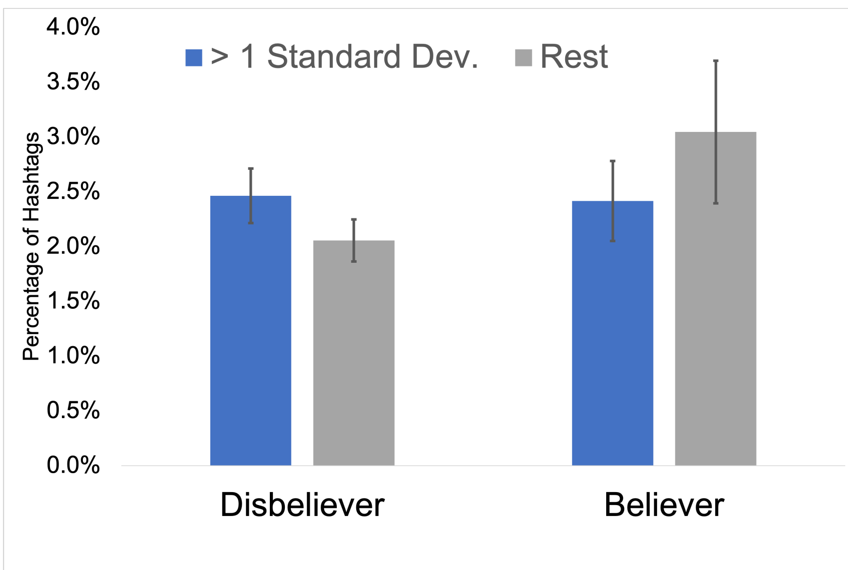


Figure 3.2: Percentage of the top 100 most frequent hashtags containing natural disaster-related words. The figure shows the percentage when the affective polarization metric is greater than 1 standard deviation or otherwise. The error bars represent ±1 standard errors.

To investigate instances where hostility between Believers and Disbelievers is high we compare those weeks with weeks where hostility is low. We define hostile weeks as those data points where l_k is more than mean plus 1 standard deviation, i.e. from figure 3.1, all the weeks where for Believers $l_k > 0.080$ and for Disbelievers $l_k > 0.140$. The number of such weeks for Disbelievers where $l_k > 0.140$ is 20 and for Believers where $l_k > 0.080$ is 12. We look further into these weeks as examples of exceptional hostilie weeks.

Next, we use natural disaster-related words as a proxy to determine how natural disasters play a role in hostility between the two groups. In figure 3.2 we look at the top 100 most frequent hashtags used within those groups to find the percentage of hashtags related to natural disasters. As expected, Believers use more natural disaster-related hashtags than Disbelievers. However, during the exceptional hostile weeks Believers use less of these hashtags (p ¡ 0.05). Interestingly, Disbelievers show the exact opposite behavior. Disbelievers use more natural disaster-related hashtags when they are more hostile towards Believers. We provide further evidence of this finding in figure 3.3. In figure 3.3, we look at the percentage of Tweets with at least one natural disaster-related word. We find similar patterns as mentioned above (p ¡ 0.05). Moreover, we find that a greater percentage of Tweets from Disbelievers mention natural disaster-related words compared to Believers. This indicates that Disbelievers are calling out natural disasters more when they are exceptionally hostile towards Believers compared to other weeks.

3.5 Discussion

Taken together, our findings suggest the importance of considering affective polarization in online discourse, particularly concerning the subject of climate change. Whereas past studies had shed light on the echo chamber dynamics which characterized intergroup communication surrounding climate change [162], we show how this polarization extends also to the realm of emotion in the form of affective polarization. We extend existing studies which highlight the role of incivility and personalized framing in encounters between climate change Believers and Disbelievers [6, 156] by introducing a scalable technique for analyzing relative intra- and

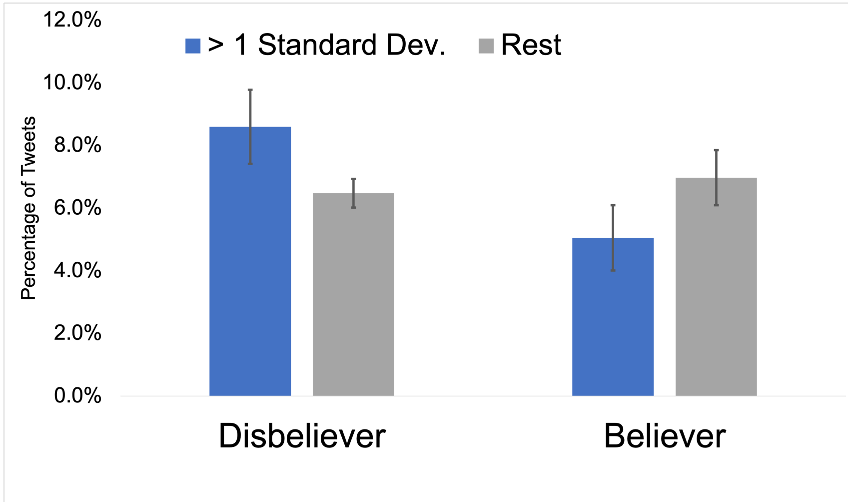


Figure 3.3: Percentage of tweets with at least one natural disaster-related word. The figure shows the percentage when the affective polarization metric is greater than 1 standard deviation or otherwise. The error bars represent ±1 standard errors.

intergroup interaction valence. This allowed us to quantify the extent of hostile communications between the two groups over a large-scale, long-term dataset - thereby validating existing findings in a generalizable manner as well as showing their relative stability over time.

Furthermore, we highlight the value of viewing polarization from an asymmetrical perspective. Related scholarship in political psychology underscores how ideological asymmetries underpin conflict dynamics across a variety of social issues [92]. In other words, the participation of two groups within polarized discourse does not necessarily mean that both groups engage in conflict in the same way. Prior work illustrates that these findings translate robustly to the digital sphere - political elites or opinion leaders who share moralized content behave in distinct ways depending on their ideological orientations [29]. The present work contributes to the literature by showing how these dynamics unfold the standpoint of the public at large concerning online climate change discourse. Indeed, higher levels of hostility from Disbelievers present a specifically notable finding for social scientific scholarship on climate change discourse. Longitudinal analysis in prior work suggests that generalized climate change beliefs over time are increasing [70, 112], and climate change Disbelievers in particular are more susceptible to potential belief change [87]. But significant cognitive barriers remain for fuller acceptance of anthropogenic causes for climate change and the corresponding urgency for responsive policy changes [56, 62]. Higher levels of hostility among climate change Disbelievers toward climate change Believers constitutes one such obstacle for further dialogue between the two groups. As past studies suggest, one psychological factor which impedes climate change Beliefs is not related to the climate at all, but anchors primarily on the feelings of dislike felt by one group towards the other [155].

Such challenges may thus persist in the form of further entrenchment of Disbelievers within interactional siloes and disengagement from intergroup communication altogether [162]. Or as emergent studies show, they can also trigger what have been called 'trench warfare dynamics' [94] - whereby Disbelievers persistently communicate with Believers but solidify their own cognitive immovability in the process.

These insights are especially important to consider given our secondary set of findings. Our analysis suggests that further asymmetries arise between Believers and Disbelievers engagement with disaster words in relation to their levels of affective polarization. Although comparable levels are seen when both groups are within average levels of our metric, moments of increased affective polarization correlate with opposite behaviors for Believers and Disbelievers. Believers appear to shift to other areas of contention, such that their aggression is characterized by non-disaster topics. In contrast, Disbelievers' increased invocation of disaster terms points to more aggressive discussion of these catastrophes, albeit positioned in resistance to explanations related to anthropogenic climate change. This introduces another layer of intractable conflict in beliefs, as major climate events do not appear to invite susceptibility of belief change for Disbelievers. Instead, they potentially incite more vigorous psychological resistance. In this paper, we conduct our analysis using different techniques as described in §3.3. We make an affective network using the sentiment of a person towards another person. Here, we develop aggregate network values over different weeks. Sentiment analysis is a non-trivial task; the network value's uncertainty would decrease as the number of communication links decreases. As the number of communication links is very high for our aggregated per week dataset, we believe it can capture different users' mean sentiments. Further detailed analysis is needed to remove the uncertain data points. Hence, our conclusions are based on average metric and higher than one standard deviation value of our novel index (l_k).

Collectively, these findings point to significant benefits to studying affective polarization in online climate change discourse. Although social media discourse does not necessarily constitute a representative sample of a particular global population [116], digital platforms like Twitter nonetheless constitute a vital space for public conversations about important issues like climate change. Hence, these findings paint a useful picture of public discourse as situated specifically in cyberspace, which may also bear implications for how digitally mediated science communication and public policy may also be designed and implemented [24, 93].

Besides the issue of demographic representativeness for online data, other limitations attend the present analysis. First, although we have a large number of tweets to characterize general affective behavior, however, it does not encompass those interactions which do not include our collection keywords. Second, the task of getting an aspect-level sentiment of each tweet towards other entities is a non-trivial task. We use Netmapper which has been used with reasonable accuracy for multiple sentiment level tasks [152, 153]. The focus of this paper is on designing a framework to get affective polarization score between two competing groups and we do not make an effort to improve aspect-level sentiment scores. Last, in our analysis we use a list of natural disaster related words. Communication about the natural disasters could also happen using specific names related to these disasters, for example using "Dorian" instead of "Hurricane Dorian". Such analysis would require a more comprehensive list of natural disasters occurring around the world during the 100 weeks. This is out of scope for the current work.

Recognizing the foregoing limitations, we also consider avenues for future work in this area. On a conceptual level, researchers may wish to expand the binary system of climate change beliefs assumed here. Affectively polarized dynamics between multiple groups may be a more challenging yet also potentially informative line of inquiry to explore given the diversity of positions held with respect to this complex issue. Acknowledging the non-neutrality of cyberspace, it would also be important to consider whether disinformation maneuvers may also be involved in shaping the wider climate change discussion. Inauthentic bot-like accounts and trolls may unduly influence different groups by manipulating the flow of information or amplifying intergroup aggressions; such factors have been seen in relation to other contentious issues and may potentially be present here as well [152]. Methodologically, computational analysis may extend our findings by performing more fine-grained characterization of the types of hostility expressed by both groups. Natural language processing (e.g., topic models) may offer one way forward in this regard.

Finally, taking flight from the digital scope of our research, further studies may fruitfully examine several hypotheses opened up by our results. For instance, social scientists may investigate actual levels of experienced hostility by climate change Believers and Disbelievers toward opposed groups. These evidence bases would be valuable to accumulate in cross-cultural settings, as well as over time - especially in connection with concurrent political shifts and natural climate-related developments like anomalous weather patterns and wider-ranging disasters [70, 112].

Chapter 4

Media Framing

4.1 Introduction

The interpretation of information depends on how that information is presented. For example, a particular topic could be presented in the news in such a way such that specific points are emphasized or de-emphasized to create confusion on scientific facts [40, 57, 103, 134]. Such a change in information distribution can significantly impact public opinion on important policy and scientific issues [40, 119].

One way to analyze how information is manipulated is by studying *frames* of the presented information, where framing presents certain information in a manner that emphasizes one issue over another. For example, news on Californian wildfires can be framed either as a natural event causing destruction of property or a human-made disaster causing socio-economic harm. More generally, framing is defined as "selecting certain aspects of a given issue and making them more salient in communication in order to 'frame' the issue in a specific way" [134].

Different approaches have been proposed in Natural Language Processing (NLP) / linguistics research to analyze frames. These approaches are broadly divided into formal/stylistic frames or content-oriented frame [134]. Formal/Stylistic frames concentrate on the structure or formal presentation of text rather than the content (e.g. Iyengar [84, 85]). Content oriented frames focus on the communicative text. Content-oriented frames can further be divided into generic frames or topical frames [47]. Topical frames are issue-specific. In NLP to analyze topical frames, we use computational models such as Latent Dirichlet Allocation (LDA)[21], Latent Semantic Analysis (LSA) [78] and more recently transformer model techniques such as Top2Vec [7]. On the other hand, generic frames are pre-defined sets of categories or patterns that transcend individual issues. For example, Semetko and Valkenburg [139] used frames such as "consequences", "responsibility", "conflict", "human interest", and "morality" on press and television news on European politics. This paper discusses and develops methods to find generic frames on news articles on a crucial socio-economic topic, i.e., climate change.

Topical frames have been used in climate change contexts (e.g. [97] and [83]). However, content framing in articles related to climate change using generic framing techniques is mostly unexplored [134]. Hence, the first research question answered in the present work is, *Which generic frames are predominant in news articles related to climate change?* We answer this research question by using news articles shared on Twitter about climate change. To investigate generic frames in climate change news articles, we discuss and develop a framework to analyze generic frames in large data using a transfer learning approach. We use different Transformer-based approaches and compare those to the approach discussed by Field et al. [61]. We use a pre-trained BERT [49] model to predict sentence level frames. For our analysis, we propose to use frames discussed in the Policy Frame Codebook [26] via a dataset annotated with these frames. The dataset is called Media Frame Corpus (MFC) [35], which is an annotated dataset of Wall Street Journal articles. The articles are annotated as per the Policy Frame Codebook's 15 frames [26] and are commonly used in multiple NLP framing analysis studies [61, 131].

Moreover, we develop a method to connect *Affect Control Theory (ACT)* with frames of news articles shared on a social network by news agencies. ACT, initially introduced by Heise [73, 74], proposes that individuals maintain their *affective* identities through their actions. The affective identities are operationalized by embedding these in Evaluation (good *vs.* bad), Potency (strong *vs.* weak), and Activity (active *vs.* passive) (EPA) space. We develop a method to embed frames in the EPA space, assuming that each frame has a particular affective meaning. Although prior work in NLP has identified ways to extract affective dimensions from pre-trained word embeddings (e.g. Field and Tsvetkov [60]), this paper discusses how we can embed and operationalize the affective dimensions of frames themselves. We then use our methodology on climate change news articles shared on Twitter[1].

Our approach of embedding the policy frames in the EPA dimension helps us determine news articles' emotional valence. In this work, we assume that a person or a news agency sharing the articles on social media with a certain frame would identify with that frame. Therefore, we discuss a mechanism where we can identify an article's emphasis (frame) in the EPA space. To the best of our knowledge, this is the first attempt at connecting computational generic framing research in NLP with ACT. By connecting frames with ACT, we draw implications for climate change communication. Hence, we find frames that are better at communicating climate change urgency as per emotional sociology. Thus, the second research question we address is, *what are the affective dimensions of the frames and which frames are more active and hence suited for communicating climate change urgency?*

Once we have found the frames' affective dimensions, we use the reshare (retweet) count of each frame to find whether each frames' emotional value or *affect* leads to more reshare. ACT states that *affect* drives individual identities and their actions. In this work, we address, *whether or not the frames' affect drive the reshare count on social media?* To answer this research question, we use the reshare count of different frames. We hence conclude which frames in climate change news articles are more likely to be reshared.

We begin by providing an overview of the framing literature and ACT in §4.2. Next, we describe our data collection (§4.3.1) and our methodology in §4.3.2. Our results (§4.4) suggest that news articles shared on social media are essentially framed using "cultural identity" frame.

[1]In Appendix 2 we discuss examples of the framing and affect values of the articles by using news articles from our dataset.

We find that MFC frames are positive and low valued in EPA space. We also find that the frames' resharing count is weakly correlated to their emotional valence. Through this research, we (1) compare different methods for framing analysis of articles using MFC, (2) develop a method to connect computational generic framing analysis with ACT, (3) show that frames with high *affect* do not necessarily drive news articles in climate change conversations on social media, and (4) present insights and implications for climate change communication framing.

4.2 Background and Related Work

4.2.1 Framing Analysis

In prior work, framing analysis is used to find bias or partisanship in news articles. Field et al. [61] use framing analysis to find media manipulation by the Russian government during economic downturns. Roy and Goldwasser [131] breaks down the policy frames [26] into more detailed sub-frames to demonstrate ideological differences between media sources. Johnson et al. [89] used weakly supervised approaches to predict frames used in political conversations on Twitter. Moreover, in a recent study, an annotation method is developed for social media framing analysis. Hartmann et al. [71] used multi-task and adversarial learning to annotate social media platforms' conversations. There has been work on predicting document and sentence level frames using MFC. Card et al. [36] predicted document level frames using a logistic regression model with latent dimensions and word-based features. This work was further improved by Ji and Smith [88], and Naderi and Hirst [117] used recurrent neural networks for sentence-level prediction of frames. In this work, we discuss BERT based novel techniques for sentence-level prediction of frames and show how framing analysis could be used in a social network setting. Apart from BERT based approaches, we use pre-trained word embeddings to decontextualize MFC corpus and predict frames on a different topic.

4.2.2 Affect Control Theory (ACT)

As framed initially by Heise [72, 73, 74], ACT was developed to explain behavior in social interaction context. Specifically, *affect* refers "to any evaluative (positive or negative) orientation toward an object" [128]. In ACT, the *affective* dimensions (EPA dimensions) could describe a persona's reaction to various situations. Each dimension in EPA space lies on the continuous interval [-4.3,+4.3]. The first dimension is "Evaluation," which describes the identity in the goodness vs. badness dimension, where a negative value indicates an identity leaning more towards bad compared to good. Similarly, "Potency" describes strong vs. week. Lastly, the "Activity" dimension describes the level of energy as active or passive. ACT theory states that it is the *affect* that we maintain during any interaction rather than a community assigned labels. For example, someone would try to maintain the affective meaning of a father ("quite good, very powerful, and somewhat lively" [128]) throughout their interactions. The affective meanings could change depending on the culture but are largely consistent. Moreover, as per ACT, a persona would minimize the deflection from its fundamental identity, reflected by its social perceptions, actions, and experiences. Social scientists codify ACT model by using a triplet of actor, behavior, and object. Each of the elements in the triplet is then measured in EPA space. In other words, each actor, behavior, and object would have an associated value in EPA space. Using the emotional signals the ACT lexicon gives, we embed the MFC frames in EPA space. This helps us understand users' perspectives while sharing the frame and connect it to wider emotional social science research.

Prior work related to ACT is rich and mostly out of scope for this paper. Therefore, we will touch upon the work which is relevant to this study [2]. Joseph et al. [91] developed methods grounded in ACT to find affective stereotypes in Twitter users who tweeted about the Michael Brown and Eric Garner tragedies. Joseph et al. [90] used ACT to predict sentiments held towards entities or behavior using a large corpus of newspaper articles. More recently, Xiang et al. [163]

[2]For more detailed discussion on ACT, please refer to Robinson et al. [128] and Heise [74]

used ACT lexicon to enhance the deep learning model for sentiment analysis. In this work, we would use a much-expanded lexicon for framing and emotional analysis of climate change news articles shared on Twitter. Moreover, we would use each article's resharing count to find which frames are more likely to be reshared.

4.3 Data and Method

In this section, we discuss our data collection and methods investigated to predict frames. Then, we discuss the evaluation of those methods. We present our method to project frames in EPA space next. Lastly, we present the formula used to find the average reshare count of frames.

4.3.1 Data

News Articles: We collected tweets using Twitter's standard API[3] with keywords "Climate Change", "#ActOnClimate", "#ClimateChange". The collection period was between August 26th, 2017 to January 4th, 2019. The collection was paused from April 7th, 2018 to May 21st, 2018, due to server errors. Hence, our results are not reflective of these periods. Our dataset consisted of 38M unique tweets and retweets from 11M unique users.

Next, we scrape all the articles shared by news agencies on Twitter using the collected tweets. To find out whether an account is from a news agency, we use a pre-trained model as described in Huang and Carley [80]. The model uses a long-short-term memory neural network [77] with an attention mechanism [15] trained on over 10,000 users. The test accuracy reported on a held-out dataset is 91.6%. We found $\sim 3\%$ percent of users as news agency account with 1.1M unique tweets and retweets. For each of the tweets, we scraped the article shared via URL. We collected 900k files shared via URL. Out of these 900k files, we removed the files which were non-text files and all the files with the error message returned from scraping the news outlet's website. After removing the unwanted files, we were left with 810k articles spread across the

[3] `https://developer.Twitter.com/en/docs/tweets/search/overview/standard`

same timeframe as the Tweets dataset [4]. We will refer to these articles as *news articles* in this paper.

Media Frames Corpus: Work by Boydstun et al. [26], also referred to as Policy Frame Codebook, defines a list of frames that are commonly used in news articles. Media Frame Corpus [35], is an annotated dataset of 22,030 wall street journal articles. The articles are annotated as per the Policy Frame Codebook's 15 frames[5]. The dataset consists of articles related to death penalty, gun control, immigration, samesex marriage, and tobacco. The annotation is done manually and could span one or more sentences. However, Media Frame Corpus does not cover climate change related annotated articles and is biased towards the Wall Street Journal's articles. Hence, we use decontextualization methods on the corpus as described in §4.3.2.

ACT Lexicon: We use the expanded EPA lexicon published by Heise [75]. The lexicon was obtained by manual annotation. We further expand the lexicon with Robinson et al. [129], Smith-Lovin et al. [141, 142] datasets, where each word has two different EPA scores; hence, we take the two scores' mean value. In the case of words appearing in multiple data sources with different EPA scores, we take the mean value of each dimension's scores.

4.3.2 Method

Frame Prediction

We use the information score based classification technique as discussed in Field et al. [61] and propose other transformer-based classifiers for sentence-level prediction of frames. In this section, first, we discuss BERT-based models to predict frames at the sentence level. Second, we discuss information score based methods. Last, we evaluate these models for sentence-level accuracy scores. For evaluation, as a benchmark model, we use the Bi-LSTM model proposed

[4]We further discuss our collection process and statistics of the dataset in Appendix 1.
[5]In Appendix 1 we describe different frames.

by Naderi and Hirst [117].

BERT based models: We use a pretrained BERT model to get embeddings of the sentences of different MFC topics [6]. Then we train (1) MLP with one hidden layer of dimension 512 and a softmax layer, (2) 1-D convolution neural network (1D-CNN) similar in dimension to Kim [96].

Information score based prediction: We use the information score based technique as used and validated in Field et al. [61]. In the study, each word is assigned an information score depending upon the frequency of that word occurring in a particular frame. Models evaluated are: (1) PMI-Non Decontextualize : Use information score of unigrams to predict each document's frame similar to Field et al. [61] but without extension of vocabulary, (2) Field et al. [61] : Use information scores but decontextualize by selecting similar words and assigning them the same score using pre-trained continuous bag of words (CBOW) language model embeddings, (3) PMI- Decontextualize (CBOW/FastText): We use the information score lexicons, but instead of adding similar words, we find words during testing which are not in our information score vocabulary. Then we assign these words a score based on a pre-trained language model (CBOW or FastText [23]). The score is assigned for each missing word based on the nearest word in our information score vocabulary.

Evaluation: To check how well the learned models transfer, we train models on four topics and test the model on a different topic. We report the 15-class average prediction accuracy for sentence-level prediction task in Table 4.1. Pretrained BERT models outperform other models. This shows the advantage of attention based models as shown by results reported for major NLP tasks in Devlin et al. [49]. For further analysis of frames we would rely on the validated PMI model [61] which gives reasonable accuracy and is much faster than other models.

[6] In Appendix 1 we provide details of the BERT model used in this paper.

Table 4.1: Prediction accuracy for sentence-level 15 class frame prediction. Given 5 topics in MFC, the accuracy values refer to the average accuracy of training on four topics and predicting on the other remaining topic. * Naderi and Hirst [117] used same topic for testing and training.

Model	**Accuracy**
BenchMark -BiLSTM	0.52*
BERT + MLP	0.53
BERT + CNN	0.54
PMI-Non Decontextualize	0.41
Field et al. [61]	0.47
PMI- Decontextualize (CBOW)	0.48
PMI- Decontextualize (FastText)	0.48

Frame Projection to EPA

Field et al. [61] gives an information score to each word based on the word belonging to one frame over the other. We use the same method to find the information score for each word. Similar to Field et al. [61], Roy and Goldwasser [131], we remove all words occurring in 2% and 98% of the articles. We enriched our lexicon using the decontextualization method used in Field et al. [61] and as benchmarked above (model (2)) with other models. For each frame F, the information score for each word is defined as follows:

$$I(F, w) = \frac{P(F, w)}{P(F)P(w)} = \frac{P(w|F)}{P(w)} \tag{4.1}$$

where $P(w|F)$ is calculated from the fraction of count of words w and count of all words in sentences annotated with frame F. Similarly, $P(w)$ is calculated from entire MFC training data. We use symbol f to denote set of words with information score associated to frame F.

Next, we use the ACT lexicon (l) to get a $[E_{w'}, P_{w'}, A_{w'}]$ score for each word $w' \in l$. We define EPA score of each frame F as:

$$[E_F, P_F, A_F] = \sum_{c \in l \cap f} \frac{I_{(F,c)} * [E_c, P_c, A_c]}{Z} \tag{4.2}$$

where Z is the normalization factor equal to the number of words in both EPA lexicon (l) and f. In Equation 4.2 each word which is in both the lexicons are weighted by their respective information score in EPA space. In Appendix 1 Table 4.2 we report the number of common words in EPA lexicons (l) and different frames (f)[7]. We find that "Capacity and Resources" frame has the least number of common words with 1210 words and "Crime and Punishment" and "Cultural Identity" with the most common words with 1756 words each.

Frame's Average Reshare Count

To find out the mean reshare count for each frame, we use each article's retweet count. For each of the 810k news articles shared via Tweets we scrape the retweet count using Twitter's standard API. We scraped the retweet count of each Tweet in January of 2021, assuming that this retweet count represents the final number of retweets. We believe that this assumption is reasonable since the last Tweet used to collect a news article was on January 4th, 2019 (refer §4.3.1). We use the retweet count and average information score calculated from the common words in framing lexicon (f) and each article to find the mean reshare count (R_F) of frame F as:

$$R_F = \frac{\sum_a r_a \frac{\sum_{c \in a \cap f} I_{(F,c)}}{\sum_F \sum_{c \in a \cap f} I_{(F,c)}}}{\#(a)} \tag{4.3}$$

Where r_a is the retweet count of each article a. In Equation 4.3, the numerator represents the weighted average of the retweet count for each frame given the information score of an

[7]Finding EPA score for words is an ongoing effort [142] and we expect as more words are added to the EPA lexicon projecting frames to EPA space would become more robust.

article. This is then summed for each article. The denominator represents the total number of articles [8].

4.4 Results

4.4.1 Frame Prediction

We find that the "Cultural Identity" frame is the most dominant frame used in climate change articles. To find the frame of a document, we use all the sentence's average score in that document. In Figure 4.1 we report the count of the number of articles with respective dominant frames. We call a frame dominant if the frame is in the top 3 of all the frames [9]. Apart from the "Cultural Identity" frame, we find that "Public Sentiment", "Political," and "Economic" frames are other considerable dominant frames. The "Cultural Identity" frame is defined as "traditions, customs, or values of a social group in relation to a policy issue" [26]. In a manual evaluation of a sample of 100 articles, we find that articles dominant in "Cultural Identity" framing are about changing current practices (eating habits, buying of estate etc.), about protests regarding climate change or changes after a natural disaster [10]. In Figure 4.1 we also report the average scores of the information scores used to calculate the dominant frame. There is a high correlation (Pearson Correlation = 0.9) between the top 3 dominant frames and the mean information scores. Frames such as "External Regulation and Reputation" and "Policy Prescription and Evaluation" show the opposite behavior. We conclude that these frames do occur regularly in different climate change articles but are more salient.

[8]Due to Tweet/user account deletion, in our second run to scrape the retweet count, we were able to collect the count for 700k articles. We use these 700k articles for our average reshare analysis.

[9]We chose the top 3 dominant frames to remove uncertainty involving the model being only 47% accurate. On using the top 3 dominant frames, the accuracy score on the training set was 74%.

[10]We discuss the details of the manual evaluation process in Appendix 2.

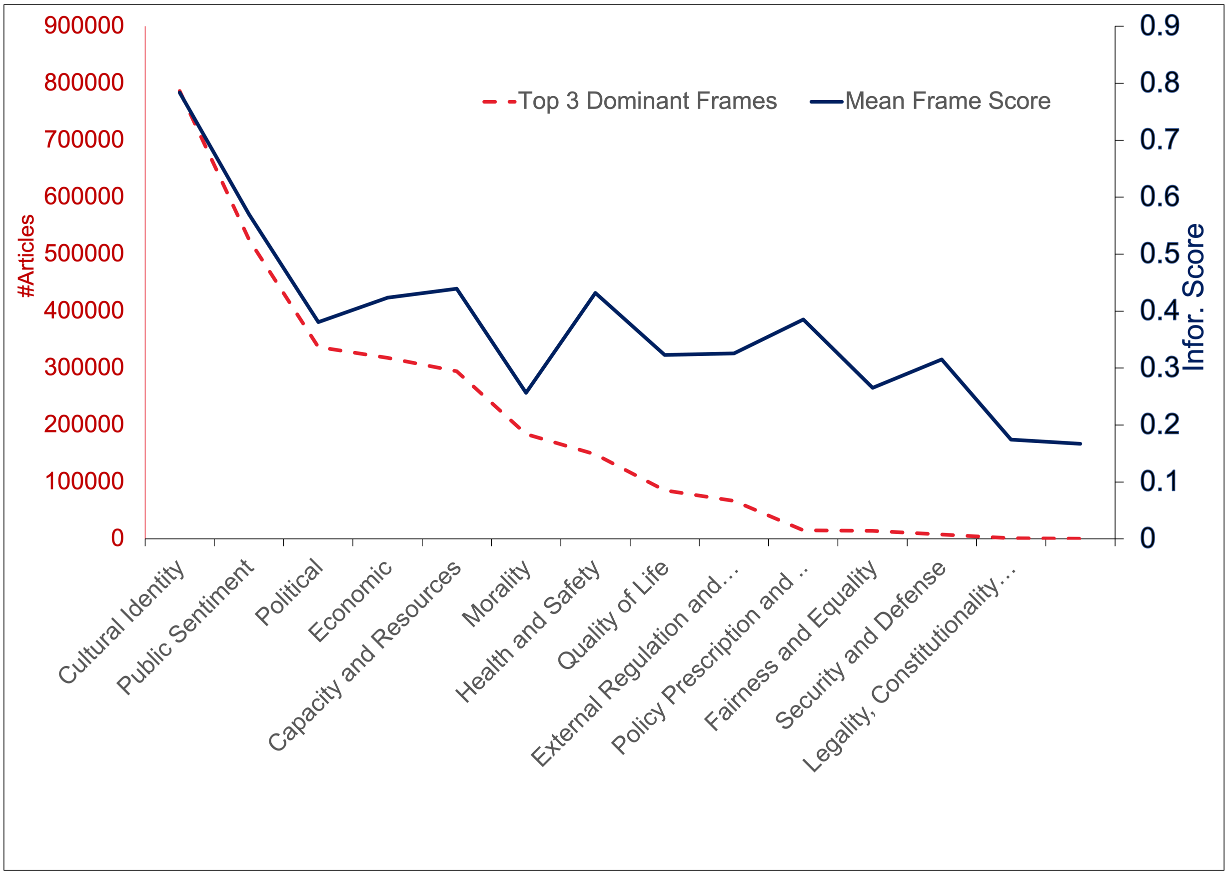


Figure 4.1: Number of articles with corresponding top 3 frames and mean of the information score (Equation 4.1) for all the articles (blue).

4.4.2 Frames in EPA

We project the frames in EPA space to find that "Capacity and Resources", "Quality of Life," and "Morality" score high in the Evaluation (good vs. bad) dimension. "Morality" also scores high in Potency (strong vs. weak). In Figure 4.2, we report each frame's EPA dimension and the centered and scaled value to better show contrast between frames. On a manual inspection of top words contributing to high Potency values of "Morality" frame, we find words related to religion such as *jesus*, *christ* and *church*. These words have a higher than usual Potency value. "External Regulation and Reputation" and "Quality of Life" frame score high in the Activity (active vs. passive) dimension. Overall, we find that all the frames are positive (leaning good,

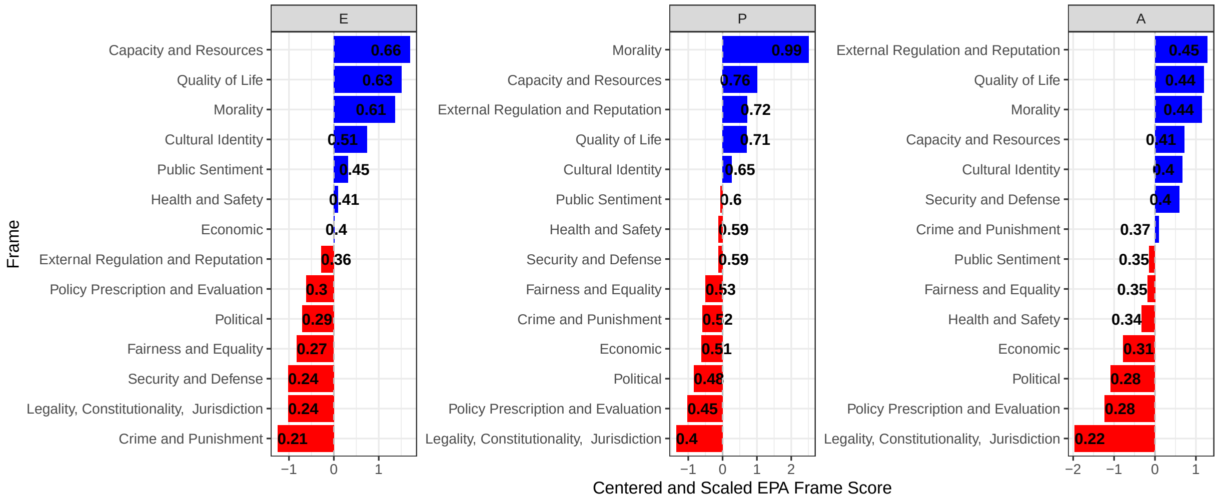


Figure 4.2: EPA values centered by mean frame score and scaled by the standard deviation values of frames in each dimension. EPA scores of frames calculated using Equation 4.2 are in boldface. EPA dimension range is [-4.3,+4.3].

strong, and active) with little variation. This is expected as frames are nuanced changes in the presentation of a topic. Moreover, the common words between the EPA lexicon and the frames represent news agencies' neutral emotions. We infer that the highly emotional words in the EPA lexicon are rare or do not occur in our framing information score lexicon.

The EPA dimensions of frames used in news articles do not vary greatly with time. Figure 4.3 reports our results. For this analysis, we aggregate articles by month to find the average information score for each frame and then convert to the EPA dimension by taking a weighted average using the base EPA dimension score of each frame from Equation 4.2. Although the number of articles in each month varies greatly, we find little or no variation with time in all three affective dimensions. Potency dimension score is the highest, followed by Evaluation and then Activity. This is explained by the fact that in figure 4.2, each frame's base score is higher for the Potency dimension.

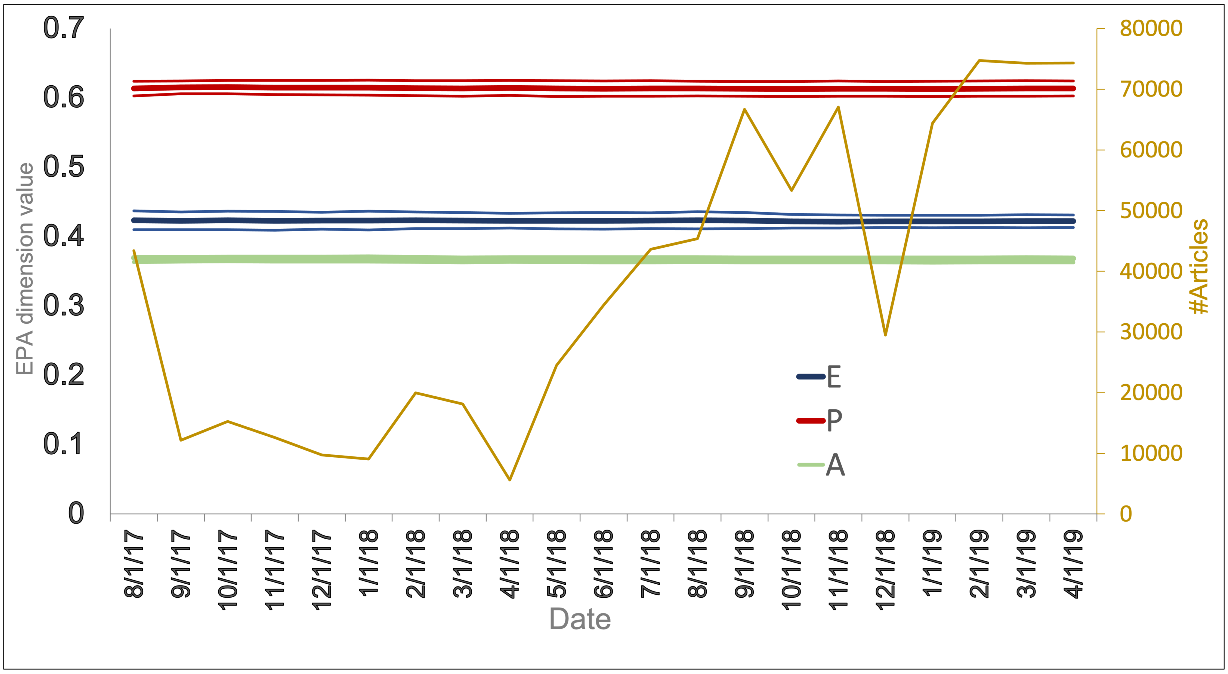


Figure 4.3: EPA dimensions of frames used in climate change news articles aggregated by month and the total number of articles in each month. Extra lines of the corresponding color represent 1 standard deviation.

4.4.3 Reshare Count of Frames

The average reshare (retweet) count varies considerably for different frames. Figure 4.4 reports the average reshare count of each frame. The average reshare is less than 1, indicating that a high percentage of articles were not shared. In fact, only 35% of the news articles were shared more than once. As described in §4.3.1, we scrape articles from all accounts that exhibit news agency like behavior based on Tweets and user account's metadata. Based on our previous experiments using the classification model, we infer that not many users follow a high percentage of the accounts labeled as news agency accounts. We suspect that some of these accounts could be bot-like. We leave the extended analysis for bot-like accounts for future work. On average, the "Cultural Identity" and "Public sentiment" frame is more than two times more reshared than the "Crime and Punishment" frame.

Next, we find out if *affect* of the frames drive their reshare activity. In figure 4.5, we report the emotional value of a frame by calculating the distance from the center (origin) of each frame in EPA dimensions. We find that "Morality", "Quality of Life" and "Capacity and Resources" frames are the most emotional, and "Legality, Constitutionality, Jurisdiction" is the least emotional frame. These results are generally consistent with common perceptions about these frames. We find that there is low correlation (Pearson Correlation = 0.15) between the emotional value of the frames and the average reshare count. This indicates that more emotional (higher *affect*) frames are not necessarily reshared more times.

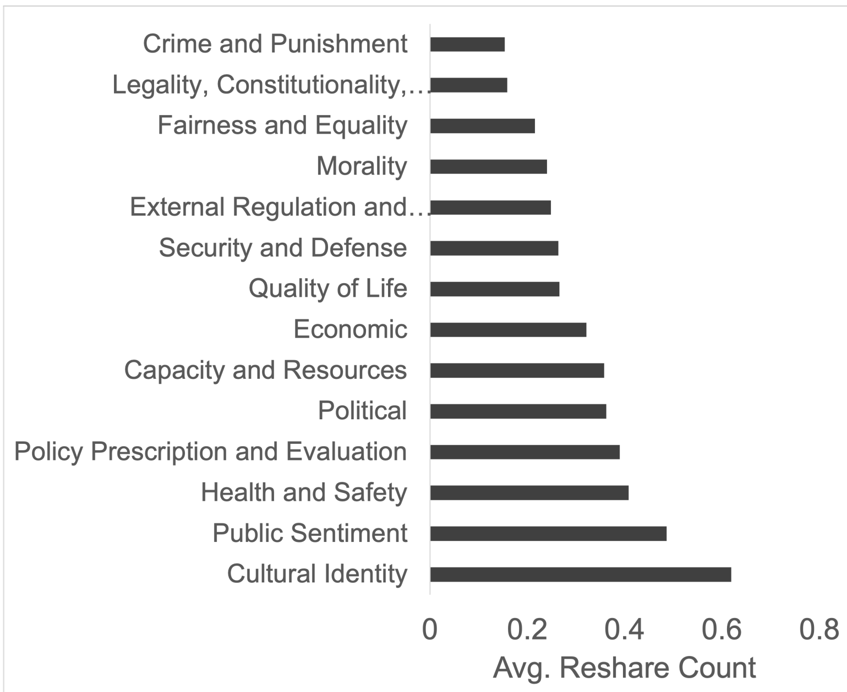


Figure 4.4: Average reshare count (R_F) of each frame calculated using Equation 4.3.

4.5 Discussion

Emphasizing and de-emphasizing certain information to manipulate public opinion has led to a growing interest in learning automated frames in articles [131]. Moreover, work done by

Figure 4.5: Total emotional value of frames calculated by finding the distance from the origin (always >0) of the projected frames in EPA dimensions. A higher value signifies more emotions or *affect* for that frame.

Kause et al. [95] indicates that difference in the framing of climate change communication could contribute to polarization in beliefs. In this work, we use MFC to find automated frames in an extensive corpus of climate change-related articles. We find that most of the articles on climate change are framed using mainly “cultural identity”, “public sentiment”, “political,” and “economic” frames. In work done by Field et al. [61] related to the articles published by the Russian government news media were mostly “External Regulation”, “Political” and “Morality” dominant while using keywords related to the U.S. In a similar work, Roy and Goldwasser [131] also classified news media articles and found the ideological differences in different news media presentation of similar topic. Given our corpus’s extensive size, we believe the dominant frame in climate change articles reflects general news articles’ nature shared on Twitter. In a manual analysis of 100 articles, we find that news articles about climate change predominantly discuss topics involving changing habits, protests, and the effects of natural disasters. Moreover, the 100 random articles used in our manual evaluation were, for the most part, from local or non-popular news sources. Thus, these news articles either address the population of a specific place or a region or are reposts of national/international news stories.

A perception exists in climate change communication that "considerable competition among (and between) scientists, industry, policymakers and non-governmental organizations (NGOs), each of whom is likely to be actively seeking to establish their particular perspectives on the issues" [5]. Previous studies have described climate change framing in "scientific uncertainty" frame [55, 109]. This framing has recently changed to "industry leadership" frame in defeating climate change [82]. Frames have also been shown to differ between countries and over time [68, 135, 136]. A recent work done by Badullovich et al. [13] suggests that scientific literature on climate change most commonly use "Scientific, Economic and Environmental" frames and are increasingly using "Public health, Disaster, and Morality/ethics" frames. In our study, instead of focusing on manual analysis, we use computational models to build on climate change communication's rich framing research. Moreover, we focus on news media to decipher perspectives as the media plays a vital role in climate change communication. Using the retweet count of each article shared via Twitter, we calculated the average number of times a frame is reshared. In this work, we show that certain frames are more reshared than others. Moreover, this resharing pattern is not correlated to different frame's emotional valence. This suggests that news stories are reshared based on other factors such as news media popularity, story type, and novelty. We recognize that the online data collected used English language keywords and did not reflect the demographic representativeness necessary to present cross-cultural conclusions.

In this paper, we develop a methodology to project frames in EPA space. We constructed a mechanism where the nuances of the article content are projected to EPA space. We do not make an effort to predict where articles themselves lie in EPA space. By projecting the frames into EPA space, we can now connect the same topic to emotional science research useful for studying group influence and belief change. Work done by Britt and Heise [30] gave clues that more active emotions could be used to incite minority groups by motivating them to participate in more extensive group activities. Our results indicate that frames such as “external regulation”, “Quality of Life” and “morality” are more emotionally active (higher activity). As climate change action becomes more urgent and necessary, a more consistent and active framing should be used to convey the policy changes needed. Moreover, multiple previous research studies on climate change discussion on social media have concluded that different belief groups exhibit “echo-chamber” type behavior [148, 162]. These different belief groups can be analyzed to find their news sources and align messages with frames that are more likely to be shared by different belief groups. A more engaged community with trustworthy news sources is likely to decrease the confusion around well established climate change facts.

Finally, taking flight from our research, further studies may fruitfully examine several hypotheses opened up by our results. Our work shows how frames could be projected to EPA space and used to find the emotional value within the subtle language used to debate a topic. Future scholarship can look at different topics to compare EPA value of the frames presented. Moreover, with the increase in the size of the EPA lexicon, deep learning approaches could be employed to project the frames. We expect our work would be an essential stepping stone for social scientists to build better communication analysis tools for future climate change communication messages.

4.6 Appendix 1

Definition of framing dimensions from Boydstun et al. [26]:

- Economic: costs, benefits, or other financial implications
- Capacity and resources: availability of physical, human or financial resources, and capacity of current systems
- Morality: religious or ethical implications
- Fairness and equality: balance or distribution of rights, responsibilities, and resources
- Legality, constitutionality and jurisprudence: rights, freedoms, and authority of individuals, corporations, and government
- Policy prescription and evaluation: discussion of specific policies aimed at addressing problems
- Crime and punishment: effectiveness and implications of laws and their enforcement
- Security and defense: threats to welfare of the individual, community, or nation
- Health and safety: health care, sanitation, public safety
- Quality of life: threats and opportunities for the individual's wealth, happiness, and well-being
- Cultural identity: traditions, customs, or values of a social group in relation to a policy issue
- Public opinion: attitudes and opinions of the general public, including polling and demographics
- Political: considerations related to politics and politicians, including lobbying, elections, and attempts to sway voters
- External regulation and reputation: international reputation or foreign policy of the U.S.
- Other: any coherent group of frames not covered by the above categories

4.6.1 Data Collection Details

As described in §4.3.1, we collected Tweets using Twitter's standard API [11] using keywords "Climate Change", "#ActOnClimate", "#ClimateChange". Table 4.3 reports statistics of the dataset. We then classify each user into a news agency account and a non-news agency account using the method described in §4.3.1. For each Tweet from a news agency account, we scrape the *news article* using the URL shared in that Tweet. To scrape the news articles, we built our software system, which uses python *requests* library to scrape the articles from the websites. We subscribed to news agencies mentioned in Pew Research's top online news media websites report [12] as these websites generally required login credentials. Extensive testing was done to ensure that we could collect as many articles as possible and circumvent possible obstacles such as AJAX calls and advertisements. Using the URLs we were able to collect 900k articles. However, some of these articles were non-text files or contained short error messages. We removed these files from our dataset. After this step, we were left with 810k news articles. In table 4.3 we report statistics of news articles used in our dataset. We further cleaned each article for any HTML tags, other non-header, or non-body text for our analysis.

4.6.2 Bert Model Details

For predicting frames at the sentence level, we use the pretrained Bert-Large-Uncased model. As per Devlin et al. [49] the model has 24-layer with 1024 hidden dimension, 16 attention heads, and 336 M parameters. The model was trained on BookCorpus [13] and English Wikipedia [14] after removing headers, tables, and lists. In this work, we predict frames for news articles, assuming

[11] https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/overview

[12] https://www.pewresearch.org/wp-content/uploads/sites/8/legacy/NIELSEN-STUDY-Copy.pdf

[13] https://yknzhu.wixsite.com/mbweb

[14] https://en.wikipedia.org/

Table 4.2: Total number of common words in each frame lexicon f and EPA lexicon l.

Frame	**Total Common Words**
Capacity and Resources	1210
Crime and Punishment	1756
Cultural Identity	1756
Economic	1690
External Regulation and Reputation	1312
Fairness and Equality	1590
Health and Safety	1712
Legality, Constitutionality, Jurisdiction	1779
Morality	1631
Policy Prescription and Evaluation	1747
Political	1768
Public Sentiment	1685
Quality of Life	1720
Security and Defense	1519

Table 4.3: Statistics of the Tweets and news articles collected as described in §4.3.1

	Tweets	**Articles**
Total Number	38M	810k
Mean per day	48,860.5	1,157.5
Min per day	2	0
Max per day	243,574	6,513

that the formal language used in books and Wikipedia generally reflects the language used in news articles. To get embedding of a sentence we concatenate the last 4 layers of the BERT model, as suggested in Devlin et al. [49]. This embedding was then passed to a MLP/1D-CNN classifier as described in §4.3.2.

4.7 Appendix 2

In this section, first, we will give some examples from our news articles dataset. Second, we use these examples to explain the frames and their projection in EPA space. Lastly, we discuss the methodology and results of our manual evaluation of 100 randomly selected news articles.

Snippets of news articles in our dataset: Snippet (a): *"STUDY REVEALS HOW CLIMATE CHANGE COULD CAUSE GLOBAL BEER SHORTAGES Severe climate events could cause shortages in the global beer supply, according to new research involving the University of East Anglia (UEA). The study warns that increasingly widespread and severe drought and heat may cause substantial decreases in barley yields worldwide, affecting the supply used to make beer, and ultimately resulting in "dramatic" falls in beer consumption and rises in beer prices."*

Snippet (b): *"Baltimore Is Suing Big Oil Over Climate Change The Supreme Court heard arguments this week in a case brought by the city of Baltimore against more than a dozen major oil and gas companies including BP, ExxonMobil and Shell. The city government argued that the fossil* *fuel giants must pay for the costs of climate change because they knew that their products cause potentially catastrophic global warming.*

Snippet (c): *Small islands use big platform to warn of climate change On the map, their homes are tiny specks in a vast sea of blue, rarely in the headlines and far removed from the centers of power. But for a few days each year, the leaders of small island nations share a podium with presidents and prime ministers from the world's most powerful nations, and their message is clear: Global warming is already changing our lives, and it will change yours too. Speaking shortly after U.S. President Donald Trump — whose fiery speech made no mention of climate change — Danny Faure told the U.N. General Assembly this week that for his country, the Seychelles, it's already a daily reality.*

The snippets from the dataset show different frames used in climate change news articles. The topic discussed in these snippets is different; moreover, these snippets are addressed towards geographically different audiences. Snippet (a) addresses the change in a decrease in yield for barley due to climate change referencing a research study. This whole news article predominantly uses "Cultural Identity" frame. Similarly, the article in snippet (b) uses more "economic" and "Legality, constitutionality and jurisprudence" frame. The article from which Snippet (c) was taken is predominantly using "Quality of life", "Capacity and Resources" and "Morality" frames. Our algorithm discussed in §4.4.2 predicts that the article of snippet (c) is high in emotional value or *affect* followed by the article of snippet (b) and then by the article of snippet (a). This order can also be followed by looking at the predominant frames of these articles.

In order to find out the general stories reported in news articles, we manually verify stories of a sample of 100 news articles from our dataset. Moreover, we also find that if the news article is from a popular news agency or not. We mark the source as popular if its name is mentioned in Pew Research's top online news media websites report [15].

[15] https://www.pewresearch.org/wp-content/uploads/sites/8/legacy/NIELSEN-STUDY-Copy.pdf

Two annotators independently annotated each news article to be related to: protests, natural disasters, social practices (such as drinking, eating, festivals, sports etc.), economic, policy or legal, flora and fauna, political, historical facts/data, satire on climate change action/lack of action, new scientific finding and, any other. We recognize these topics do not constitute all possible topics in the context of climate change. Using these groups, we were able to generalize the dominant stories in the climate change discussion to explain the dominant frames.

We find that social practices (15/13), protests (10/12) and, natural disaster-related (9/10) stories are most prominent. The "any other" (11/8) category was also prominent. The least prominent stories were in satire (1/4) and historical facts (2/2) group. The values in the bracket represent the number of stories as marked by each annotator. The % agreement between annotators is relatively high at 77%. A thorough topical analysis of a large sample of news articles would give a more robust insight into the dynamics of frames in climate change news articles. Future work could further the NLP frames research by connecting generic and topical frames/topics in large datasets. We also find that only 2 news articles are from popular news agencies as listed by Pew Research.

4.8 Appendix 3

In this section, first, we will examine the differences between our method with another validated method to predict frames using MFC. Second, we will discuss the results using our method on training topics. Last, we will also provide brief examples of our method for predicting Tweets' frames and for other types of datasets.

Similar to Field et al. [61], we use information scores as described in Equation 4.1. We discard all words that occur in fewer than 0.5% of documents or more than 98% of documents. In Field et al. [61], the initial information score-based total vocabulary size is limited to 250 words with the highest information scores. In our analysis, we use all the words and corresponding information scores. In our method, to decontextualize, we assign the nearest neighbors in the background corpus (K =1000) the same score as found from the information score. We also do not change the score if the word already exists in our lexicon. We believe this gives a fairer decontextualization and, as per our experiments, gives us a slightly better accuracy score.

To help us understand the model's training accuracy, we use our decontextualized information score-based model to predict the training topics' dominant frames. We classify the frame for a sentence as predicted correctly if the frame's gold label is in the top 3 predicted frames. We find that, on average, $\sim$ 74% of the sentences are predicted correctly for the five training topics. We consider this as a reasonable accuracy given our 15-way prediction task. Hence, in our results, we discuss the top 3 dominant frames in climate change-related news articles. Moreover, our results indicate that "cultural identity" and "public sentiment" are very dominant compared to other frames. Thus, we do not expect that uncertainty around the prediction of frames would change our general results.

As our method relies on decontextualized lexicons, we can potentially use our method to Tweets themselves. For example, we compare frames used by Justin Trudeau (liberal) and Andrew Scheer (Conservative) in all their tweets on the day of the Canadian elections (October 21, 2019). We find that Mr. Trudeau used *External Regulation and Reputation, Cultural Identity* as dominant frames, and Mr. Scheer used *Security and Defense, Economic and Capacity and Resources* as dominant frames. A manual analysis validates that these were indeed the dominant frames. We suspect that our framing analysis methods could be used in other datasets and social media platforms such as Reddit.

Chapter 5

Climate Change Conspiracy Theories

5.1 Introduction

There is a virtually 100% consensus among scientists that greenhouse gas emissions from human activity cause climate change [50]. Despite the overwhelming evidence, much public discourse shows open skepticism with many popular contrarian voices [28, 56, 121]. In fact, it is believed that between 20% to 40% of the U.S. population considers climate change as a hoax or do not believe in its anthropogenic cause [151].

Contrarian voices on climate change can be divided among different categories. For instance, there is a category of people who argue that climate change is real but is not caused by human activity. Another example would be people who believe that climate change is real, but they dispute the anthropogenic cause. However, the most alarming category is climate change deniers who outright reject climate science findings or the data as a hoax. Different ideologies drive most people who describe climate science findings and data as hoax [151]. One such facet of ideology is conspiratorial thinking. Previous studies have suggested that conspiratorial thinking is associated with beliefs about climate change. In other words, individuals who believe in conspiracies are more likely to refute the anthropogenic cause of climate change [151].

Conspiracy theories are “unsubstantiated explanations of events or circumstances that accuse powerful malevolent groups of plotting in secret for their own benefit against the common good” [151]. People who believe in conspiracy theory might want to derive an explanation for any complex scientific fact from these theories. Conspiracy theories can be interlinked with each other, although they might not have any logical basis. In this paper, we discuss some of the major conspiracy theories in climate change that are popular on a social media platform. These conspiracy theories present a significant challenge in removing the false narratives around climate change. Thus, it becomes essential to analyze these conspiracy theories.

Previous work on climate change and conspiracy theories suggests that people believing in conspiracy theories are likely to believe that climate change is a hoax [151]. That work relied on manual survey-based collection methods. Surveys are limited in finding nuanced beliefs and in studying extensive social network structures. This paper uses extensive Twitter data to link beliefs about climate change and sharing of conspiracy related text. Thus, the main research question we answer is, *In climate change discussion, do climate change Disbelievers share more conspiracy related terms compared to Believers?* To answer this research question, we scrape Twitter data for all the Tweets containing climate change and conspiracy related keywords. We then use a state-of-the-art stance detection method to find climate change Believers and Disbelievers [1].

Moreover, conspiracies about climate change could be promulgated by bot-like accounts - automated user accounts - in addition to human actors. These bot-like accounts can further create confusion on well established climate change realities. Moreover, previous studies have suggested that “that bots seek to create false amplification of contentious issues with the intention to create discord” [148]. This paper examines whether or not bot-like accounts are more active in sharing conspiratorial messages in different belief groups.

This paper begins by defining some of the common climate change-related conspiracy the-

[1]We define Believers as people who cognitively accept anthropogenic causes of climate change Disbelievers as those who reject the same.

ories §5.2. Next, we discuss the method used to identify individual beliefs, keywords used to identify the conspiracy theories, and method used to find bot-like accounts §5.3. We present our results in §5.4. Our results (§5.4) suggest that Disbelievers share most conspiracy theory related Tweets. Conspiracy theory related to chemtrails and geo-engineering is most popular in our dataset. However, conspiracy theory related to flat earth is most popular among Believers but rather used as sarcasm. We also find that most Disbelievers share only one or two different conspiracy theories with climate change discussion. Finally, we discuss our findings and their implications in §5.5.

5.2 Major Conspiracies about Climate Change

Conspiracy theories evolve with time and do not follow logical arguments. This section covers the most well-known conspiracy theories related to climate change and gives brief backgrounds about each. The list was created based on the author's findings and readings of previous work on the same topic [2] [148, 151, 161].

1. Deep state: Followers of this conspiracy theory agree that there is a hidden government within the legitimately elected government that controls the state. Climate change is a hidden agenda of the deep state to further the deep's states motives.
2. Chem Trails: The condensation trails from the jet engines of an aircraft are erroneously recognized as consisting of chemical or biological agents. The theory posits that these trails are responsible for climate change.
3. Sunspots: Sunspots are a temporary phenomenon of reduced temperature on the Sun's surface [133]. This theory asserts that sunspots and not human activity are causing climate change.
4. Directed Energy Weapon (DEW): A human-made weapon that damages its target by a

[2] In §5.6 we present few example Tweets of each type of conspiracy.

highly focussed beam of energy. As per the proponents of this theory, the usage of DEWs is causing climate change.

5. Flat Earth: Advocates of this conspiracy theory do not believe that the earth is a sphere but rather believe that the earth is a flat disc. Climate is hence not governed by the standard scientific laws, and climate change is a hoax.
6. Geo Engineering: Enthusiasts of this conspiracy theory believe that governmental experiments cause climate change.
7. Unknown Planet: A ninth planet with a vast orbit and unknown to humanity is causing climate change. The effect of the planet will keep on increasing as it goes through its perigee.

5.3 Data Collection and Method

In this section, we first describe our data collection in §5.3.1. Second, in §5.3.2, we describe our method to find climate change belief stance and the keywords used to find conspiracy theory related Tweets.

5.3.1 Data Collection

We collected tweets using Twitter's standard API[3] with keywords "Climate Change", "#ActOnClimate", "#ClimateChange". Our dataset was collected between August 26th, 2017 to September 14th, 2019. Due to server errors, the collection was paused from April 7th, 2018 to May 21st, 2018, and again from May 12th, 2019 to May 16th, 2019. We ignore these periods in our analysis. After deduplicating tweets, our dataset consisted of 38M unique tweets and retweets from 7M unique users.

[3] `https://developer.Twitter.com/en/docs/tweets/search/overview/standard`

5.3.2 Method

This section will first discuss the stance detection method used to identify climate change Believers and Disbelievers. Then, the keywords used to identify conspiracy related Tweets.

Stance Detection: Labeling each user as a climate change Believer or a Disbeliever is a non-trivial task. The broader field of labeling users based on the position the user takes on a particular topic is called *stance mining* [113]. We use state-of-the-art stance mining method which uses weak supervision to find Believers and Disbelievers [102]. The model uses text signals from Tweets along with retweet and hashtag network features using a co-training approach with label propagation [167] and text classification. A set of seed hashtags are provided as a pro and anti stance signals to the model. The model then labels seed users based on the usage of these seed hashtags at the end of the tweet (endtags). The labeled and unlabeled users are then taken as input to the co-training algorithm. In each step, a combined user-retweet and user-hashtag network is used to propagate labels to unlabelled users. Concurrently, the text classifier uses the seed user's tweets to train an SVM [44] based text classifier to predict unlabeled users. A common set from text classification and label propagation of highly confident labels are then used as seed labels for the next iteration. The final classification is based on the prediction of the joint model using the combined confidence scores.[4] The model has been shown to be above 80% accurate with multiple datasets.

We select hashtag *#ClimateHoax* and *#ClimateChangeIsNotReal* as Disbeliever seed hashtags and *#ClimateChangeIsReal* and *#SavetheEarth* as Believer seed hashtags. Hashtags *ClimateHoax* has been shown to be used mostly by Disbelivers [148]. We found similar results on using other Disbeliever hashtags reported in [148]. We use *ClimateChangeIsReal* and *SavetheEarth* as Believer hashtags because of their semantics. Out of the 7M users, we classified 3.1M as disbelievers and 3.9M as believers [5].

[4]We use the parameter values as defined in [102] as $\{k = 5000, p = 5000, \theta^I = 0.1, \theta^U = 0.0, \theta^T = 0.7\}$.

[5]We provide details of manual validation of stance results and the parameters in our project repository

Conspiracy Keywords: We use the following keywords to identify if a Tweet is a conspiracy related Tweet.

1. Deep state: club of rome, clubofrome, clubrome, pizzagate, lizard people, lizardpeople, illuminati, deepstate, deep state, qanon
2. Chem Trails: chemtrail, chem trail
3. Sunspots: sunspot
4. Directed Energy Weapon: dew, directed energy weapon, directedenergy
5. Flat Earth: flat earth, flatearth
6. Geo Engineering: geo engineering, geoengineering, weather modification, weathermodification
7. Unknown Planet: planet x, niburu

Bot Detection: We label an account as bot-like or not using CMU's Bot-Hunter [18, 19]. Bot-Hunter's output is a probability measure of bot-like behavior assigned to each account. Unless otherwise stated, we report our analysis for a probability threshold of 0.6 [6]. In other words, we classified an account as bot-like if the output probability from Bot-Hunter was greater than 0.6.

5.4 Results

Climate change Disbelievers share more conspiracy related Tweets than climate change Believers. We report the number of Tweets and Retweets shared by climate change Believers and Disbelievers in Table 5.1. We see an order of magnitude difference between the activity of the groups. Disbelievers overwhelmingly share Tweets related to conspiracy theories. Interestingly, for both groups, conspiracy theory related Tweets are Tweeted more than Retweeted.

`https://github.com/amantyag/affectiveclimatechange`

[6] We use 0.6 as this probability threshold gives us a lower false-positive rate than generally used 0.5.

This behavior is contrary to the findings of most studies on Twitter which conclude that users prefer Retweeting to Tweeting [25]. More Tweeting activity than Retweet activity suggests that although conspiracy related Tweets can be found in climate change discussion, not many users are re-sharing the message.

Once we know that Disbelievers are predominantly sharing the conspiracy theory related Tweets, next, we find which conspiracy theory is most popular. We break down the Tweets/Retweets with the respective type of conspiracy theory in figure 5.1. As expected, Disbelievers are sharing conspiracy theories more than Believers. The most popular conspiracy theory among Disbelievers is Geo-engineering and Chemtrails related conspiracy theory. On the other hand, Believers are sharing Flat Earth conspiracy theory more than other conspiracies. A manual analysis of 100 randomly selected Tweets shows that the Flat Earth conspiracy theory is used as a sarcastic comment or to make fun of the other group. We provide further evidence by finding the average sentiment towards conspiracy related keywords [7]. Figure 5.2 reports the average sentiment in Tweets towards conspiracy theory related words. Flat Earth conspiracy theory stands out with negative sentiment, more so when shared by Disbelievers. In other words, irrespective of beliefs about climate change, the Flat Earth conspiracy theory is viewed negatively. Interestingly, Believers have a higher positive sentiment towards ChemTrails and Geo-Engineering conspiracy theories compared to Disbelievers. We suspect that this could be attributed to Believers explaining the actualities of these theories. More robust sentiment analysis with a labeled dataset is needed to draw sentiment level conclusions; such analysis is out of scope for this work. We also test our sentiment results in comparison to the rest of the unfiltered climate change discussion. In table 5.2, we report our result. We find that Believers and Disbelievers Tweets exhibit higher positive sentiment when used with conspiracy-related keywords (p-value = 0.0005), which is

[7] To find sentiment towards keywords, we utilize Netmapper [38] which uses a word-level sentiment computation based on the average of known valences of surrounding words within a sliding window. The output values are between -1 and 1, where a negative value represents a negative sentiment, and a positive value represents positive sentiment.

Table 5.1: Number of unique Tweets and Retweets shared by Disbelievers and Believers containing conspiracy theory related keywords.

	Disbeliever	Believer
Tweet	31084	4830
Retweet	14369	3576

Table 5.2: Average sentiment scores of Tweets related to conspiracies and the rest of the Tweets in our unfiltered climate change dataset. We randomly sampled Tweets from the unfiltered dataset to find the rest of the dataset's sentiment score.

Belief (Tweets)	Conspiracy Related	Rest
Disbelievers (40,389)	0.108	0.0364
Believers (1,799)	0.3484	0.0097

truer for Believers than for Disbelievers. This finding further strengthens our argument that Believers are using conspiracy keywords to debunk conspiracies.

After analyzing the origin of different conspiracy theories, next, we look at the correlation of different conspiracy theories shared by each user. In table 5.3, we report the correlation between two different conspiracy theories by finding the number of times different conspiracy keyword is used by each user. We find that most conspiracies are highly correlated with each other, indicating that users who Tweet about one conspiracy also tweet about other conspiracies. The Chemtrails and Unknown Planet conspiracies are least likely to be Tweeted by a user who tweets other conspiracy theories. To further gain insight into the sharing pattern, in figure 5.3 we report the number of users sharing unique conspiracy theories. Even on a log scale, we see a steep decline in the number of Believers and Disbelievers sharing different types of conspiracy theories.

Next, we find whether or not the same Tweet has more than one conspiracy discussed. In

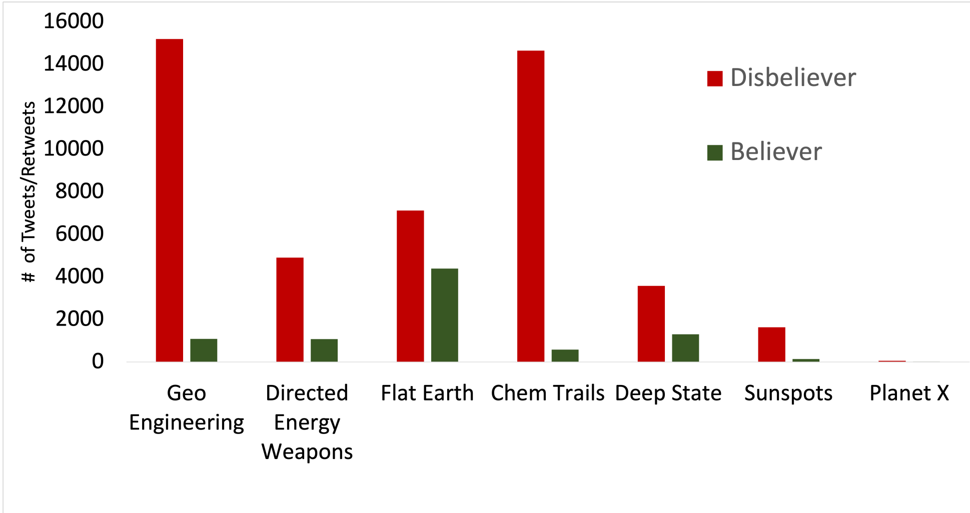


Figure 5.1: Number of unique Tweets and Retweets shared by Disbelievers and Believers containing different conspiracy theory related keywords defined in §5.3.2.

table 5.4, we report the correlation between two different conspiracy theories by finding the number of times different conspiracy keyword occurs in the same Tweet. We notice in table 5.4 that there is a weak negative or close to zero correlation between all the keywords belonging to different conspiracy theories. Twitter users prefer using conspiracy theory keywords independent of using other conspiracy theory keywords in a Tweet. Moreover, conspiracy theories related to Flat Earth and Geo Engineering are most negatively correlated (-0.338). In other words, Twitter users using keywords related to Flat Earth do not use keywords related to Geo Engineering in the same Tweet. Thus, we conclude that most users share one or two types of conspiracy theory and most Tweets have keywords related to one type of conspiracy. Moreover, this behavior does not differ from a change in climate change belief.

Lastly, we look at whether bot-like accounts drive the conspiracy theory related discussion. In table 5.5 we report the fraction of Disbeliever and Believer accounts labeled as bot-like at different probability thresholds [8] We find that even at 0.7 probability cutoff, about a quarter of all users exhibit bot-like characteristics. We also find that there is not much difference

[8] Refer §5.3.2 Bot Detection.

Table 5.3: Correlation matrix of conspiracy theories related keywords used by different users. We find the correlation between two different conspiracy theories by calculating the number of respective keywords used by each user.

	Deep State	Chem Trails	Sunspots	Directed Energy Weapons	Flat Earth	Geo Engineering	Planet X
Deep State	1.000	0.360	0.713	0.953	0.958	0.892	0.161
Chem Trails	0.360	1.000	0.195	0.333	0.321	0.361	0.041
Sunspots	0.713	0.195	1.000	0.751	0.740	0.676	0.156
Directed Energy Weapons	0.953	0.333	0.751	1.000	0.982	0.903	0.202
Flat Earth	0.958	0.321	0.740	0.982	1.000	0.915	0.231
Geo Engineering	0.892	0.361	0.676	0.903	0.915	1.000	0.184
Planet X	0.161	0.041	0.156	0.202	0.231	0.184	1.000

Table 5.4: Correlation matrix of conspiracy theories related keywords occurring in a single Tweet. We find the correlation between two different conspiracy theories by calculating the number of respective keywords used in each Tweet.

	Deep State	Chem Trails	Sunspots	Directed Energy Weapons	Flat Earth	Geo Engineering	Planet X
Deep State	1.000	-0.179	-0.056	-0.106	-0.151	-0.190	-0.010
Chem Trails	-0.179	1.000	-0.114	-0.218	-0.309	-0.284	-0.021
Sunspots	-0.056	-0.114	1.000	-0.065	-0.096	-0.119	-0.006
Directed Energy Weapons	-0.106	-0.218	-0.065	1.000	-0.183	-0.227	-0.012
Flat Earth	-0.151	-0.309	-0.096	-0.183	1.000	**-0.338**	-0.017
Geo Engineering	-0.190	-0.284	-0.119	-0.227	**-0.338**	1.000	-0.022
Planet X	-0.010	-0.021	-0.006	-0.012	-0.017	-0.022	1.000

Table 5.5: Fraction of users labeled as bot-like accounts at different probability thresholds.

Threshold	Believers	Disbeliever
0.5	0.45	0.46
0.6	0.35	0.36
0.7	0.24	0.27

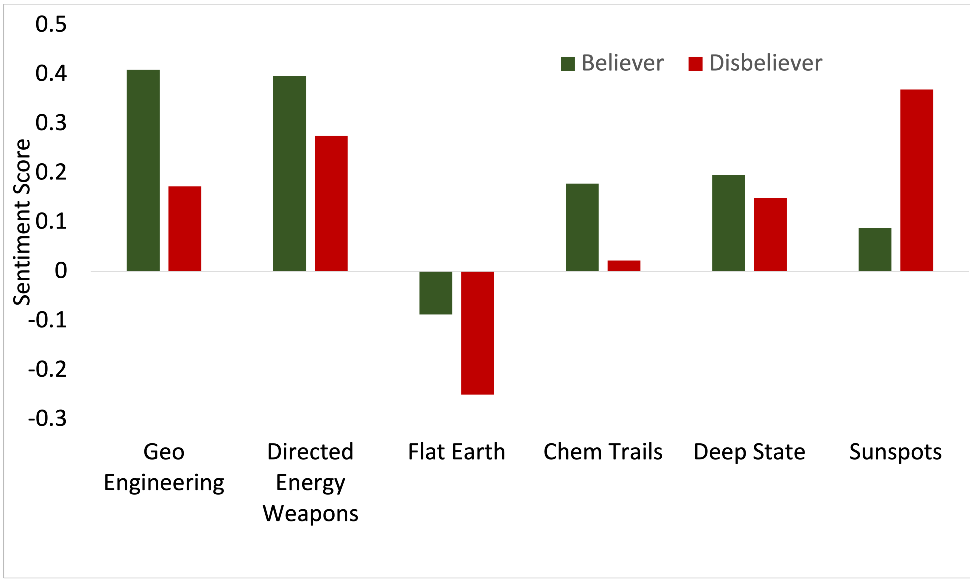


Figure 5.2: Average sentiment score towards the keywords related to different conspiracy theories. A negative value means a negative sentiment and a positive value means a positive sentiment towards the conspiracy keywords.

in the activity between Disbelievers and Believers. Moreover, we find that most bots (∼88%) share only one type of conspiracy theory. This conclusion is similar to the results described in figure 5.3, where we report the distribution without separating bot-like accounts. Moreover, bot-like accounts also show a similar pattern with regards to sharing the type of conspiracy theories. Bot-like accounts showing behavior akin to Disbelievers share more conspiracy theories related to Geo-engineering and Chem Trails. On the other hand, bot-like accounts showing behavior akin to Believers share more conspiracy theories with Flat Earth related keywords. We use a pre-trained model as described in Huang and Carley [80] to find the percentage of bot-like accounts that show behavior akin to news agencies. We find that at 0.6 probability threshold value, ∼12.3% news agency accounts show bot-like activity. This value is lower than the overall ∼35% bot-like accounts that we find for all users, indicating that a lesser number of bot-like accounts show activity similar to news agencies. On the other hand, the number of bot-like accounts in climate change discussion unfiltered for conspiracy theories is $\sim 4\%$. In terms of percentage of total accounts, more bot-like accounts behave as news agencies and engage in conspiracy theory-related discussions than other climate change discussions.

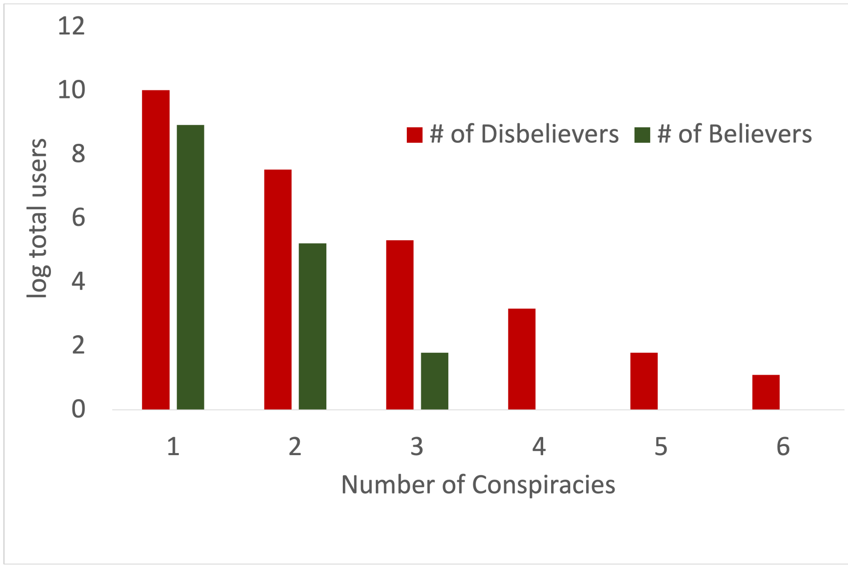


Figure 5.3: Distribution of Disbeliever and Believer users (in log scale) sharing unique conspiracy theories.

5.5 Discussion

Understanding people's underlying beliefs helps understand the constructs by which people could be attracted or repelled by different messaging. People believing in conspiracy theories are more likely to believe that a conspiracy theory is a possible explanation of climate change [151]. Hence, conspiracy theories could be used as a potential recruitment tool by Disbeliever lobbyists. Celebrities and politicians have been vocal about their criticism of science, even using conspiracy theories as possible explanations for climate change [151]. These reasons make the study of conspiracy theories in the climate change context even more relevant.

Conspiracy theories are a means for people to justify the actions of a powerful entity or a person, mostly when those actions are not relatable [32, 33]. Uscinski et al. [151] argue that the influence of elites interacting with the masses predispositions explains conspiracy thinking and why there is a partisan divide in such thinking. Moreover, President Donald Trump's election has further enhanced this effect and could potentially lead to mass radicalization [2]. Conspiracy belief is thus linked to people's justification of predisposed climate change belief. Future research on conspiracy theories warrants these explanations to be looked at from the lens of psychology and social science. In this paper, we find that climate change Disbelievers are more likely to share conspiracy theories. The conspiracy theories range from deep state conspiracy theory, which portrays climate change as an agenda of individual actors or *deep state* to possible explanatory theories such as sunspots and chemtrails. Future research should look at these theories from the lens of explanatory or motivated by partisanship.

Climate change communication research should look to evolve messaging in ways that take into account different beliefs. Conspiratorial thinking and reasoning to justify climate change will dampen the global effort to decrease climate change effects. In this paper, we show that most people sharing conspiracies in the context of climate change only share one or two types of conspiracies. The most popular conspiracy theories are related to Chem Trails or Geo-Engineering. Policymakers should focus on delivering targeted messages to Disbelievers about the scientific practicalities of these conspiracies. Moreover, climate change Believers using Flat Earth conspiracy theory to target Disbelievers or their belief does not help clear scientific facts. Our results suggest that Flat Earth conspiracy theory is not the most popular conspiracy theory among Disbelievers.

Previous studies have concluded that Bot-like accounts stir conversations in differently politically aligned belief groups rather than concentrating on conversations in one belief group [20, 148]. In this study, we further provide evidence that Bot-like accounts were similarly active in sharing conspiracy related messages irrespective of whether they showed activity akin to a Disbeliever or a Believer. These bot-like accounts aim at widening the divide between belief groups and pose a danger of creating confusion on scientific facts [20, 31, 59]. As more and more people consume information via social media, it becomes imperative for these platforms to identify and remove bot-like accounts.

To the best of our knowledge, this study is the first attempt to find conspiracy theories in climate change in a large social media dataset. We find that some conspiracy theories are more popular and used widely to justify climate change compared to others. Future psychology and social science scholarship should divide conspiratorial thinking into different types of conspiracies. This will help find the underlying constructs and motivations, knowing which helps target climate change communication messaging.

Besides the demographic representativeness of the data, there are other limitations in this analysis. First, although we have many tweets about climate change conspiracies, it does not encompass those interactions that do not include our collection keywords. Second, we use a proxy of keywords to classify Tweets as conspiracy related or not. We do not make an effort to find if sarcasm or negation is used to call out conspiracies; we leave this to future scholarship. Last, we focused on the conspiracy theories recorded in media or found during our search. Many more conspiracy theories could be widespread in the climate change debate. Nevertheless, we believe that we were able to analyze the main conspiracy theories more widely popular among general Twitter users.

5.6 Appendix

5.6.1 Different Types of Inaccurate Information

Definition and Examples of different types of inaccurate information:

- Fake News : Defined as"fabricated information that mimics news media content in form but not in organizational process or intent. Fake-news outlets, in turn, lack the news media's editorial norms and processes for ensuring the accuracy and credibility of information. Fake news overlaps with other information disorders, such as misinformation (false or misleading information) and disinformation (false information that is purposely spread to deceive people)" [104]. Example: " 1,000 scientists and 13 federal agencies confirmed that there has been less than a degree increase in global temperatures in over 100 years." this information is followed by graphs and website URLs proving the same.

- Misinformation: Defined as any false or inaccurate information given by media or a person regardless of intent. Although it is hard to know the exact intent of the information source, generally, if the source retracts their comments or clarifies, that could be a sign of nonmalicious intent. Example: "Buttigieg, Sept. 4: And for me and everybody I know, for the children that we hope to have, for the people who will be alive at the turn of the century, when if we don't change what we're doing, we could lose half the world's oxygen because of what's going on in the oceans. That is unthinkable." This is an inaccurate fact [9].

- Disinformation: Any false or inaccurate information spread intentionally for gain. Disinformation is hence a subset of misinformation. Example: "Green new deal would cost $100 trillion." An inaccurate figure, mostly used by the Republican Party of the USA[10].

5.6.2 Example Tweets Related to Conspiracy Theories

Below are the example Tweets of each type of conspiracies:

1. Deep state: "We know the deepstate behind mass shootings creating narratives, to blame you an ban 2-amendment. We know they creating climate Change, to blame you! " "Climate change is a shell game for how the DeepState moves money."

2. Chem Trails: "#lookup #notclouds #skystripes #chemtrails #toxic #poison plastering the sky, covering the sun! #geoengineering #climatechange" , "Bill gates announces plans to use 'Chemtrails' to 'solve global warming' This moronic, psychopath is determined to

[9] https://www.factcheck.org/2019/09/buttigieg-wrong-about-climate-changes“-effect-on-oceans/

[10] https://www.factcheck.org/2019/03/how-much-will-the-green-new-deal-cost/

reduce world population to 500,000. the goal of Elite."

3. Sunspots: "Solar Magnetic Field Oscillations Confirm Global Cooling is Upon Us #ClimateChange #GlobalWarming #GlobalCooling #MiniIceAge #SolarCycles #WeatherCycles #SunSpots #SolarMinimums #Volcanos #Earthquaks #PoleShifts #Taxes #Pseudoscience" , "What's the sun have to do with anything? The leaders of the world can change the temperature. Just ask them. Lack of sunspots to bring record cold, warns NASA scientist - Ice Age Now."

4. Directed Energy Weapon: "Deborah Taveres California Fires, Directed Energy Weapons & the Globalists Push for #Genocide #DEW" "CIA director admitted #DEW #climatechange"

5. Flat Earth: "You seriously wanna debate #ClimateChange? What? have you given up on considering the existence of gravity and still believe in flat earth…huh?"

6. Geo Engineering: "CANT THE WORLD SEE? CLIMATE CHANGE IS MAN MADE NOT THE WAY THEY PRESENT IT TO US! ITS CALLED WEATHER WEAPONRY HURRICANES FIRES DROUGHT EARTHQUAKES-DEW WHY CANT TRUMP TELL US! ITS IS CAUSED BY SPRAYING MASSIVE ALUMINUM AND BARIUM", "Scientists want to synthetically create a volcanic winter, are they just nuts? #climatechange #pseudoscience #geoengineering #weathermodification"

7. Unknown Planet: "Climate change is due to Planet X being next to the sun. It will go away. Ask the Donald what he is going to do about the ET wars in CA."

Chapter 6

Conclusion

Social media suffers from excessive bias, which can lead to the spread of misinformation, the formation of echo-chambers, and confusion of scientific facts. In my research, I studied the challenges in communication for one of society's most significant problems today, i.e., climate change. In my thesis, I analyzed climate change discussions on Twitter to study users confined to different belief groups : (a) users who believe in the anthropogenic cause of Climate Change (Believers); and (b) users who don't (Disbelievers). Firstly, I analyze the interactional polarization among these competing groups. Secondly, I analyze the affective polarization to quantify the hostility between the two groups. Thirdly, I find the framing bias of news articles on climate change shared on Twitter. Lastly, I compare the spread of conspiracies in Believers and Disbelievers. I use climate change as a case study and expect that the model and the analysis developed in this research can be extended to other socio-economic topics. Apart from the frameworks designed to study social media challenges, this work also contributes to climate change messaging research. In chapter 2, I conclude that formal climate change messaging on social media should be neutral in meaning. For example, during the UN's conference of parties (COP) in 2018, *#Takeyourseat* was adopted to promote the conference [118]. This neutral meaning hashtag was reshared by climate change Disbelievers and Believers. Chapter 3 reasons that Disbelievers use natural disasters related words to resist explanations related to the anthropogenic cause of climate change. Climate change messages invoking natural disasters are not likely to change the opinion of Disbelievers. Due to climate change urgency, climate messages should use moral framing to make arguments. Moreover, I infer in chapter 4 that using arguments that invoke threats from external powers (such as Russia profiting from climate change) are more likely to make climate change an urgent issue in people's perception. Lastly, in chapter 5, I find that users only share one or two types of conspiracy theories. A more focused messaging may be better served than using blanket messaging for all conspiracies to remove suspicions on climate realities.

Social media research on detecting fake news and media bias is an upcoming field. The growing interest in the field has led to a plethora of papers and conference venues where the work on challenges in social media could be presented. In spite of the wave of work related to social media challenges, there is a gap in works that connects the social media challenges to actual socio-economic problems. This gap is apparent in the case of climate change communication research on social media. As part of the different chapters in this thesis, I listed the works in climate change communication. There has been work on social media but lack the big data context and usage of state-of-the-art computational methods. This thesis analyzes the challenges using years of Twitter data and million news articles on climate change. As part of the thesis, I also develop methods that could be feasibly applied to the above datasets.

In chapter 2, I develop a method using the semantic cues of hashtags to find competing communities within climate change discussion. The method can find polarized communities in a weakly supervised manner. The results show that Believers and Deniers do not interact with each other as often and show "echo-chamber-ly" behavior. The lack of interaction between communities, or in other words, interactional polarization leads to radicalization of viewpoints. I develop a framework to study the network structure of the two competing communities. The framework suggests that Disbeliever tend to be more structured in the way they communicate on Twitter. My conclusions suggest that appropriately targeting messaging and changing influencer's beliefs could help clear doubts on scientific facts.

Interactional polarization presents an incomplete picture of polarization on social media. Users could be connected but can always be hostile, leading to extreme attitudes. In chapter 3, I develop a framework to quantify affective polarization between Believers and Disbelievers. Results suggest that Disbelievers are more hostile towards Believers than vice-a-versa. Higher hostility among two different belief groups could be constituted as an obstacle for further dialogue between the two groups. I believe that the hostility is not anchored in climate change belief but rather on general or even partisan dislike felt by one group towards the other. The affective and positional polarization is an effect; in the next chapter we look at the cause, which is how or what information is received by both the groups.

Chapter 4 analyzes 810k articles on climate change to find framing bias in climate change news. I come up with a method to find the framing bias in news articles at scale. I find that climate change news articles are framed in cultural identity frame. By a manual inspection of a sample, I conclude that most of the articles on climate change are published by local news agencies and are specific to a place or a region. I also conclude that using more active frames might be helpful given climate urgency.

Lastly, in chapter 5, I analyze the typical conspiracy theories in climate change discussion. I find that Disbelievers are majorly responsible for sharing messages with conspiracy theory related keywords. Overall, this thesis analyzes and connects climate change communication challenges on social media with climate change messaging research.

In figure 6.1, I present the data pipeline used in this thesis. The Twitter dataset was collected using keywords related to climate change and deduplicated to remove any redundant Tweets. This deduplicated dataset was then used in each chapter. As machine learning methods evolve, I suspect that we can use a similar data pipeline with datasets from different media platforms. My thesis focusses on climate change discussion, the data pipeline (figure 6.1) can be used for other socio-economic topics. Data would need to be collected for other topics using the keywords

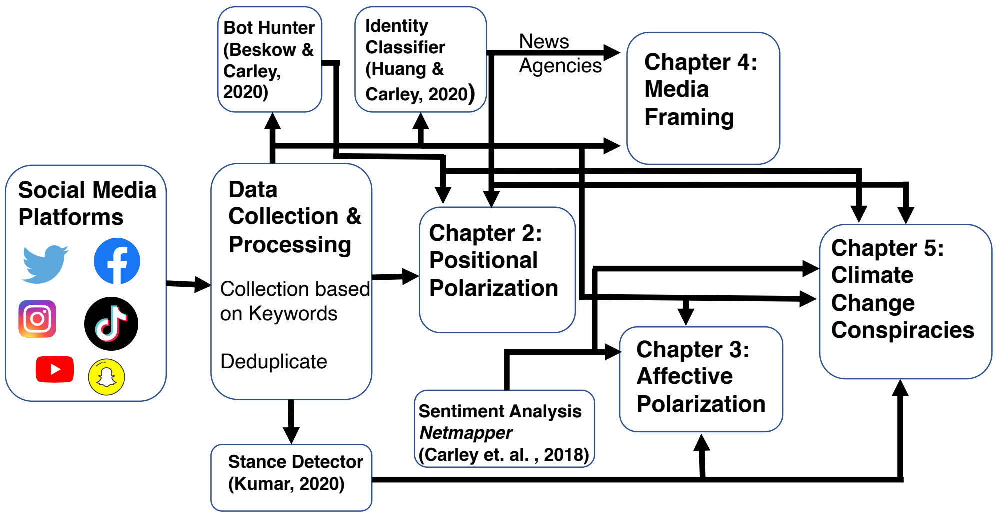


Figure 6.1: Data pipeline used in this thesis. The thesis uses the Twitter dataset with different state-of-the-art methods to inform the analysis in different chapters.

related to the relevant topic, and only minor adjustments would need to be made to perform analysis done in each chapter.

6.1 Implications

Historically, one of the deciding factors to get information was socioeconomic status. Media platforms, or the internet in general, closed the gap between the haves and have-nots regarding access to information. Social media platforms have become a critical source of information and areas of public discourse. However, these same platforms have also magnified the skepticism on scientific facts. In my thesis, I study the challenges that social media platforms face today. As we look into ways to mitigate the effects of climate change and de-carbonize the production of our needs and wants, specific measures must be taken to sensitize social media platforms. These measures will not only help in the adoption of a low-carbon lifestyle but would also indirectly incentivize policymakers to implement policies to stop climate change disaster. Thus, social media platforms need to acknowledge the platform challenges and work towards mitigating them. Implications from my thesis for social media platforms are:

- Polarized and hostile communities exist within climate change discussion. Media platforms should take measures to show messages to educate users with different beliefs.
- Extreme hostility between different groups is harmful to a fruitful discussion. A need thus arises to block toxic messaging and hate speech.
- Bot-like accounts tend to stir conversations, to further the divide and increase "echo-chamberness" among different belief groups. Media platforms need to weed out these automated accounts.
- Obvious factual inconsistencies, such as conspiracies, are present on major platforms such as Twitter. Steps should be taken to regulate the sharing of such inconsistencies.

As media platforms suffer from multiple challenges, policymakers need to be sensitive about the new information medium. The implications from my thesis for government and independent agencies are:

- Presence of divided communities in the climate change discussion means that active social media presence by governmental/non-governmental agencies might not be enough to change beliefs or provoke action.
- For effective climate change communication, agencies should focus on targeted messaging tailored to different belief groups and delivered by influencers in each belief group.
- Policymakers and news media should use more active frames to communicate climate messaging. Using more "Moral", "Quality of life", "External reputation or foreign policy" frames to describe climate change is more likely to make climate change an urgent issue in public perception.
- Conspiracy theories are a real threat to effective climate change messaging. Climate

change messaging should not indiscriminate all conspiracy theories but tackle the popular ones and alienate the unpopular ones.

Framing and targetting of information would matter in a polarized and hostile social media environment. Government or independent agencies educating people on climate change facts should use more "active" frames, given the climate change urgency. As active frames could potentially be used in false information campaigns, agencies would need to build trust with social media users by constant engagement. One way to increase engagement and reduce polarization is by recruiting influencers in different belief groups, especially among Disbelievers. Social media platforms could also alter their recommendation systems to show news or stories followed by other groups. Moreover, as more social media platforms play a more prominent role in shaping society's perspective, policymakers would want to adopt innovative methods to regulate these platforms. One such method would be to use open-source technologies for regulating false information, which would make decision-making on these platforms more transparent and trustworthy. Thus, policymakers should use the internet's scale and speed to make policies for these platforms, rather than relying on traditional methods.

6.2 Limitations and Future Work

The thesis work suffers from several limitations. First, I collect data from only Twitter social media platform. Although Twitter as a platform represents more than 7% of the total social media users in the US [14], different platforms could present many other challenges that are out of scope for this thesis.

Second, the Twitter dataset and the news articles collected are biased towards the English language. We collected tweets using Twitter's standard API[1] with English language keywords. These Tweets were then used to collect news articles used for analysis in Chapter 4. The Tweets and articles hence collected only represent the messages with those keywords. Although 38M

[1] `https://developer.Twitter.com/en/docs/tweets/search/overview/standard`

unique Tweets are collected as part of this thesis's data collection, I believe it represents a significant sample size to make generally valid conclusions.

Last, in the chapters, I discuss results based on validating a sample of the articles or Tweets. In chapters 2,3,4 and 5, I manually validated more than 1000 users. In chapter 4, I manually annotated 100 articles to check the validity of framing analysis results. As most methods in this thesis use weakly supervised approaches, validation using a bigger sample size or future research involving labeled datasets would further strengthen the results.

As more people consume information via social media, the field of scientific communication needs to adapt. As scientific communication would move to new media, more research is needed to understand the new platforms. This thesis concentrates on Twitter social media. Future scholarship should focus on recent social media platforms such as Tik-Tok, Snapchat, Gab, and dark web-based sources. Analyzing these social media platforms would require the use of multimodal datasets and reliance on video and audio-based sources. Current work focuses on the use of text and social network based signals to draw conclusions. Future work should build on the current work and use computer vision and speech translation-based techniques to analyze the recent platform's challenges. Moreover, the current work does not use any traditional survey based method to strengthen the results. I recognize this as a potential future work that would further help connect the present work to traditional climate change communication science.

The present work focuses on climate change due to a lack of work and urgency in this area. The methods and framework developed in this thesis can be used in other issues of socio-economic importance where there is a scientific consensus, but public opinion is still unclear. Such topics include vaccination, immigration, same-sex marriage, among others. The framework developed in this thesis could also be used to analyze partisan topics. These topics could range from analyzing polarization along partisan lines to analyzing framing bias in right and left-wing news sources. As part of my Ph.D. research work, I have used the algorithms developed in this thesis to analyze some of these topics. In [149], I used the work done in chapter 2 to analyze polarization in Indian and Pakistani social media. Similarly, I used the work done in chapter 2 and 3 to analyze division in vaccine communities [111, 147].

In the future, with the increase in social media usage as a news consumption platform, the effect of social media problems will only become more acute. My work in this thesis suggests the need for using a combination of technical and policy tools to mitigate the social media challenges affecting internet users and their opinion.

* * *

“A fire broke out backstage in a theatre. The clown came out to warn the public; they thought it was a joke and applauded. He repeated it; the acclaim was even greater. I think that’s just how the world will come to an end: to general applause from wits who believe it’s a joke.” - Soren Kierkegaard

Bibliography

[1] Eugene Agichtein, Eric Brill, and Susan Dumais. Improving web search ranking by incorporating user behavior information. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 19–26, 2006. 1

[2] HANNAH ALLAM. Right-wing embrace of conspiracy is 'mass radicalization,' experts warn. *NPR*, 2020. 5.5

[3] Hunt Allcott, Matthew Gentzkow, and Chuan Yu. Trends in the diffusion of misinformation on social media. *Research & Politics*, 6(2):2053168019848554, 2019. 2.1

[4] Neal Altman, Kathleen M Carley, and Jeffrey Reminga. Ora user's guide 2018. *Carnegie-Mellon Univ. Pittsburgh PA Inst of Software Research International, Tech. Rep.*, 2018. (document), 2.2, 2.3.2

[5] Alison Anderson. Media, politics and climate change: Towards a new research agenda. *Sociology compass*, 3(2):166–182, 2009. 4.5

[6] Ashley A Anderson and Heidi E Huntington. Social media, science, and attack discourse: How Twitter discussions of climate change use sarcasm and incivility. *Science Communication*, 39(5):598–620, 2017. 3.1, 3.2.1, 3.2.2, 3.5

[7] Dimo Angelov. Top2vec: Distributed representations of topics. *arXiv preprint arXiv:2008.09470*, 2020. 4.1

[8] Ian Angus. *Facing the Anthropocene: fossil capitalism and the crisis of the earth system.* NYU Press, 2016. 2.1

[9] Demetris Antoniades and Constantine Dovrolis. Co-evolutionary dynamics in social networks: A case study of twitter. *Computational Social Networks*, 2(1):14, 2015. 2.4

[10] Dorothee Arlt, Imke Hoppe, Josephine B Schmitt, Fenja De Silva-Schmidt, and Michael Brüggemann. Climate engagement in a digital age: Exploring the drivers of participation in climate discourse online in the context of cop21. *Environmental communication*, 12(1): 84–98, 2018. 2.4

[11] Matthew Babcock, David M Beskow, and Kathleen M Carley. Beaten up on twitter? exploring fake news and satirical responses during the black panther movie event. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 97–103. Springer, 2018. 2.2.1

[12] Matthew Babcock, David M Beskow, and Kathleen M Carley. Different faces of false: The spread and curtailment of false information in the black panther twitter discussion. *Journal of Data and Information Quality (JDIQ)*, 11(4):1–15, 2019. 2.2.1

[13] N Badullovich, WJ Grant, and RM Colvin. Framing climate change for effective communication: a systematic map. *Environmental Research Letters*, 15(12):123002, 2020. 4.5

[14] Jay Baer. 7 surprising statistics about twitter in america, 2020. URL `https://www.convinceandconvert.com/social-media-strategy/7-surprising-statistics-about-twitter-in-america/`. 6.2

[15] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014. 1.1.2, 2.2.5, 4.3.1

[16] Antoine Banks, Ernesto Calvo, David Karol, and Shibley Telhami. # polarizedfeeds: Three

experiments on polarization, framing, and social media. *The International Journal of Press/Politics*, page 1940161220940964, 2020. 3.2.1

[17] Pablo Barberá, John T Jost, Jonathan Nagler, Joshua A Tucker, and Richard Bonneau. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10):1531–1542, 2015. 3.1, 3.2.1

[18] David Beskow, Kathleen M Carley, Halil Bisgin, Ayaz Hyder, Chris Dancy, and Robert Thomson. Introducing bothunter: A tiered approach to detection and characterizing automated activity on twitter. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation. Springer*, 2018. (document), 2.2.5, 2.5, 5.3.2

[19] David M Beskow and Kathleen M Carley. You are known by your friends: Leveraging network metrics for bot detection in twitter. In *Open Source Intelligence and Cyber Crime*, pages 53–88. Springer, 2020. 2.2.5, 5.3.2

[20] Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 us presidential election online discussion. *First Monday*, 21(11-7), 2016. 2.4, 5.5

[21] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003. 2.3.1, 4.1

[22] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008. 2.3.2

[23] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *arXiv preprint arXiv:1607.04606*, 2016. 4.3.2

[24] Toby Bolsen and Matthew A Shapiro. The US news media, polarization on climate change, and pathways to effective communication. *Environmental Communication*, 12(2):149–163, 2018. 3.2.2, 3.5

[25] Danah Boyd, Scott Golder, and Gilad Lotan. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *2010 43rd Hawaii international conference on system sciences*, pages 1–10. IEEE, 2010. 5.4

[26] Amber E. Boydstun, Dallas Card, Justin Gross, Paul Resnick, and Noah A. Smith. Tracking the development of media frames within and across policy issues, Jun 2018. URL `https://kilthub.cmu.edu/articles/journalcontribution/TrackingtheDevelopmentofMediaFrameswithinandacrossPolicyIssues/6473780/1`. 4.1, 4.2.1, 4.3.1, 4.4.1, 4.6

[27] Maxwell T Boykoff. Public enemy no. 1? understanding media representations of outlier views on climate change. *American behavioral scientist*, 57(6):796–817, 2013. 2.4

[28] Maxwell T Boykoff and Jules M Boykoff. Balance as bias: Global warming and the us prestige press. *Global environmental change*, 14(2):125–136, 2004. 5.1

[29] William J Brady, Julian A Wills, John T Jost, Joshua A Tucker, and Jay J Van Bavel. Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28):7313–7318, 2017. 3.5

[30] Lory Britt and David Heise. From shame to pride in identity politics. *Self, identity, and social movements*, 5:252–68, 2000. 4.5

[31] David A Broniatowski, Amelia M Jamison, SiHua Qi, Lulwah AlKulaib, Tao Chen, Adrian Benton, Sandra C Quinn, and Mark Dredze. Weaponized health communication: Twitter bots and russian trolls amplify the vaccine debate. *American journal of public health*, 108 (10):1378–1384, 2018. 2.2.1, 5.5

[32] Rob Brotherton. *Suspicious minds: Why we believe conspiracy theories*. Bloomsbury Publishing, 2015. 5.5

[33] Robert Brotherton, Christopher C French, and Alan D Pickering. Measuring belief in conspiracy theories: The generic conspiracist beliefs scale. *Frontiers in psychology*, 4:279,

2013. 5.5

[34] Suzan Burton, Ann Dadich, and Alena Soboleva. Competing voices: Marketing and counter-marketing alcohol on twitter. *Journal of Nonprofit & Public Sector Marketing*, 25(2):186–209, 2013. 2.2.1

[35] Dallas Card, Amber Boydstun, Justin H Gross, Philip Resnik, and Noah A Smith. The media frames corpus: Annotations of frames across issues. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 438–444, 2015. 1.2, 4.1, 4.3.1

[36] Dallas Card, Justin Gross, Amber Boydstun, and Noah A. Smith. Analyzing framing through the casts of characters in the news. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1410–1420, Austin, Texas, November 2016. Association for Computational Linguistics. doi: 10.18653/v1/D16-1148. URL `https://www.aclweb.org/anthology/D16-1148`. 4.2.1

[37] Kathleen M Carley. Ora: A toolkit for dynamic network analysis and visualization., 2017. (document), 2.2, 2.3.2

[38] L Richard Carley, Jeff Reminga, and Kathleen M Carley. ORA & Netmapper. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*. Springer, 2018. 3.3.3, 7

[39] Jason T Carmichael and Robert J Brulle. Elite cues, media coverage, and public concern: an integrated path analysis of public opinion on climate change, 2001–2013. *Environmental Politics*, 26(2):232–252, 2017. 3.2.2

[40] Dennis Chong and James N Druckman. Framing theory. *Annu. Rev. Polit. Sci.*, 10:103–126, 2007. 4.1

[41] Jake Cigainero. Who are france's yellow vest protesters, and what do they want?,

Dec 2018. URL https://www.npr.org/2018/12/03/672862353/who-are-frances-yellow-vest-protesters-and-what-do-they-want. 2.3.1

[42] Emily M Cody, Andrew J Reagan, Lewis Mitchell, Peter Sheridan Dodds, and Christopher M Danforth. Climate change sentiment on twitter: An unsolicited public opinion poll. *PloS one*, 10(8), 2015. 2.1

[43] M. Coletto, A. Esuli, C. Lucchese, C. I. Muntean, F. M. Nardini, R. Perego, and C. Renso. Sentiment-enhanced multidimensional analysis of online social networks: Perception of the Mediterranean refugees crisis. In *Proc. of ASONAM*, pages 1270–1277, 2016. 2.2.3

[44] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3): 273–297, 1995. 5.3.2

[45] Kareem Darwish. Quantifying polarization on twitter: The Kavanaugh nomination. In *Proc. SocInfo*, pages 188–201, 2019. 2.2.3, 3.2.1

[46] Abhinandan S Das, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web*, pages 271–280, 2007. 1

[47] Claes H De Vreese. News framing: Theory and typology. *Information design journal & document design*, 13(1), 2005. 4.1

[48] Dorottya Demszky, Nikhil Garg, Rob Voigt, James Zou, Jesse Shapiro, Matthew Gentzkow, and Dan Jurafsky. Analyzing polarization in social media: Method and application to tweets on 21 mass shootings. In *Proc. of NAACL*, pages 2970–3005, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1304. URL https://www.aclweb.org/anthology/N19-1304. 2.2.4, 3.2.1

[49] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-

training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018. 4.1, 4.3.2, 4.6.2

[50] Peter T Doran and Maggie Kendall Zimmerman. Examining the scientific consensus on climate change. *Eos, Transactions American Geophysical Union*, 90(3):22–23, 2009. 5.1

[51] Mark Dredze, David A Broniatowski, and Karen M Hilyard. Zika vaccine misconceptions: A social media analysis. *Vaccine*, 34(30):3441, 2016. 2.2.1

[52] Mark Dredze, David A Broniatowski, Michael C Smith, and Karen M Hilyard. Understanding vaccine refusal: Why we need social media now. *American journal of preventive medicine*, 50(4):550–552, Apr 2016. ISSN 1873-2607. doi: 10.1016/j.amepre.2015.10.002. URL https://pubmed.ncbi.nlm.nih.gov/26655067. 1

[53] Mark Dredze, Zachary Wood-Doughty, Sandra Crouse Quinn, and David A Broniatowski. Vaccine opponents' use of twitter during the 2016 us presidential election: Implications for practice and policy. *Vaccine*, 35(36):4670–4672, 2017. 2.2.1

[54] James N Druckman and Matthew S Levendusky. What do we measure when we measure affective polarization? *Public Opinion Quarterly*, 83(1):114–122, 2019. 3.1, 3.2.1

[55] Riley E Dunlap and Aaron M McCright. Organized climate change denial. *The Oxford handbook of climate change and society*, 1:144–160, 2011. 4.5

[56] Riley E Dunlap, Aaron M McCright, and Jerrod H Yarosh. The political divide on climate change: Partisan polarization widens in the US. *Environment: Science and Policy for Sustainable Development*, 58(5):4–23, 2016. 2.1, 2.3.1, 3.1, 3.2.2, 3.2.3, 3.5, 5.1

[57] Robert M. Entman. Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4):51–58, 02 2006. ISSN 0021-9916. doi: 10.1111/j.1460-2466.1993.tb01304.x. URL https://doi.org/10.1111/j.1460-2466.1993.tb01304.x. 4.1

[58] Ash Evans. Stance and identity in twitter hashtags. *Language@ internet*, 13(1), 2016. 2.2.3

[59] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *Communications of the ACM*, 59(7):96–104, 2016. 2.1, 5.5

[60] Anjalie Field and Yulia Tsvetkov. Entity-centric contextual affective analysis. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2550–2560, Florence, Italy, July 2019. Association for Computational Linguistics. doi: 10.18653/v1/P19-1243. URL `https://www.aclweb.org/anthology/P19-1243`. 4.1

[61] Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov. Framing and agenda-setting in russian news: a computational analysis of intricate political strategies. *arXiv preprint arXiv:1808.09386*, 2018. 1, 4.1, 4.2.1, 4.3.2, 4.3.2, 4.3.2, 4.1, 4.3.2, 4.5, 4.8

[62] Dana R Fisher, Joseph Waggle, and Philip Leifeld. Where does political polarization come from? Locating polarization within the US climate change debate. *American Behavioral Scientist*, 57(1):70–92, 2013. 3.1, 3.2.2, 3.5

[63] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978. 2.3.2

[64] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Quantifying controversy on social media. *ACM Transactions on Social Computing*, 1(1):1–27, 2018. 3.1, 3.2.1

[65] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Quantifying controversy on social media. *Trans. Soc. Comput.*, 1(1), 2018. ISSN 2469-7818. doi: 10.1145/3140565. 2.2.3

[66] Daniel Geschke, Jan Lorenz, and Peter Holtz. The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *British Journal of Social Psychology*, 58(1):129–149, 2019. 3.1

[67] Katherine Giuffre. Cultural production in networks. In James D. Wright, editor, *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*, pages 466 – 470. Elsevier, Oxford, second edition edition, 2015. ISBN 978-0-08-097087-5. doi: https://doi.org/10.1016/B978-0-08-097086-8.10445-3. URL http://www.sciencedirect.com/science/article/pii/B9780080970868104453. 2.3.2

[68] Joyeeta Gupta. A history of international climate change policy. *Wiley Interdisciplinary Reviews: Climate Change*, 1(5):636–653, 2010. 4.5

[69] Sean Hackbarth. What does 'keep it in the ground' really mean? read this., Dec 2017. URL https://www.uschamber.com/series/above-the-fold/what-does-keep-it-the-ground-really-mean-read. 2.3.1

[70] Lawrence C Hamilton, Joel Hartter, Mary Lemcke-Stampone, David W Moore, and Thomas G Safford. Tracking public beliefs about anthropogenic climate change. *PloS One*, 10(9):e0138208, 2015. 3.1, 3.2.2, 3.2.3, 3.5

[71] Mareike Hartmann, Tallulah Jansen, Isabelle Augenstein, and Anders Søgaard. Issue framing in online discussion fora. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 1401–1407, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1142. URL https://www.aclweb.org/anthology/N19-1142. 4.2.1

[72] David R Heise. *Understanding events: Affect and the construction of social action*. CUP Archive, 1979. 4.2.2

[73] David R Heise. Affect control theory: Concepts and model. *Journal of Mathematical Sociology*, 13(1-2):1–33, 1987. 4.1, 4.2.2

[74] David R Heise. *Expressive order: Confirming sentiments in social actions*. Springer Science & Business Media, 2007. 4.1, 4.2.2, 2

[75] David R Heise. *Surveying cultures: Discovering shared conceptions and sentiments*. John Wiley & Sons, 2010. 4.3.1

[76] Frank L Hitchcock. The distribution of a product from several sources to numerous localities. *Journal of Mathematics and Physics*, 20(1-4):224–230, 1941. 3.3.3

[77] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997. 1.1.2, 2.2.5, 4.3.1

[78] Thomas Hofmann. Probabilistic latent semantic analysis. *arXiv preprint arXiv:1301.6705*, 2013. 4.1

[79] Peter D Howe, Matto Mildenberger, Jennifer R Marlon, and Anthony Leiserowitz. Geographic variation in opinions on climate change at state and local scales in the usa. *Nature Climate Change*, 5(6):596–603, 2015. 2.1

[80] Binxuan Huang and Kathleen M. Carley. Discover your social identity from what you tweet: a content based approach. *Disinformation, Misinformation, and Fake News in Social Media - Emerging Research Challenges and Opportunities*, 2020. 1.1.2, 2.2.5, 4.3.1, 5.4

[81] Binxuan Huang, Yanglan Ou, and Kathleen M Carley. Aspect level sentiment classification with attention-over-attention neural networks. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 197–206. Springer, 2018. 3.1, 3.2.1

[82] Øyvind Ihlen. Business and climate change: the climate response of the world's 30 largest corporations. *Environmental Communication*, 3(2):244–262, 2009. 4.5

[83] Ana Ivanova. *Transnationalisierung von Öffentlichkeiten*. Springer, 2017. 4.1

[84] Shanto Iyengar. *Is anyone responsible?: How television frames political issues*. University of Chicago Press, 1994. 4.1

[85] Shanto Iyengar. Framing responsibility for political issues. *The Annals of the American Academy of Political and Social Science*, 546(1):59–70, 1996. 4.1

[86] S Mo Jang and P Sol Hart. Polarized frames on "climate change" and "global warming" across countries and states: Evidence from twitter big data. *Global Environmental Change*, 32:11–17, 2015. 2.2.1, 3.1

[87] Hank C Jenkins-Smith, Joseph T Ripberger, Carol L Silva, Deven E Carlson, Kuhika Gupta, Nina Carlson, Ani Ter-Mkrtchyan, and Riley E Dunlap. Partisan asymmetry in temporal stability of climate change beliefs. *Nature Climate Change*, 10(4):322–328, 2020. 3.2.2, 3.2.3, 3.5

[88] Yangfeng Ji and Noah A. Smith. Neural discourse structure for text categorization. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 996–1005, Vancouver, Canada, July 2017. Association for Computational Linguistics. doi: 10.18653/v1/P17-1092. URL `https://www.aclweb.org/anthology/P17-1092`. 4.2.1

[89] Kristen Johnson, Di Jin, and Dan Goldwasser. Leveraging behavioral and social information for weakly supervised collective classification of political discourse on Twitter. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 741–752, Vancouver, Canada, July 2017. Association for Computational Linguistics. doi: 10.18653/v1/P17-1069. URL `https://www.aclweb.org/anthology/P17-1069`. 4.2.1

[90] Kenneth Joseph, Wei Wei, Matthew Benigni, and Kathleen M Carley. A social-event based approach to sentiment analysis of identities and behaviors in text. *The Journal of Mathematical Sociology*, 40(3):137–166, 2016. 4.2.2

[91] Kenneth Joseph, Wei Wei, and Kathleen M Carley. Girls rule, boys drool: Extracting semantic and affective stereotypes from twitter. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 1362–1374, 2017. 4.2.2

[92] John T Jost. Ideological asymmetries and the essence of political psychology. *Political psychology*, 38(2):167–208, 2017. 3.5

[93] Dan M Kahan, Hank Jenkins-Smith, Tor Tarantola, Carol L Silva, and Donald Braman. Geoengineering and climate change polarization: Testing a two-channel model of science communication. *The ANNALS of the American Academy of Political and Social Science*, 658 (1):192–222, 2015. 3.2.2, 3.2.3, 3.5

[94] Rune Karlsen, Kari Steen-Johnsen, Dag Wollebæk, and Bernard Enjolras. Echo chamber and trench warfare dynamics in online debates. *European Journal of Communication*, 32 (3):257–273, 2017. 3.1, 3.2.1, 3.5

[95] Astrid Kause, Tarlise Townsend, and Wolfgang Gaissmaier. Framing climate uncertainty: Frame choices reveal and influence climate change beliefs. *Weather, Climate, and Society*, 11(1):199–215, 2019. 4.5

[96] Yoon Kim. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014. 4.3.2

[97] Andrei P Kirilenko and Svetlana O Stepchenkova. Climate change discourse in mass media: application of computer-assisted content analysis. *Journal of Environmental Studies and Sciences*, 2(2):178–191, 2012. 4.1

[98] Andrei P Kirilenko, Tatiana Molodtsova, and Svetlana O Stepchenkova. People as sensors: Mass media and local temperature influence climate change discussion on twitter. *Global Environmental Change*, 30:92–100, 2015. 2.2.1

[99] Naomi Klein. *This changes everything: Capitalism vs. the climate.* Simon and Schuster, 2015. 2.1

[100] Philip Kotler and Eduardo L Roberto. *Social marketing: Strategies for changing public behavior.* The Free Press, 1989. 2.1

[101] David Krackhardt and Robert N Stern. Informal networks and organizational crises: An

experimental simulation. *Social Psychology Quarterly*, pages 123–140, 1988. 3.1, 3.3.3, 3.3.3

[102] Sumeet Kumar. *Social Media Analytics for Stance Mining A Multi-Modal Approach with Weak Supervision.* PhD thesis, Carnegie Mellon University, 2020. 3.2.1, 3.3.2, 5.3.2, 4

[103] George Lakoff and S. Ferguson. The framing of immigration. *Berkeley, CA: The Rockridge Institute*, 2006. 4.1

[104] David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. The science of fake news. *Science*, 359 (6380):1094–1096, 2018. ISSN 0036-8075. doi: 10.1126/science.aao2998. URL `https://science.sciencemag.org/content/359/6380/1094`. 5.6.1

[105] Joshua Leigh. The evolution of the media landscape (and why you probably won't finish this article. *Method London*, 2017. 1

[106] Frank J Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78, 1951. 3.3.3

[107] Antonis Matakos, Evimaria Terzi, and Panayiotis Tsaparas. Measuring and moderating opinion polarization in social networks. *Data Mining and Knowledge Discovery*, 31(5): 1480–1505, 2017. 3.2.1

[108] Paul Matthews. Why are people skeptical about climate change? some insights from blog comments. *Environmental Communication*, 9(2):153–168, 2015. doi: 10.1080/17524032.2014.999694. 2.1

[109] Aaron M McCright and Riley E Dunlap. Defeating kyoto: The conservative movement's impact on us climate change policy. *Social problems*, 50(3):348–373, 2003. 4.5

[110] Aaron M McCright and Riley E Dunlap. The politicization of climate change and polar-

ization in the American public's views of global warming, 2001–2010. *The Sociological Quarterly*, 52(2):155–194, 2011. 3.2.2

[111] Shahan Ali Memon, Aman Tyagi, David R. Mortensen, and Kathleen M. Carley. Characterizing sociolinguistic variation in the competing vaccination communities. In Robert Thomson, Halil Bisgin, Christopher Dancy, Ayaz Hyder, and Muhammad Hussain, editors, *Social, Cultural, and Behavioral Modeling*, pages 118–129, Cham, 2020. Springer International Publishing. 6.2

[112] Taciano L Milfont, Marc S Wilson, and Chris G Sibley. The public's belief in climate change and its human cause are increasing over time. *PloS one*, 12(3):e0174246, 2017. 3.1, 3.2.2, 3.2.3, 3.5

[113] Saif M Mohammad, Parinaz Sobhani, and Svetlana Kiritchenko. Stance and sentiment in tweets. *ACM Transactions on Internet Technology (TOIT)*, 17(3):1–23, 2017. 5.3.2

[114] Ban-Ki Moon. Full text: Ban-ki moon speech. *UN COP 13*, 2007. 1

[115] AJ Morales, Javier Borondo, Juan Carlos Losada, and Rosa M Benito. Measuring political polarization: Twitter shows the two sides of venezuela. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(3):033114, 2015. 3.2.1

[116] Fred Morstatter, Jürgen Pfeffer, Huan Liu, and Kathleen M Carley. Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose. In *7th International AAAI Conference on Weblogs and Social Media, ICWSM 2013*, pages 400–408. AAAI press, 2013. 3.5

[117] Nona Naderi and Graeme Hirst. Classifying frames at the sentence level in news articles. *Policy*, 9:4–233, 2017. (document), 4.2.1, 4.3.2, 4.1

[118] United Nations. Conference of the parties (cop), 2020. URL `https://unfccc.int/process/bodies/supreme-bodies/conference-of-the-parties-cop/`. 2.2.2, 2.3.1, 6

[119] Thomas E Nelson, Rosalee A Clawson, and Zoe M Oxley. Media framing of a civil liberties conflict and its effect on tolerance. *American Political Science Review*, pages 567–583, 1997. 4.1

[120] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999. 2.3.2

[121] James Painter and Neil T Gavin. Climate skepticism in british newspapers, 2007–2011. *Environmental Communication*, 10(4):432–452, 2016. 5.1

[122] Warren Pearce, Kim Holmberg, Iina Hellsten, and Brigitte Nerlich. Climate change on twitter: Topics, communities and conversations about the 2013 ipcc working group 1 report. *PloS one*, 9(4):e94785, 2014. 2.2.1

[123] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011. 2.2.5

[124] Richard M Perloff. *The dynamics of persuasion: Communication and attitudes in the twenty-first century*. Routledge, 2020. 2.1

[125] Laura Pollacci, Alina Sîrbu, Fosca Giannotti, Dino Pedreschi, Claudio Lucchese, and Cristina Ioana Muntean. Sentiment spreading: An epidemic model for lexicon-based sentiment analysis on Twitter. In Floriana Esposito, Roberto Basili, Stefano Ferilli, and Francesca A. Lisi, editors, *Proc. of AI*IA*, pages 114–127, 2017. 2.2.3

[126] James Powell. Scientists reach 100% consensus on anthropogenic global warming. *Bulletin of Science, Technology & Society*, 37(4):183–184, 2017. 1

[127] Aaditya Ramdas, Nicolás García Trillos, and Marco Cuturi. On wasserstein two-sample testing and related families of nonparametric tests. *Entropy*, 19(2):47, 2017. 3.3.3, 3.3.3

[128] Dawn T. Robinson, Lynn Smith-Lovin, and Allison K. Wisecup. *Affect Control Theory*, pages 179–202. Springer US, Boston, MA, 2006. ISBN 978-0-387-30715-

2. doi: 10.1007/978-0-387-30715-2_9. URL https://doi.org/10.1007/978-0-387-30715-29. 4.2.2, 2

[129] Dawn T. Robinson, Lynn Smith-Lovin, Bryan C. Cannon, Jesse K. Clark, Robert Freeland, Jonathan H. Morgan, and Kimberly B. Rogers. Mean affective ratings of 932 identities, 810 behaviors, and 660 modifiers by university of georgia undergraduates in 2012-2014. *University of Georgia: Distributed at UGA Affect Control Theory Website*, 2016. URL http://research.franklin.uga.edu/act/. 4.3.1

[130] Manuel Gomez Rodriguez, Krishna Gummadi, and Bernhard Schoelkopf. Quantifying information overload in social media and its impact on social contagions. *arXiv preprint arXiv:1403.6838*, 2014. 1

[131] Shamik Roy and Dan Goldwasser. Weakly supervised learning of nuanced frames for analyzing polarization in news media. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7698–7716, Online, November 2020. Association for Computational Linguistics. URL https://www.aclweb.org/anthology/2020.emnlp-main.620. 4.1, 4.2.1, 4.3.2, 4.5

[132] David Runciman. A climate of conspiracy, Sep 2015. URL https://www.opendemocracy.net/en/climate-of-conspiracy/. 2.2.3

[133] Alexander Ruzmaikin. Origin of sunspots. *Space Science Reviews*, 95(1-2):43–53, 2001. 3

[134] Mike S Schäfer, Saffron O'Neill, Matthew Nisbet, Shirley Ho, Ezra Markowitz, and Jagadish Thaker. Frame analysis in climate change communication: approaches for assessing journalists' minds, online communication and media portrayals. *Oxford Research Encyclopedia of Climate Science*, 2017. 1, 4.1

[135] Inga Schlichting. Strategic framing of climate change by industry actors: A meta-analysis. *Environmental Communication: A Journal of Nature and Culture*, 7(4):493–511, 2013. 4.5

[136] Andreas Schmidt, Ana Ivanova, and Mike S Schäfer. Media attention for climate change around the world: A comparative analysis of newspaper coverage in 27 countries. *Global Environmental Change*, 23(5):1233–1248, 2013. 4.5

[137] Roy Scranton. *Learning to Die in the Anthropocene: Reflections on the End of a Civilization*. City Lights Publishers, 2015. 2.1

[138] Alexandra Segerberg and W. Lance Bennett. Social media and the organization of collective action: Using twitter to explore the ecologies of two climate change protests. *The Communication Review*, 14(3):197–215, 2011. doi: 10.1080/10714421.2011.597250. 1, 2.1

[139] Holli A Semetko and Patti M Valkenburg. Framing european politics: A content analysis of press and television news. *Journal of communication*, 50(2):93–109, 2000. 4.1

[140] Michael Shellenberger and Ted Nordhaus. *The death of environmentalism: global warming politics in a post-environmental world*. Verlag nicht ermittelbar, 2004. 2.1

[141] Lynn Smith-Lovin, Dawn T. Robinson, Bryan C. Cannon, Jesse K. Clark, Robert Freeland, Jonathan H. Morgan, and Kimberly B. Rogers. Mean affective ratings of 929 identities, 814 behaviors, and 660 modifiers by university of georgia and duke university undergraduates and by community members in durham, nc, in 2012-2014. *University of Georgia: Distributed at UGA Affect Control Theory Website*, 2016. URL `http://research.franklin.uga.edu/act/`. 4.3.1

[142] Lynn Smith-Lovin, Dawn T. Robinson, Bryan C. Cannon, Brent H. Curdy, and Jonathan H. Morgan. Mean affective ratings of 968 identities, 853 behaviors, and 660 modifiers by amazon mechanical turk workers in 2015. *University of Georgia: Distributed at UGA Affect Control Theory Website*, 2019. URL `http://research.franklin.uga.edu/act/`. 4.3.1, 7

[143] Leo G Stewart, Ahmer Arif, and Kate Starbird. Examining trolls and polarization with a retweet network. In *Proc. ACM WSDM, workshop on misinformation and misbehavior*

mining on the web, 2018. 3.2.1

[144] Cass R. Sunstein. *Republic.com 2.0*. Princeton University Press, 2007. ISBN 9780691143286. URL `http://www.jstor.org/stable/j.ctt7tbsw`. 2.4

[145] Cass R Sunstein. *# Republic: Divided democracy in the age of social media*. Princeton University Press, 2018. 1

[146] Henri Tajfel, John C Turner, William G Austin, and Stephen Worchel. An integrative theory of intergroup conflict. *Organizational identity: A reader*, 56:65, 1979. 2.4

[147] Aman Tyagi and Kathleen M Carley. Divide in vaccine belief in covid-19 conversations: Implications for immunization plans. *medRxiv*, 2020. 3.1, 6.2

[148] Aman Tyagi, Matthew Babcock, Kathleen M. Carley, and Douglas C. Sicker. Polarizing tweets on climate change. In Robert Thomson, Halil Bisgin, Christopher Dancy, Ayaz Hyder, and Muhammad Hussain, editors, *Social, Cultural, and Behavioral Modeling*, pages 107–117, Cham, 2020. Springer International Publishing. 2, 3.1, 3.2.2, 3.3.2, 4.5, 5.1, 5.2, 5.3.2, 5.5

[149] Aman Tyagi, Anjalie Field, Priyank Lathwal, Yulia Tsvetkov, and Kathleen M. Carley. A computational analysis of polarization on Indian and Pakistani social media. In *Social Informatics*. Springer, 2020. 1, 2.2.3, 3.1, 3.2.1, 6.2

[150] Aman Tyagi, Joshua Uyheng, and Kathleen M Carley. Affective polarization in online climate change discourse on twitter. *arXiv preprint arXiv:2008.13051*, 2020. 3

[151] Joseph E Uscinski, Karen Douglas, and Stephan Lewandowsky. Climate change conspiracy theories. In *Oxford Research Encyclopedia of Climate Science*. Oxford University Press, 2017. 1, 2.1, 2.3.2, 2.4, 5.1, 5.2, 5.5

[152] Joshua Uyheng and Kathleen M Carley. Bot impacts on public sentiment and community structures: Comparative analysis of three elections in the Asia-Pacific. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behav-*

ior Representation in Modeling and Simulation, pages 12–22. Springer, 2020. 3.3.3, 3.5

[153] Joshua Uyheng, Thomas Magelinski, Ramon Villa-Cox, Christine Sowa, and Kathleen M Carley. Interoperable pipelines for social cyber-security: Assessing Twitter information operations during NATO Trident Juncture 2018. *Computational and Mathematical Organization Theory*, pages 1–19, 2019. 3.5

[154] Thomas W Valente. Network interventions. *Science*, 337(6090):49–53, 2012. 2.4

[155] Leaf Van Boven, Phillip J Ehret, and David K Sherman. Psychological barriers to bipartisan public support for climate policy. *Perspectives on Psychological Science*, 13(4):492–507, 2018. 3.2.2, 3.2.3, 3.5

[156] Christel W van Eck, Bob C Mulder, and Art Dewulf. Online climate change polarization: Interactional framing analysis of climate change blog comments. *Science Communication*, page 1075547020942228, 2020. 3.1, 3.2.2, 3.5

[157] Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, 2018. 1, 2.2.4

[158] Stanley Wasserman, Katherine Faust, et al. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994. 2.3.2, 2.3.2

[159] Ingmar Weber, Venkata R. Kiran Garimella, and Alaa Batayneh. Secular vs. Islamist polarization in Egypt on Twitter. In *Proc. of ASONAM*, pages 290––297, 2013. doi: 10.1145/2492517.2492557. 2.2.3, 3.2.1

[160] Wikipedia contributors. What is qanon? the origins of bizarre conspiracy theory spreading online, 2020. URL `https://www.independent.co.uk/life-style/gadgets-and-tech/news/qanon-explained-what-is-trump-russia-“investigation-pizzagate-a8845226.html`. [Online; accessed 29-October-2010]. 2.3.2

[161] Wikipedia contributors. Wikipedia, the free encyclopedia, 2020. URL https://en.wikipedia.org/wiki/Listofconspiracytheories. [Online; accessed 22-October-2020]. 2.2.4, 5.2

[162] Hywel TP Williams, James R McMurray, Tim Kurz, and F Hugo Lambert. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change*, 32:126–138, 2015. 2.1, 2.4, 3.1, 3.2.2, 3.2.3, 3.5, 4.5

[163] Rong Xiang, Yunfei Long, Mingyu Wan, Jinghang Gu, Qin Lu, and Chu-Ren Huang. Affection driven neural networks for sentiment analysis. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 112–119, Marseille, France, May 2020. European Language Resources Association. ISBN 979-10-95546-34-4. URL https://www.aclweb.org/anthology/2020.lrec-1.14. 4.2.2

[164] Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik. Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. *Political Communication*, pages 1–42, 2020. 3.1, 3.2.1

[165] Sarita Yardi and Danah Boyd. Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of science, technology & society*, 30(5):316–327, 2010. 2.2.1

[166] Marta Zajac and Australian Conservation Foundation. Stop adani: how a grassroots environmental movement is ticking sdg boxes in australia, Feb 2018. URL https://impakter.com/stop-adani-grassroots-environmental-movement-ticking-“sdg-boxes-australia/. 2.3.1

[167] Xiaojin Zhu and Zoubin Ghahramani. Learning from labeled and unlabeled data with label propagation. *Technical Report CMU-CALD-02-107, Carnegie Mellon University*, 2002. 5.3.2