DIGITAL DEMOCRACY
THE STRUCTURE AND DYNAMICS OF POLITICAL COMMUNICATION IN A LARGE SCALE SOCIAL MEDIA STREAM

Michael D. Conover

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the School of Informatics and Computing
Indiana University
February, 2013

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

Alessandro Flammini, Ph.D.

Filippo Menczer, Ph.D.

Luis Rocha, Ph.D.

Michael Trosset, Ph.D.

February 4th, 2013

ACKNOWLEDGEMENTS

I am forever grateful to my family for their unflagging love and support. Mom & Dad, Rob & Amy: Never am I so whole as when I laugh in your company.

This work would not have been possible without the guidance of my creative, insightful, and above all, patient, faculty mentors, Alessandro Flammini and Filippo Menczer. You taught me how to reason. Thank you.

To my strange and brilliant partner in crime, Jacob Ratkiewicz, I owe more than is fit to print. Your insight made the most challenging problems tractable and your friendship the most tedious days a laugh. To Vigor!

My committee members Michael Trosset and Luis Rocha taught me how to see things differently. The world is clearer and more fascinating for having known you. Thank you.

I am especially thankful to my peers and co-authors for their invaluable contributions to the studies on which this thesis is based. Jacob Ratkiewicz, Bruno Gonçalves, Mark Meiss, Emilio Ferrara, Karissa McKelvey, Clayton Davis & Alex Rudnick: I'm so proud of the work we've done. Thank you.

Finally, I want to express my love for the souls whose laughter travels with me the world over. Mike, Ben, & Andy: You are my brothers. Forty-Five Hi-Fi Lives. Kory & Brandon: For the dust and the heat in my bones. You made the southwest my home. Audra, Julia & Celia: For teaching me how to be a man. Daniel: For the networks yet to be built. Artemy & Seth: A toast! To love, to beauty, and wanderlust! Chris: For the hustle, for travel. Nicola, Bruno & Rossano: Many thanks for the colorful foreign language education. Steve: Your guidance has been a beacon. Erin: For everything. Will: For day one. Azadeh & Laura: Because it's never too late to start cooking.

ABSTRACT

Digital Democracy
The Structure and Dynamics of Political Communication
in a Large Scale Social Media Stream

by

Michael D. Conover

Low cost networked communication technologies have precipitated fundamental changes in the market forces governing the production and consumption of information. These developments, in turn, have shaped the character of political discourse, enabling citizens to engage in anonymous, homophilous, geographically-unconstrained communication with a potentially global audience. Seeking to understand the effects of these changes, this research employs the tools of complex network analysis, text mining, and machine learning to quantify the structure and dynamics of political communication in a high volume social media stream. In doing so, we shed light on the network signatures of automated propaganda campaigns, the polarized nature of domestic political communication, partisan asymmetries in online political activity, and the geospatial structure and temporal evolution of social movement communication networks. Blending quantitative results with theory from the political and social sciences, this work provides a detailed accounting of the structure and dynamics of political communication in a high profile social media stream.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table

CHAPTER I

Introduction

"Science is magic that works."

Kurt Vonnegut

Hidden within the statistical folds of large-scale behavioral trace data are insights that will transform society in fundamental and lasting ways. Generated implicitly by our interaction with the world's technological systems, these data represent an unprecedented opportunity to study society itself, enabling scientists, entrepreneurs, and policy makers to apply the full force of quantitative reasoning to systems whose scale and complexity has been historically intractable.

Such systems are often characterized by behaviors that arise from interactions among large numbers of individuals, a property that confounds the analytical techniques traditionally employed in the study of social dynamics. In many cases, these emergent phenomena owe to interactions that are expressly physical, as in the case of urban gridlock or the crush of a panicked crowd. A more interesting class of interactions, perhaps, involves the transmission of information between agents through a common media, as with the chemical signaling mechanisms of social insects or the collaborative editing of digital encyclopedias. What makes these interactions exceptional is that they are largely temporally and spatially unconstrained. Whereas the extent of an individual's influence on a seething throng is limited to his immediate surroundings, the impact of Vonnegut and Twain extends across space and time as a result of the preservation and dissemination of their writing. Complex processes involving such interactions, while not uniquely human, are for this reason fascinating and their effects on social systems far reaching.

The evolution of human communication has been punctuated by major, transformative changes in society's ability to record and share information [10]. For much of history the spread of information has operated at expressly local scales, with oral traditions and gossip acting as the primary vectors of cultural transmission. Even among the world's literate elite, the difficulty of storing information, whether in ink or stone, rendered it impossible in all but a handful of cases to amass large repositories of knowledge [91]. The invention of the printing press in the 1400's, however, ushered in dramatic decreases in the cost of reproducing and archiving information, and on this development the fate of empire and religion turned.

While movable type made it easy to store and reproduce knowledge, its spread was still largely dominated by local interactions, with physical access to printed materials a requisite for experiencing an author's intended effect. Such resources, not amenable to simultaneous consumption, are said to exhibit *rivalry*, a property characteristic of information diffusion through the dawn of broadcast communication in the late nineteenth century [10]. With the advent of radio and television humanity transcended the physical laws forcing consumers to compete for knowledge, as the marginal cost transmitting an electromagnetic signal to one additional receiver is effectively zero. Moreover, the speed with which information could travel was no longer governed by geographical constraints, with instantaneous, long-distance transmission becoming a facet of everyday life in a span of less than fifty years. Over the course of the twentieth century these changes lead to the rise of a global culture, with the world's population consuming from and adapting to a common pool of information and creative works [10, 38].

Despite these advances, the dividends of broadcast communication were not evenly distributed among the world's people. As with the printing press, content production by means of television and radio required access to financial capital and physical infrastructure. A consequence of this fact was that these markets exhibited high levels of concentration, with a small number of states and corporations exercising control over the most highly visible information channels. From this, it follows that the content of broadcast communication tends to reflect the political and economic interests of these information gatekeepers [10, 94, 42].

With the advent of networked communication technologies, however, such privilege would become rapidly and dramatically devalued. For much of the world's population, low cost computing platforms have effectively eliminated the barriers to entry in the market for information production, catalyzing a host of critical changes in the structure and dynamics of information spreading [10]. Most notably, the volume and diversity of information available to broad swaths of the world's population has increased by orders of magnitude. This increase has led, in turn, to complex, collectively-driven processes for assessing the value and reliability of information resources, functions traditionally performed by the political and economic elite. Moreover, high volume networked storage has decreased even further the cost of archiving and duplicating information, a development central to the availability of high resolution behavioral trace data [39].

Citizens the world over can now engage in critical and creative culture at the local, national, and global level. Judgments made collectively and at a societal scale now function as the primary mechanism by which value is attributed to information, and nowhere are the effects of this development more pronounced than with respect to political communication. Information that would historically had little chance of bypassing mass media filtering mechanisms is now widely available, and, once introduced into the networked ecosystem, all but impossible to erase. What's more, the very properties that enable citizens to engage in public political discourse are those responsible for the internet's capacity as an economic engine, making it tremendously difficult for governments to suppress critical content while still enjoying the benefits of a networked economy [3].

Predictably, these developments have not had an unambiguously positive effect on the character of civic discourse. In this new information ecology, catchiness and sensationalism often function as the principle animating forces behind the spread of political information. This, combined with the ease with which content can be created and shared, sustains an environment in which misinformation propagates quickly and easily, corroding the foundations of a deliberative democracy.

The wealth of information available to citizens, likewise, is at once a force capable of effecting positive social change while simultaneously exacerbating the effects of processes that lead to a fractured, intolerant public. Faced with limitless opportunities for content consumption and social connectivity, individuals tend to prefer information and peers that reinforce their pre-existing beliefs [67]. Such homophilous social connectivity can lead individuals to adopt increasingly extreme positions, eroding the shared frames of reference that support productive political debate [93].

So motivated, this work is centrally concerned with understanding the role of these developments in shaping modern political communication. Chapter II describes the mathematical and conceptual frameworks underlying much of the research presented herein. Chapter III explores how collective filtering and anonymity affect the spread of propaganda and political misinformation. When an an organization, government, or individual can exert centralized, anonymous control over large numbers of accounts, the concern is that they may be able to shape the public discourse surrounding news and policy issues. Using the tools of machine learning and information visualization, we demonstrate that the statistical signatures of communication networks associated with this kind of activity exhibit distinctive properties, a result underscoring the power of analyses based on connectivity rather than content, a theme that finds currency throughout this work.

In Chapter IV we investigate how information abundance and homophilous social connectivity shape American political communication on Twitter. Analyzing communication networks from a six week period preceding a midterm congressional election we reveal a highly segregated, partisan community structure in which left- and right-leaning individuals rarely rebroadcast content produced by ideologically opposed users. While this finding dovetails nicely with existing theoretical and empirical work in this area, we also report on novel results that show unexpectedly frequent interactions among partisans in the form of cross-ideological mentions. To explain the existence of this structure we propose and validate a mechanism of action based on unique communication affordances of the Twitter platform, suggesting that ideological balkanization is not an unavoidable terminal state for political communication online.

Related research on political knowledge discovery, presented throughout Chapters IV & V, explores how the advent of high volume, high resolution information archival can lead to surprising insights into the opinions and behaviors of politically active social media users, specifically with respect to the issues of demographic profiling and political engagement. The ability of campaigns to monitor, in real time, the opinions of large numbers of American voters will undoubtedly have a significant effect on the character of the electoral process. Troublingly however, whether this development will lead to more responsive, insightful campaigns or populist, lowest-common-denominator marketing drivel is as yet unknown.

Finally, we conclude by examining how low cost, long distance communication helped to shape the evolution of the American anticapitalist social movement Occupy Wall Street. Motivated by canonical work in the study of social movement organizations, we find that participants utilize the Twitter platform differently depending on whether they are communicating with individuals who are geographically close as opposed to those across state boundaries. These empirical findings suggest new theoretical avenues in the study of social movements, mirroring the way in which theoretical work on polarization and extremism breath life into the analyses presented in Chapter IV. This paradigm, in which sociological theory suggests computational analyses, and empirical results suggest new sociological theory, represents a virtuous cycle of scholarship which may come to characterize the emerging discipline of computational social science.

These inquiries, in their modest computational and statistical sophistication, represent only early, tentative steps in a broader research program, undertaken by a global scientific community, that will change the course of history. Cities themselves will rise and fall on knowledge gleaned from the study of human mobility; heretofore unimagined disciplines will grow, like so many flowers in spring, from the fertile science of bibliometrics; vast stores of untapped human potential will be realized through the study of economic and intellectual flows operating at a global scale. The past hundred years have borne witness to revolution upon revolution; the story of the coming century is ours to conceive. In mathematics and computation we hold the cipher to our greatest challenges, and with these tools humanity will transform, as a butterfly in its chrysalis – forever and fundamentally changed.

CHAPTER II

Background

"Essentially, all models are wrong, but some are useful."

George E.P. Box

In this section we touch on the core techniques and modeling paradigms that form the methodological foundation for nearly every aspect of this work. As this thesis is concerned primarily with understanding the structure of digital political communication, essentially a function of the relationships among individuals producing text, it follows that we must rely on robust computational models for representing both language and social ties.

2.1 Text Mining

Combining computational approximations of human semantic judgments with the speed and scaling capabilities of automated data analysis we can make deep inferences about features of large document collections that may not be obvious or comprehensible to a human reader.

2.1.1 Representation

The digital representation of language is at the core of the text mining problem, and here we describe one of the oldest and most popular frameworks for addressing this issue, the Salton vector space model. Introduced in 1975, this model represents documents as points in an n-dimensional space. Each of the n dimensions corresponds to a token, and each document is represented by a vector. Consequently, an entire corpus represented by a matrix, M, where M_{ij} represents the relevance of term i to document j.

The most common technique for measuring the relevance of a term to a given document is known as Term Frequency-Inverse Document Frequency (TFIDF). In the context of our study of political communication, TF measures the relative importance of term i in the set of documents produced by an individual j, and is defined as:

$$TF_{ij} = \frac{n_{ij}}{\sum_k n_{k,j}}, \tag{2.1}$$

where n_{ij} is the number of times i occurs in all documents produced by j, and $\sum n_{\cdot j}$ is the total number of terms in all documents produced by individual j. IDF discounts terms with high overall prominence across all political actors, and is defined as

$$IDF_i = \log \frac{|U|}{|U_i|}, \tag{2.2}$$

where U is the set of all individuals, and U_i is the subset of persons who produced term i. A term produced by everyone has no discriminative power and its IDF is zero. The product $TF_{ij} \cdot IDF_i$ measures the extent to which term i occurs frequently in the documents produced by j without occurring in the communication of too many other individuals.

2.1.1.1 Bag of Words

One of the chief concessions of many text analysis techniques, including the Salton Vector Space Model, is the 'bag of words' assumption, which asserts that documents' contents are generated by drawing conditionally independent tokens from some unobserved distribution across the vocabulary of possible terms.

Problems associated with this concession manifest themselves in a number of ways. Principal among them is the loss of information about semantic structure. A canonical example compares the statements 'That cat devoured a mouse.' and 'That mouse devoured a cat.' These are logically distinct statements, but the bag of words model treats them as containing identical information. Despite these drawbacks, the assumption is tenable in practice, as documents with common vocabularies often address the similar subject matter.

2.1.1.2 Synonymy & Polysemy

The Salton model also exhibits limitations with respect to its ability to account for polysemy and synonymy. In the context of this model, polysemy can be understood as conflating in a single axis variation that should be described in terms of multiple independent dimensions. Likewise, synonymy can be interpreted as encoding in multiple different dimensions variation that should exist along a single 'true' axis.

2.1.2 Similarity

Given a mechanism for representing a document corpus it is useful to define a way to quantify relationships between its contents. Two popular approaches to this problem are the measures of cosine and Jaccard similarity [92]. Measuring the cosine of the angle between two document vectors produces a value ranging from -1 to 1. Computed as

$$c(A, B) = \frac{A \cdot B}{|A||B|}, \tag{2.3}$$

where $A \cdot B$ is the dot product of the two document vectors, and $|A|$ and $|B|$ are the lengths of each document vector, respectively. A pair of documents has cosine similarity 1 if they point in exactly the same direction, -1 if they point in exactly opposite directions, and 0 if the are perfectly orthogonal.

The Jaccard coefficient is a somewhat simpler measure than cosine similarity, thought the two are are often highly correlated. The Jaccard coefficient measures the extent to which the sets of terms appearing in two documents tend to overlap, and is computed as

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}. \tag{2.4}$$

Not limited to terms, throughout this work the Jaccard coefficient is employed to measure similarities between many types of sets.

2.1.3 Latent Semantic Analysis

Latent semantic analysis is a text mining technique based on the singular value decomposition (SVD). Mathematically, LSA produces a d-dimensional least-squares best-fit approximation of the original term-document matrix, a representation argued to address issues of polysemy, synonymy, and lexical noisiness [61]. LSA has also been argued to approximate the underlying semantic or conceptual space from which a observations in a corpus are sampled, identifying linear combinations of textual features that correspond to unobserved semantic features, or topics [60].

Latent semantic analysis, through the use of SVD, creates a factorization of the original term-document matrix ($U\Sigma V^t$) that describes two sets of orthogonal basis vectors; the left singular vectors provide a vector basis for terms in the factorized representation (‘eigendocuments’), and the right singular vectors provide an ‘eigenterm’ basis for documents, with the singular values of matrix Σ acting as scaling factors that identify the amount of variance captured by each component.

LSA is argued to capture information about 'not just the summed contiguous pairwise (or tuple-wise) co-occurrences of words, but the detailed patterns of occurrences of very many words over very large numbers of local meaning-bearing contexts' [60]. Returning to the description above, let us focus on the left singular vectors, or 'eigendocument' basis for terms. Because the original term-document matrix, M, is not symmetric, it cannot be diagonalized. However, the symmetric matrix XX^t has a spectral decomposition of $V\Sigma^t\Sigma V^t$, and we observe that the eigenvectors of XX^t are identical to the left singular vectors. This matrix, XX^t, is very similar to the covariance matrix of X, differing only in the respect that the column values have not been centered. [7] Thus, we see that the LSA factorization is based on the eigenvectors of a matrix that describes the co-occurrence patterns of terms and documents in all possible contexts.

2.2 Network Analysis

Graph representations of complex phenomena have been shown to be a useful modeling paradigm in many contexts, from metabolic networks to the large-scale structure of global transportation systems. In its most basic form a network is comprised of a set of vertices, V and a set of edges among them, E, where each vertex represents a single element of the system and edges characterize relationships among components. Mathematically, undirected networks with n nodes can be represented by a symmetric $n \times n$ adjacency matrix, M, where M_{ij} is 1 if nodes i & j share an edge and 0 otherwise. Directed matrices are defined similarly, but do away with the requirement that M be symmetric. If the relationship between nodes is not binary, M is a weighted network, and values of M_{ij} can take any value.

2.2.1 Network Structure

The formulation described above allows for the development of statistics which can be useful for describing the structure of complex networks. At the local level we have degree, the number of edges incident to a given node. In the case of a weighted network, we have node strength, defined as $\sum w_k$, where w_k is the weight of each edge adjacent to a given node.

A series of edges connecting two nodes defines a path between them. Path-based properties include the shortest path length between two nodes, as well as network diameter, the longest length of the shortest path connecting any pair of nodes. Betweenness centrality, another path-based property that plays a role in determining the robustness and resilience of networks, is defined in terms of the number of shortest paths that lie along a given node or edge. Formally, for vertex i

$$b_i = \sum_{h \neq j \neq i} \frac{\sigma_{hj}(i)}{\sigma_{hj}}, \tag{2.5}$$

where σ_{hj} is the total number of shortest paths between nodes h and j and $\sigma_{hj}(i)$ is the number of these paths that pass through i.

The extent to which a local neighborhood exhibits mutual connectivity is characterized by a node's clustering coefficient, which is most commonly defined for a node i in undirected graphs according to the Watts and Strogatz formulation as

$$C_i = \frac{2|\sum e_{jk}, v_j \in N_i|}{k_i(k_i - 1)}, \tag{2.6}$$

where N_i is the set of nodes in the neighborhood of i, k_i is the degree of node i, and e_{jk} represents an edge between node j to node $k \in N_i$ [102].

Aggregating these local attributes across the vertices in many complex networks gives rise to statistical distributions that exhibit interesting properties. Among the most famous of these is the scale-free degree distribution, whereby the probability that a random node has degree k as defined as

$$P(k) \sim ck^{-\gamma}, \tag{2.7}$$

where γ is a value that ranges between 2 and 3 for most networks and c is a normalization constant that ensures the integral of $P(k)$ is one. As a result of this property, scale-free networks exhibit high levels of average resilience to damage, low average clustering coefficient, and a small average shortest path length [8].

2.2.2 Community Detection

The relationships among the vertices of a network can be used to identify communities of nodes that tend to exhibit unexpectedly high levels of in-group assortativity. However, the identification of graph communities shares many challenges with other clustering techniques, most importantly the lack consensus as to what exactly constitutes a cluster. All community-detection algorithms implicitly assert a definition of what constitutes a 'community', and a given mechanism may be appropriate in some situations and not applicable in others.

One popular measure of community structure, Newman-Girvan modularity, compares, for a given partitioning of a network, the number of observed inter-cluster edges to the number that should be expected given a randomly rewired configuration model. For a graph with two communities this value can be computed as

$$Q = \frac{1}{4m}\sum_{ij}(M_{ij} - \frac{k_i k_j}{2m})\delta(s_i, s_j), \tag{2.8}$$

where $m = \frac{1}{2}\sum k_i$, the total number of edges in the network, k_i and k_j are the degrees of node i and j, respectively, and $\delta(s_i, s_j)$ is the Kronecker delta, which is equal to 1 if i and j are in the same community and 0 otherwise. The expression $\frac{k_i k_j}{2m}$ describes the number of edges that should be expected between a pair of nodes in a randomly rewired network, and from this we see how modularity captures the actual deviation from this expected value.

2.2.3 Information Diffusion and Influence

Of central interest to this analysis is the question of how information spreads between individuals as a result of their interactions. However, though processes of social influence and information spreading are readily observed in our day to day interactions, the development of models that correctly represent these processes is a challenging and open area of research.

Early attempts to describe the spread of information employed agent based modeling to study the behavior of idealized systems in which individuals exert conformist pressure on their neighbors [37, 6]. Other approaches have been inspired by the similarity between cultural transmission and epidemiological spreading, with results giving rise to the notion of 'viral' phenomena [14, 52]. It's worth noting, however, that the way in which information spreads through social networks network exhibits important differences from the spread of infectious diseases [70]. For example, where a single exposure to a disease is often sufficient to infect an individual, there is evidence to suggest that a person may require repeated exposures to an idea before adopting it as his or her own [86]. More recent work has employed historical data from large scale social networking systems to study the diffusion of information and processes of social influence [21, 8, 63, 64, 2].

Troublingly, there is evidence to support the notion that it is impossible to distinguish, using such observational data, changes in individuals' behavior owing to homophily versus changes in behavior owing to social influence [90]. Consequently, commercial entities have employed large scale social networking platforms as laboratories for studying these processes in a more methodologically rigorous way. One such study, undertaken by Facebook, exposed one group of 60 million individuals to an interface element indicating which of their friends had voted in the 2010 Congressional elections. Comparing the self-reported voting behavior of these individuals to the behavior of approximately

600,000 users who did not receive the social treatment, the researchers were able to conclude that the presence of a channel for social influence precipitated a small but statistically significant increase in voting activity [17].

2.3 Social Media & Political Activity

2.3.1 The Twitter Platform

Twitter is a popular social networking and microblogging site where users can post 140-character messages containing text and hyperlinks, called *tweets*, and interact with one another in a variety of ways. In the present section we describe several of the platform's key features. Twitter allows each user to broadcast tweets to an audience of users who have elected to subscribe to the stream of content he or she produces. The act of subscribing to a user's tweets is known as *following*, and represents a directed, non-reciprocal social link between two users. From a content consumption perspective, each user can sample tweets from a variety of content streams, including the stream of tweets produced by the users he or she follows, as well as the set of tweets containing specific keywords known as hashtags.

A hashtag is a token prepended with a pound sign (*e.g.*, `#token`) which, when displayed, functions as a hyperlink to the stream of recent tweets containing the specified tag [49]. While they can be used to specify the topic of a tweet (*e.g.*, `#oil` or `#taxes`), when used in political communication hashtags are commonly employed to identify one or more intended audiences, as in the case of the most popular political hashtags, `#tcot` and `#p2`, acronyms for "Top Conservative on Twitter" and "Progressives 2.0," respectively. In this way, hashtags function to broaden the audience of a tweet, extending its visibility beyond a person's immediate followers to include all users who seek out content associated with the tag's topic or audience.

In addition to broadcasting tweets to the public at large, Twitter users can interact directly with one another in two primary ways: retweets and mentions. Retweets often act as a form of endorsement, allowing individuals to rebroadcast content generated by other users, thus raising the content's visibility. Mentions allow someone to address a specific user directly through the public feed, or, to a lesser extent, refer to an individual in the third person.

2.3.2 Data Mining & Twitter

Twitter has evoked tremendous interest from the academic community, warranting several studies focused expressly on the platform and its users [49, 57, 44, 48, 58, 18]. In more applied work, Twitter has been used to study influence [7], the reliability of information sources [22, 85], and social network structure [26, 36]. Additional studies have demonstrated that communication on the Twitter platform is a useful resource for data mining, facilitating, for example, predictions of box office success [5] and the outcomes of political elections [98]. Other work has shown that some correlation exists between the sentiment expressed by Twitter users and important worldwide events, including stock market fluctuations [15, 16].

Another important line of work relates to the identifying information relating to unforseen events as they occur in real time. One group of researchers developed an automated breaking news detection system based on the linking behavior of Twitter users [81], while others have used Twitter data to approximate the epicenter of earthquakes in Japan by treating users as a geographically-distributed sensor network [87]. Still others have used Twitter to study information sharing practices during emergencies [46], with an focus on identifying relevant, actionable information [28, 68].

While its large scale and streaming character make Twitter a useful platform for data mining research, many of these same characteristics have also made it a prime target for spammers, and the detection of spam and automated content production is an active area of research [105, 100, 58].

2.3.3 Political Activity Online

Digitally-mediated communication has become an integral part of the American political landscape, providing citizens access to an unprecedented wealth of information and organizational resources for political activity [11, 10, 94, 32, 3, 98, 75]. So pervasive is the influence of digital communication on the political process that almost one quarter (24%) of American adults got the majority of their news about the 2010 midterm congressional elections from online sources, a figure that has increased three-fold since the Pew Research Center began monitoring the statistic during the 2002 campaign [78]. Relax the constraint that a majority of a person's political news and information must come from online sources and the figure jumps to include the 54% of adult Americans who went online in 2010 to get political information. Critically, this activity precipitates tangible changes in the beliefs and behaviors of voters, with 35% of Internet users who voted in 2010 reporting that political information they saw or read online made them decide to vote for or against a particular candidate [78].

In terms of political organization and engagement, the benefits of social media use are many. For voters, social media make it easier to share political information, draw attention to ideological issues, and facilitate the formation of advocacy groups with low barriers to entry and participation [97, 34]. The ease with which individual voters can connect with one another directly also makes it easier to aggregate small-scale acts, as in the case of online petitions, fundraising, or web-based phone banking [59]. Together, these features contribute to the widespread use of social media for political purposes among the voting public, with as many as 21% of online adults using social networking sites to engage with the 2010 congressional midterm elections [80]. Moreover, a survey by the Pew Internet and American Life Project finds that online political activity is correlated with more traditional forms of political participation, with individuals who use blogs or social networking sites as a vehicle for civic engagement being more likely to join a political or civic group, compared to other Internet users [79].

2.3.4 Data Mining and Political Speech

Predictably, formal political speech and activity have been a target for data mining applications. The seminal work of Poole and Rosenthal applied multidimensional scaling to congressional voting records to quantify the ideological leanings of members of the first 99 United States Congresses [82]. Similar work by Thomas *et al.* used transcripts of floor debates in the House of Representatives to predict whether a speech segment was provided in support of or opposition to a specific proposal [96]. Related efforts have been undertaken for more informal, web-based political speech, such as that found on blogs and blog comments [31, 30].

2.4 Overview

Together these tools constitute the analytical foundation on which we base our study of political communication on Twitter. Graph theory allows us to examine the relationships among large numbers of individuals, enabling us to quantify the structure of their social and communication networks. Similarly, text and data mining techniques allow us to make quantitative inferences about the qualitative character of these individuals' speech. Combined with theory from the political and social sciences, these statistical and computational techniques help shed light on the structure and dynamics of human communication.

CHAPTER III

Propaganda & Misinformation

"In the case of news, we should always wait for the sacrament of confirmation."

Voltaire

For all the drawbacks of mass media communication, historically, the press have played an important role in establishing the credibility of information and sources on which they report. On social media, however, the role of such centralized gatekeepers has been greatly diminished. Instead, the fact checking process takes place in a distributed manner, with content consumers and producers negotiating a shared consensus about information credibility and value. While the benefits of this process are numerous, the ease with which individuals can propagate information they encounter online virtually ensures that ' or misleading information has the potential to reach a very large audience.

Complicating this problem is the fact that lowered barriers to entry in the content production market have made it significantly easier to engage in effectively anonymous communication. Although a concerted effeort can almost invariably reveal the true identity of a content producer, for most users it is prohibitively difficult to verify the identity of an information source. Consequently, a motivated party can create large numbers of centrally controlled accounts, forging the illusion of consensus or credibility, in hopes that the content promoted by those accounts will be widely rebroadcast by an unwitting audience. In this chapter, we describe two complementary approaches to the problem of identifying behavioral signatures associated with this type of activity, one based on machine learning techniques and the other on visualization and statistical summarization.

While motivations for engaging such persuasion campaigns can vary widely, it is clear that social media content can exhibit significant influence on the discourse surrounding public figures and policy issues. An illustrative example can be drawn from the 2010 midterm election, when several major news organizations picked up on a viral tweet relating to the allocation of stimulus funds, succinctly describing a study of decision making in drug-addicted macaques as "Stimulus $$$ for coke monkeys" [95].

While the "coke monkeys" meme developed organically from the attention dynamics of thousands of users, the work of Mustafaraj & Metaxas suggests that, with some measure of luck, a motivated attacker can initiate a similar information cascade by deliberately employing deceptive communication techniques. In their 2009 article, Mustafaraj & Metaxas describe a concerted, deceitful attempt to cause a specific URL to rise to prominence on Twitter through the use of a network of nine fake user accounts. These accounts produced 929 tweets over the course of 138 minutes, all of which included a link to a website smearing one of the candidates in the 2009 Massachusetts special election. Targeting the tweets to hundreds of users who had previously expressed interest in the election, the initiators sought to trigger an information cascade that would lend a sense of credibility and grassroots enthusiasm to a specific political message. Within hours, a substantial portion of the targeted users retweeted the link, resulting in significant spreading that was quickly detected by Google's realtime search machinery. As a result, in the days immediately preceding the election, the URL in question was promoted to the top of the Google results page for queries containing the candidate's name — a so-called *Twitter bomb*.

While both of these examples demonstrate the ability of social media content to influence a public well beyond the confines of an individual social network, we are specifically interested in intentionally deceptive activity similar to that described by Musafaraj & Metaxas. Some features of this activity, such as the mass creation of accounts, user impersonation, and the posting of deceptive content are behaviors common to both spam and political astroturf. However, political astroturf is not exactly the same as spam. While the primary objective of a spammer is often to persuade users to click a link, someone interested in promoting an astroturf message wants to establish a false sense of group consensus about a particular idea. Related to this process is the fact that users are more likely to believe a message that they perceive as coming from several independent sources, or from an acquaintance [47]. Spam detection systems often focus on the content of a potential spam message — for instance, to see if the message contains a certain link or set of tags. In detecting political astroturf, we focus on *how* the message is delivered rather than on its content. Further, many legitimate users may be unwittingly complicit in the propagation of astroturf, having been themselves deceived. Spam detection methods that focus solely on properties of user accounts, such as the number of URLs in tweets from an account or the interval between successive tweets, may therefore be unsuccessful in finding such abuse.

For the purposes of the work described in this chapter, we adopt the term *truthy* to discriminate falsely-propagated information from organic grassroots memes. The term was coined by comedian Stephen Colbert to describe something that a person believes based on emotion rather than facts. We can then define our task as the detection of truthy memes in the Twitter stream. Not every truthy meme will result in a viral cascade like the one documented by [71], but we wish to test the hypothesis that the initial stages exhibit identifiable signatures.

3.1 Analytical Framework

Before discussing the computational and mathematical infrastructure of the Truthy system, we first provide an overview of the conceptual models used for the data under study.

3.1.1 Meme Types

To study the diffusion of information on Twitter it is necessary to identify a specific topic, or 'meme,' as it propagates through the social substrate. While there exist sophisticated statistical techniques for modeling the topics underlying bodies of text, the small size of each tweet and context drift frequently present in high-throughput data streams create significant complications [101]. Fortunately, several conventions employed by Twitter users, namely hashtags, mentions, and embedded hyperlinks, reveal useful information about the topical content of a tweet. In addition to these platform-specific identifiers, we also consider the entire text of the tweet itself, once all Twitter metadata, punctuation, and URLs have been removed. Relying on these conventions we are able to focus on the ways in which a large number of memes propagate through the Twitter social network. It's important to note that a tweet may be included in several of these categories, for example a tweet containing two hashtags and a URL would count as a member of each of the three resulting memes.

3.1.2 Network Edges

To represent the flow of information through the Twitter community, we construct a directed graph in which nodes are individual user accounts. An edge is drawn from node A to B when either B is observed to retweet a message from A, or A mentions B in a tweet. The weight of an edge is increased each time we observe an event connecting two users. In this way, either type of edge can be understood to represent a flow of information from A to B. Observing a retweet at node B provides implicit confirmation that information from A appeared in B's Twitter feed, while a mention of B

originating at node A explicitly confirms that A's message appeared in B's Twitter feed.

Retweet and reply/mention information parsed from the text can be ambiguous, as in the case when a tweet is marked as being a 'retweet' of multiple people. Rather, we rely on Twitter metadata, which designates the user replied to or retweeted by each message. Thus, while the text of a tweet may contain several usernames, we draw an edge only to the user explicitly designated in the metadata. Note that this is separate from our use of mentions as memes, which we parse from the text of the tweet.

3.2 System Architecture

Based on the analytical framework described above we developed a data management and processing pipeline and in this section we describe several of its core elements. These components monitor the Twitter data stream, collect tweets matching themes of interest, detect relevant memes, and produce statistical features characterizing the structure of meme diffusion networks. These statistical features are used as the input to the classification and visualization infrastructure, and together with the components described above are collectively known as "Truthy."

3.2.1 Meme Detection

The task of the meme detection component is the identification of sets of tweets that (a) contain content related to the political elections, and (b) are of sufficiently general interest. We satisfy both of these criteria in the following way. To identify politically relevant tweets, we used a hand-curated collection of approximately 2500 keywords relating to the 2010 U.S. midterm elections. This keyword list contains the names of all candidates running for federal office, as well as any common variations, known Twitter account usernames, and many popular political hashtags. We consider as a potentially interesting 'candidate' tweet any meme that co-occurs with a keyword in this list, thus enabling the discovery of political meme identifiers not known to our system *a priori*. To address the second criteria we rely on a parameterized filtering mechanism that tracks the rate at which each candidate is observed during the course of the previous hour. If the occurrence rate rises above a pre-specified threshold, all tweets containing that meme are automatically stored in the Truthy database. If the rate at which we observe a meme drops below the parametric threshold, the meme is no longer considered active and we stop storing its associated tweets in the Truthy database.

3.2.2 Network Analysis

To characterize the structure of each meme's diffusion network we compute several statistics based on the topology of the largest connected component of the retweet/mention graph. These include the number of nodes and edges in the graph, the mean degree and strength of nodes in the graph, mean edge weight, mean clustering coefficient across nodes in the largest connected component, and the standard deviation and skew of each network's in-degree, out-degree and strength distributions. Additionally we track the out-degree and out-strength of the most prolific broadcaster, as well as the in-degree and in-strength of the most focused-upon user. We also monitor the number of unique injection points of the meme, reasoning that organic memes (such as those relating to news events) will be associated with larger number of originating users.

3.2.3 Sentiment Analysis

We also utilize a modified version of the Google-based Profile of Mood States (GPOMS) sentiment analysis method [15] in the analysis of meme-specific sentiment on Twitter. The GPOMS tool assigns to a body of text a six-dimensional vector with bases corresponding to different mood attributes (*Calm, Alert, Sure, Vital, Kind,* and *Happy*). To produce scores for a meme along each of the six dimensions, GPOMS relies on a vocabulary taken from an established psychometric evaluation instrument extended with co-occurring terms from the Google n-gram corpus. We applied the GPOMS methodology to the collection of tweets, obtaining a six-dimensional mood vector for each meme.

3.3 Visualization

One of the major challenges facing this research program relates to producing training data on which to experiment. We argue that interactive visualization tools are a natural solution to this problem as they surface large volumes of information about a system, allowing users to make rapid judgements about large numbers of examples.

Within the visualization interface we developed for this purpose, Truthy, collections of related memes are algorithmically grouped into top-level categories ('themes') representing the most coarse-grained level of analysis available on the platform. Within a given theme, users can search for memes containing specific keywords or sort content based a variety of statistical features. Navigating a theme, users are presented with a concise visual representation of each meme, characterized in terms of a multiplex information diffusion network (Figure 3.2) and sparklines representing activity volumes over time.

At the level of a specific meme the user is presented with a high-resolution image of the meme's information diffusion network and a variety of statistics about the activity and connectivity of users who have produced content associated with that meme (Figure 3.3.) These features include the number of users and tweets associated with the meme, diffusion network statistics such as mean degree and largest connected component size, and user-specific statistics such as the most retweeted user and the number of unique injection points for the meme. Additionally, users can interact with a zoomable historical timeseries of activity volume and produce animations of relevant meme-meme co-occurrence patterns.

Together these interface elements provide multiple, complementary perspectives on the activity associated with clusters of related content. In practice, this visual analytics interface found use in two primary ways. The first and most obvious was as a tool for inspecting the features associated with individual memes as they arose in different contexts. For example, during the development of the training and evaluation dataset describe in Section 3.4 it was essential to be able to inspect the network structure, content, and temporal activity associated with individual memes to make an informed decision about its trustworthiness. Consider the case of automated accounts that produce content on a regular interval-based schedule. Such cyclical activity patterns that are immediately evident upon inspection of the meme's tweet volume time series would be substantially more difficult to identify through the inspection of timestamped tweet events alone.

The second use case relates to providing high-level insights into non-obvious classes of activity. For example, Figure 3.4 shows two diffusion networks associated with the high-profile @whitehouse and @michelleobama accounts. Common to both of these networks is a dense cluster of inbound mentions targeting the account in question and a broad outbound cascade of retweets and mentions generated by discussion and sharing activity related to these accounts. In the case of @michelleobama, we can see another dense cluster of mentions, surrounding what is likely a politically related account, potentially @barackobama. Similarly, inspecting Figure 3.5 we see a dense network of interior nodes that target peripheral nodes with mentions promoting a nightclub event. By creating the appearance of social interconnectedness, these presumably centrally controlled accounts are able to lend an aura of credibility to their advertising campaign. In both of these examples we can immediately see high-level network topologies that may be difficult to identify in more coarse-grained summary statistics.

The final validation of visualization as an approach to the study of information diffusion in large scale social media streams comes from Figure 3.6. Originally observed in the Truthy theme-level overview (Figure 3.2,) these network layouts all exhibit a characteristic two cluster structure. Specifically, we see two communities of individuals who tend to retweet one another preferentially, but who engage across the community divide using mentions. As these networks represent political communication, a natural hypothesis that emerges from this observation is that these clusters correspond to users from the political left and right. Explored in detail in Chapter IV, this process is characteristic of the way in which visualization and statistical summarization can catalyze the generation of theoretically-grounded hypotheses about complex sociological phenomena.

3.4 Automatic Classification

In the previous section we saw, visually, how certain memes have network structures that appear suspicious. In this section, we describe the development of a hand-curated training and evaluation dataset, the results two classification apparatuses tasked with identifying truthy content, and the qualitative interpretation of the most discriminative features associated with truthy communication.

For the purpose of training and evaluating our machine learning apparatus we developed a custom dataset that partitions political memes into three classes — 'truthy,' 'legitimate,' and 'remove.' To accomplish this partitioning, random political memes were presented to multiple human reviewers who were asked to place each meme in one of the three categories. A meme was classified as 'truthy' if a significant portion of the users involved in that meme appeared to be spreading it in misleading ways — e.g., if a number of the accounts tweeting about the meme appeared to be robots or sock puppets, the accounts appeared to follow only other propagators of the meme (clique behavior), or the users engaged in repeated reply/retweet exclusively with other users who had tweeted the meme. 'Legitimate' memes were described as representing normal use of Twitter — several non-automated users conversing about a topic. The final category, 'remove,' was used for memes in a non-English language or otherwise unrelated to U.S. politics (`#youth`, for example). These memes were not used in the training or evaluation of classifiers.

Upon collecting annotations for 252 memes, we found a significant imbalance in our labeled data (231 legitimate and only 21 truthy). Rather than simply resampling from the smaller class, as is common practice in the case of class imbalance, we performed a second round of human annotations on previously-unlabeled memes predicted to be 'truthy' by the classifier trained in the previous round, gaining 103 more annotations (74 legitimate and 40 truthy). We note that the human classifiers knew that the additional memes were possibly more likely to be truthy, but that the classifier did not perform well at this point due to the paucity of training data and indeed was often contradicted by the human classification. This bootstrapping procedure allowed us to manually label a larger portion of truthy memes with less bias than resampling. Our final training dataset consisted of 366 training examples — 61 'truthy' memes and 305 legitimate ones. In a few cases where multiple reviewers disagreed on the labeling of a meme, we determined the final label by reaching consensus in a group discussion among all reviewers.

As comparing different learning algorithms is not our goal, we report on the results obtained with just two well-known classifiers: AdaBoost with DecisionStump, and SVM [41] [41]. Each classifier was provided with 31 features describing each meme, as shown in Table 3.1. Measures relating to 'degree' and 'strength' refer to the nodes in the diffusion network of the meme in question — that is, the number of people that each user retweeted or mentioned, and the number of times these connections were made, respectively. We defined an 'injection point' as a tweet containing the meme which was not itself a retweet; our intuition was that memes with a larger number of injection points were more likely to be legitimate. For this experiment none of the features were normalized.

As the incidence of truthy memes was well below that of legitimate ones we also experimented with resampling the training data to balance the classes prior to classification. The performance of the classifiers is shown in Table 3.2, as evaluated by their accuracy and the area under their ROC curves (AUC). The latter is an appropriate evaluation measure in the presence of class imbalance. In all cases these preliminary results are quite encouraging, with accuracy around or above 90%. The best results are obtained by AdaBoost with resampling: better than 96% accuracy and 0.99 AUC. Table 4.10 further shows the confusion matrices for AdaBoost. In this task, false negatives (truthy memes incorrectly classified as legitimate, in the upper-right quadrant of each matrix) are less desirable than false positives (the lower-left quadrant). In the worst case, the false negative rate is 4%. Given the relatively low dimensionality of the space and high initial classification accuracy, we did not perform any feature selection or other optimization; the classifiers were provided with all the features computed for each meme (Table 3.1).

Table 3.4 shows the 10 most discriminative features, as determined by χ^2 analysis. One sees that mean weight and mean strength are highly discriminative features, along with the number of edges present in the network. This result is sensible, as many of the memes in the truthy class exhibit multiple unconnected injection points with few interactions among actors producing otherwise identical text. Consequently, these networks have very low mean edge weight and strength, leading to these being highly discriminative features.

Table 3.1: **Features used in truthy classification.**

`nodes`	Number of nodes
`edges`	Number of edges
`mean_k`	Mean degree
`mean_s`	Mean strength
`mean_w`	Mean edge weight in largest connected component
`max_k(i,o)`	Maximum (in,out)-degree
`max_k(i,o)_user`	User with max. (in,out)-degree
`max_s(i,o)`	Maximum (in,out)-strength
`max_s(i,o)_user`	User with max. (in,out)-strength
`std_k(i,o)`	Std. dev. of (in,out)-degree
`std_s(i,o)`	Std. dev. of (in,out)-strength
`skew_k(i,o)`	Skew of (in,out)-degree distribution
`skew_s(i,o)`	Skew of (in,out)-strength distribution
`mean_cc`	Mean size of connected components
`max_cc`	Size of largest connected component
`entry_nodes`	Number of unique injections
`num_truthy`	Number of times 'truthy' button was clicked
sentiment scores	Six GPOMS sentiment dimensions

3.5 Discussion

In this chapter we have detailed the infrastructure and data models of the Truthy data management and processing pipeline, as well as the computational and analytical underpinnings of its two principle applications: an automated system for detecting political astroturf and a visual analytics interface for the inspection of high-throughput social media content. Our simple classification system was able to accurately detect 'truthy' memes based on features extracted from the topology of the diffusion networks. Though few of these exhibit the explosive growth characteristic of true viral memes, they are nonetheless clear examples of coordinated attempts to deceive Twitter users. Truthy memes are often spread initially by bots, causing them to exhibit, when compared with organic memes, pathological diffusion graphs. These networks exhibit a number of signature features, including high numbers of unique injection points with few or no connected components, strong star-like topologies characterized by high average degree, and most tellingly large edge weights between dyads.

The visual analytics interface also provided insight into structural regularities at a system-wide level. Using the theme-level view, we were able to identify a biclustered network structure characteristic of many popular political hashtags, and though a thorough analysis of the origin and character of this structure required significant analytical investment, the high-level insight that spawned the research described in Chapter IV was a direct result of this visualization platform. Together, these applications illustrate the way that computational techniques can be used to produce actionable, meaningful insight into the content and structure of large-scale social media streams.

Table 3.2: **Performance of two classifiers with and without resampling training data to equalize class sizes.** All results are averaged based on 10-fold cross-validation.

Classifier	Resampling	Accuracy	AUC
AdaBoost	No	92.6%	0.91
AdaBoost	Yes	96.4%	0.99
SVM	No	88.3%	0.77
SVM	Yes	95.6%	0.95

Table 3.3: **Confusion matrices for a boosted decision stump classifier with and without resampling.** The labels on the rows refer to true class assignments; the labels on the columns are those predicted.

	No resampling		With resampling	
	Truthy	Legitimate	Truthy	Legitimate
T	45 (12%)	16 (4%)	165 (45%)	6 (1%)
L	11 (3%)	294 (80%)	7 (2%)	188 (51%)

Table 3.4: **Top 10 most discriminative features, according to a χ^2 analysis under 10-fold cross validation.** Intervals represent the variation of the χ^2 or rank across the folds.

Feature	χ^2	Rank
`mean_w`	230 ± 4	1.0 ± 0.0
`mean_s`	204 ± 6	2.0 ± 0.0
`edges`	188 ± 4	4.3 ± 1.9
`skew_ko`	185 ± 4	4.4 ± 1.1
`std_si`	183 ± 5	5.1 ± 1.3
`skew_so`	184 ± 4	5.1 ± 0.9
`skew_si`	180 ± 4	6.7 ± 1.3
`max_cc`	177 ± 4	8.1 ± 1.0
`skew_ki`	174 ± 4	9.6 ± 0.9
`std_ko`	168 ± 5	11.5 ± 0.9

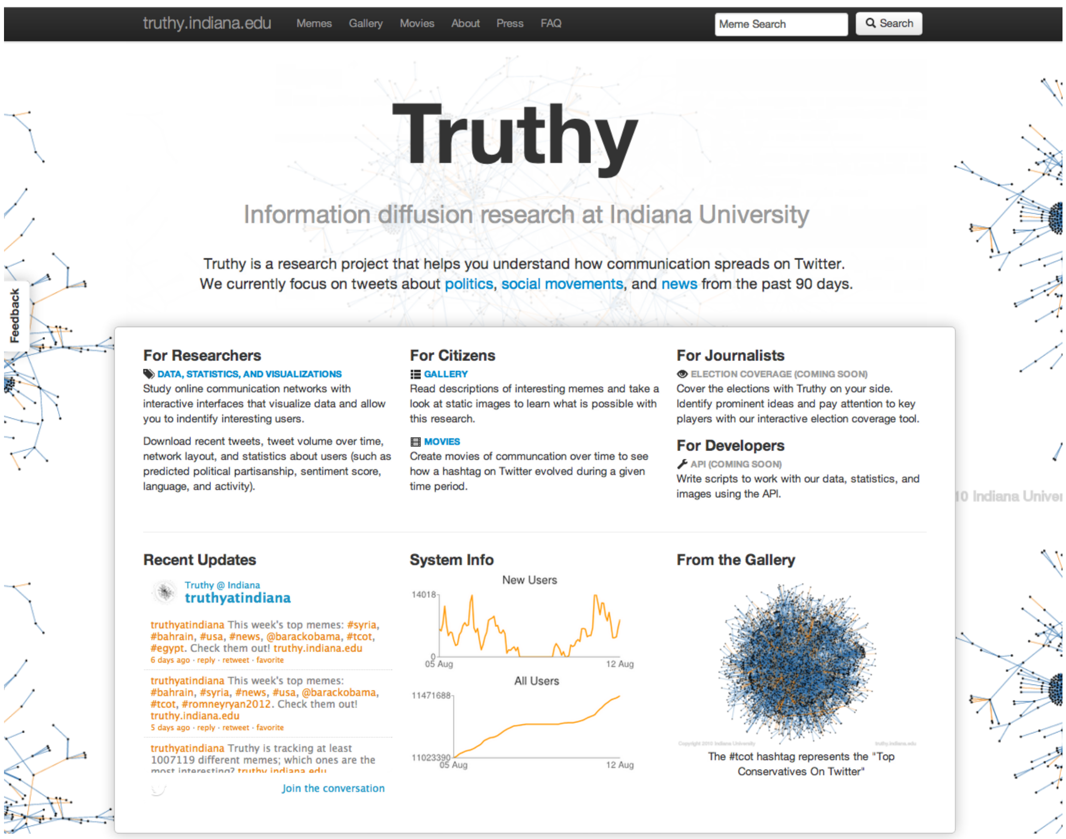

Figure 3.1: **Landing page for *truthy.indiana.edu*.** Users are presented with system-level statistics and a slideshow of popular information diffusion networks.

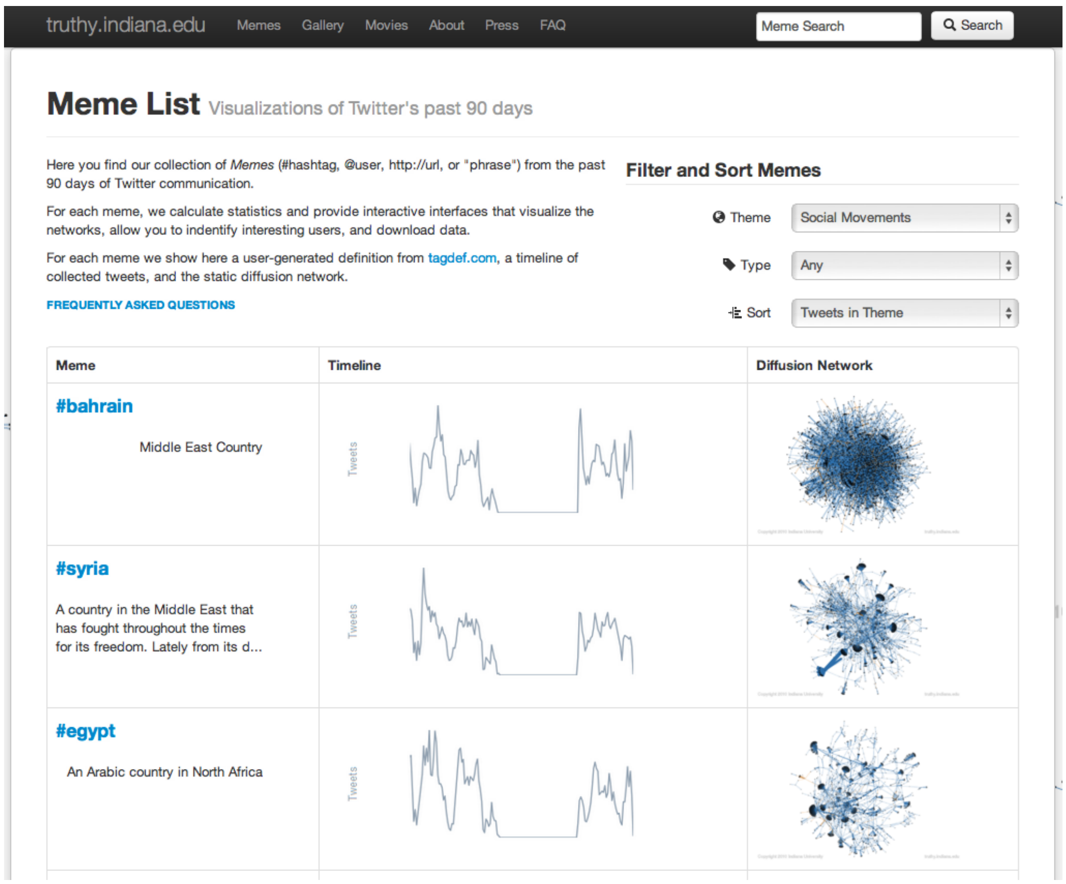

Figure 3.2: **A high-level visual presentation of multiple memes related to a single theme.** This interface provides the ability to sort and filter based on criteria relating to the meme's diffusion characteristics, and includes spark lines describing activity volumes and multiplex network diagrams.

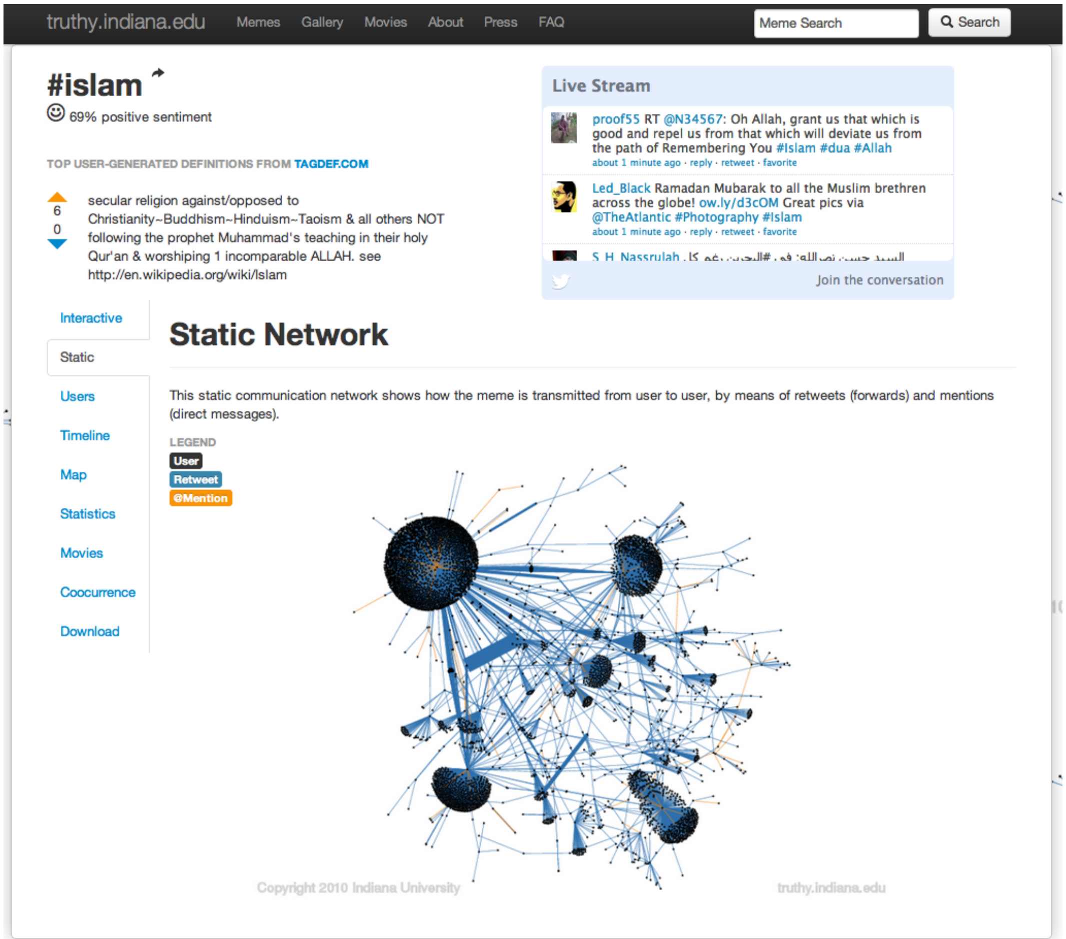

Figure 3.3: **Multi-faceted, detailed accounting of the activity and statistics associated with an individual meme tracked by the Truthy system.** Users can inspect accounts associated with the meme, and interactive network diagram, time series and geospatial data.

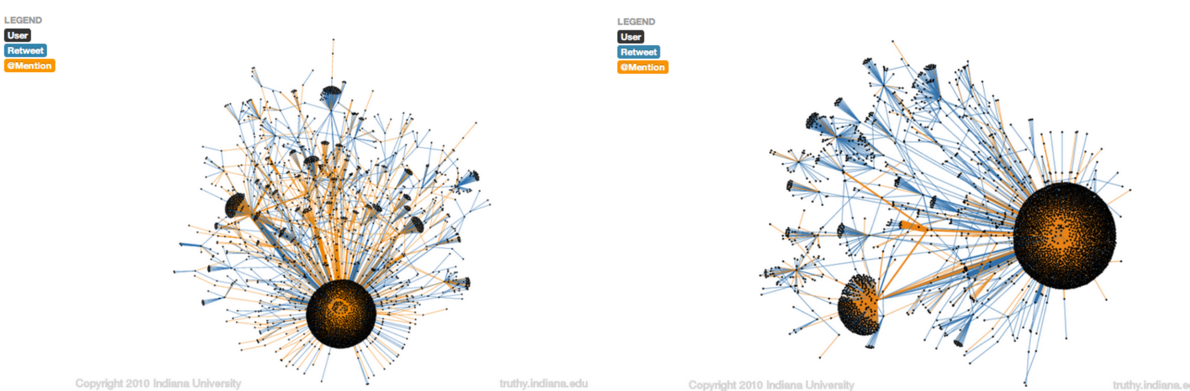

Figure 3.4: **Information diffusion networks for @whitehouse (left) and @michelleobama (right).** These networks both exhibit structural features that are characteristic of activity related to high-profile Twitter users and public figures. Notice the broad outbound retweet cascades and high volume of inbound-mentions.

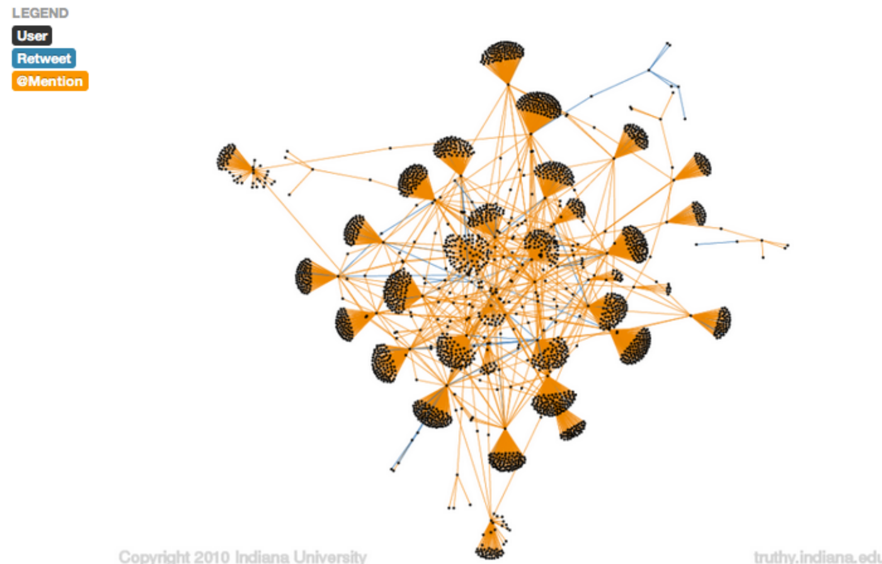

Figure 3.5: **Information diffusion networks for #rsvp, a meme used to promote special events at a Miami nightclub.** While the content of tweets produced by the accounts in this network are diverse, visual inspection immediately suggests suspicious activity. Notice the highly interconnected subgraph of hub accounts, all of which target large numbers of peripheral accounts with mentions containing the promotion.

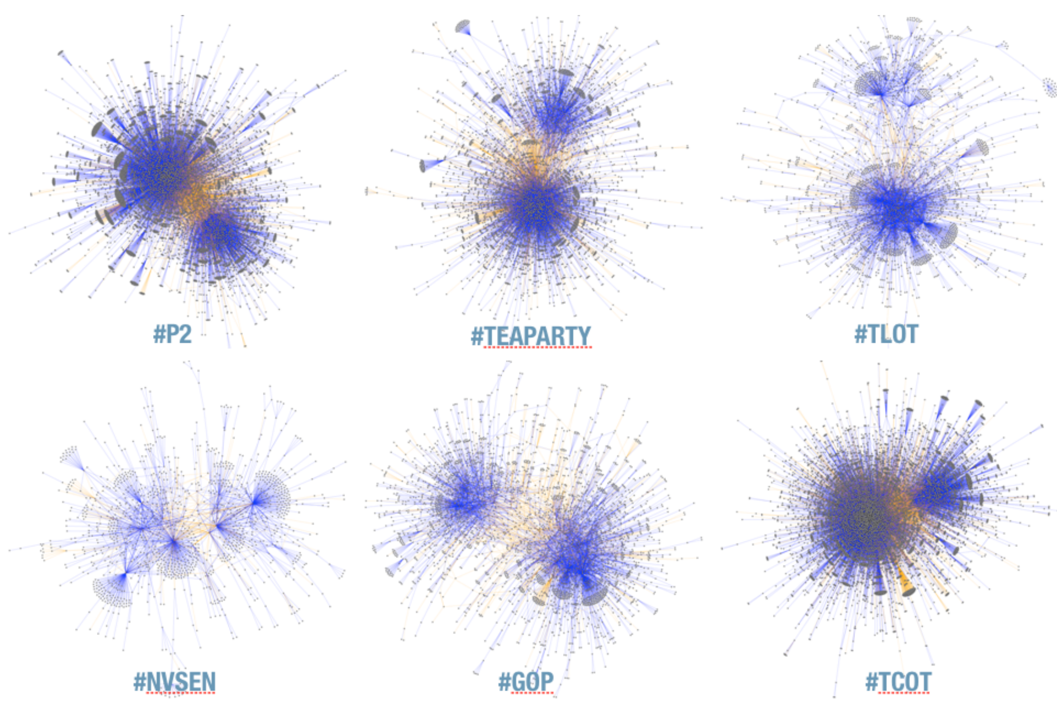

Figure 3.6: **Information diffusion networks for political memes with clear biclustered structure.** One of the key contributions of the Truthy visualization infrastructure was the ability to easily compare and identify patterns in large numbers of political memes. The insights gleaned from using the tool in this way were a direct catalyst for the research described in Chapter IV.

CHAPTER IV

Polarization & Political Knowledge Discovery

"Democracy is a device that ensures we shall be governed no better than we deserve."

George Berdnard Shaw

In Chapter III we dealt with the automated identification of degenerate behaviors resulting from anonymous, low-cost networked communication. Here we turn our attention to understanding how collective filtering and vanishing archival costs affects political communication online. The first portion of this chapter deals with how general interest intermediaries like print and television outlets have been supplanted in large part by distributed, social information sharing processes, a development that fosters homogeneous, ideologically polarized content ecosystems. The second portion deals with how high-resolution behavioral trace data describing a person's communication activity can be leveraged to produce insights into the opinions and behaviors of large numbers of political actors.

4.1 Polarization

Despite the benefits associated with heightened levels of online political engagement, some empirical evidence suggests that politically active web users tend to organize into insular, homogenous communities segregated along partisan lines. Adamic and Glance famously demonstrated that political blogs preferentially link to other blogs of the same political ideology [1], a finding supported by the work of Hargittai and boyd[104]. Consumers of online political information tend to behave similarly, choosing to read blogs that share their political beliefs, with 26% more users doing so in 2008 than 2004 [77].

In its own right, the formation of online communities is not necessarily a serious problem. The concern is that when politically active individuals can avoid people and information they would not have chosen in advance, their opinions are likely to become increasingly extreme as a result of being exposed to more homogeneous viewpoints and fewer credible opposing opinions. The implications for the political process in this case are clear. A deliberative democracy relies on a broadly informed public and a healthy ecosystem of competing ideas. If individuals are exposed exclusively to people or facts that reinforce their preexisting beliefs, democracy suffers [93, 94].

4.1.1 Data & Methods

In this study we examine networks of political communication on the Twitter microblogging service during the six weeks prior to the 2010 U.S. midterm elections. Sampling data from the Twitter 'gardenhose' API, we identified 250,000 politically relevant messages (*tweets*) produced by more than 45,000 users. From these tweets we isolated two networks of political communication — the *retweet* network, in which users are connected if one has rebroadcast content produced by another, and the *mention* network, where users are connected if one has mentioned another in a post, including the case of tweet replies.

4.1.2 Identifying Political Content

Let us define a political communication as any tweet containing at least one politically relevant hashtag. To identify an appropriate set of political hashtags and to avoid introducing bias into the sample, we performed a simple tag co-occurrence discovery procedure. We began by seeding our sample with the two most popular political hashtags, `#p2` ("Progressives 2.0") and `#tcot` ("Top Conservatives on Twitter"). For each seed we identified the set of hashtags with which it co-occurred in at least one tweet, and ranked the results using the Jaccard coefficient. For a set of tweets S containing a seed hashtag, and a set of tweets T containing another hashtag, the Jaccard coefficient between S and T is

$$\sigma(S,T) = \frac{|S \cap T|}{|S \cup T|}. \tag{4.1}$$

Thus, when the tweets in which both seed and hashtag occur make up a large portion of the tweets in which either occurs, the two are deemed to be related. Using a similarity threshold of 0.005 we identified 66 unique hashtags (Table 4.1), eleven of which we excluded due to overly-broad or ambiguous meaning (Table 4.2). This process resulted in a corpus of 252,300 politically relevant tweets. There is substantial overlap between streams associated with different political hashtags because many tweets contain multiple hashtags. As a result, lowering the similarity threshold leads to only modest increases in the number of political tweets in our sample — which do not substantially affect the results of our analysis — while introducing unrelated hashtags.

4.1.3 Representativeness

In Section 4.1.2 we identified many high-profile political hashtags, and with them the majority of tweets and users associated with domestic political communication on Twitter. Supporting this claim, Figure 4.1 shows a roughly exponential decay in hashtag popularity as measured in terms of number of users or tweets associated with the hashtag. This sharp decay in the tag popularity indicates that the inclusion of additional political hashtags is not likely to substantially increase the size or alter the structure of the corpus.

This claim is also supported by Figure 4.2, which shows that there is a strong effect of diminishing returns with respect to the observed number of unique users and tweets as the number of hashtags included in our analysis increases. This effect is due to the fact that many tweets are annotated with multiple hashtags, and many users utilize several different hashtags over the course of the study period. As a result, the inclusion of a single hashtag may result in the inclusion of many tweets and users also redundantly associated with other hashtags.

To further support the claim that sampling based on this set of hashtags produces a representative set of political tweets, we selected all the tweets in the gardenhose from the study period that included any one of 2500 hand-selected political keywords related to the 2010 elections [85]. We considered only the $312,560$ tweets in this set containing a hashtag because we use this characteristic to define public political communication on Twitter. We found that 26.4% of these tweets are covered by our target set of hashtags. Furthermore, among the ten most popular hashtags not included in our target set (`#2010memories`, `#2010disappointments`, `#ff`, `#p2000`, `#2010`, `#business`, `#uk`, `#newsjp`, `#asia`, `#sports`), only one is explicitly political and its volume accounts for less than 2% of public political communication. This coverage confirms that we have isolated a substantial and representative sample of political communication on Twitter.

4.1.4 Political Communication Networks

From the tweets containing any of the politically relevant hashtags we constructed networks representing political communication among Twitter users. Focusing on the two primary modes of public user-user interaction, mentions and retweets, we define communication links in the following ways. In the retweet network an edge runs from a node representing user A to a node representing user B if B retweets content originally broadcast by A, indicating that information has propagated from A to B. In the mention network an edge runs from A to B if A mentions B in a tweet, indicating that information may have propagated from A to B (a tweet mentioning B is visible in B's *timeline*). Both networks therefore represent potential pathways for information to flow between users.

The retweet network consists of 23,766 non-isolated nodes among a total of 45,365. The largest connected component accounts for 18,470 nodes, with 102 nodes in the next-largest component. The mention network is smaller, consisting of 10,142 non-isolated nodes out of 17,752 total. It has 7,175 nodes in its largest connected component, and 119 in the next-largest. Because of their dominance we focus on the largest connected components for the rest of our analysis. We observe that the retweet and mention networks exhibit very similar scale-free topology (power-law degree distribution not shown), with a number of users receiving or spreading a huge amount of information.

4.1.5 Cluster Analysis

Initial inspection of the retweet network suggested that users preferentially retweet other users with whom they agree politically, while the mention network appeared to form a bridge between users of different ideologies. We explore this hypothesis in several stages. In § 4.1.6 we use network clustering algorithms to demonstrate that the retweet network exhibits two highly segregated communities of users, while the mention network does not. Finally, in § 4.2.1, by manually annotating users, we show that the retweet network is polarized on a partisan basis, while the mention network is much more politically heterogeneous.

4.1.6 Community Structure

To establish the large-scale political structure of the retweet and mention networks we performed community detection using a label propagation method for two communities.[1] Label propagation [83] works by assigning an initial arbitrary cluster membership to each node and then iteratively updating each node's label according to the label that is shared by most of its neighbors. Ties are broken randomly when they occur. Label propagation is a greedy hill-climbing algorithm. As such it is extremely efficient, but can easily converge to different suboptimal clusters dependent on initial label assignments and random tie breaking. To improve its effectiveness and stability, we seeded the algorithm with initial node labels determined by the leading-eigenvector modularity maximization method for two clusters [73].

To confirm that we can produce consistent clusters across different runs we executed the algorithm

[1] While the partisan nature of U.S. political discourse makes two a natural number of clusters, in § 4.2.1 we describe the effect on our analysis of increasing the target number of communities.

one hundred times for each network and compared the label assignments produced by every run. Table 4.3 reports the high average agreement between the resulting cluster assignments for each graph, as computed by the Adjusted Rand Index [45]. Such a high agreement suggests that the clusters are consistent, and therefore we avoid resorting to consensus clustering for simplicity.

Figure 4.4 shows the retweet and mention networks, laid out using a force-directed layout algorithm [33], with node colors determined by the assigned communities. The retweet network exhibits two distinct communities of users, while the mention network is dominated by a single massive cluster of interconnected users. Modularity [74] resulting from the cluster assignments offers a first measure of segregation, and reinforces the qualitative finding above. The modularity induced by the communities in the retweet and mention networks have values of 0.48 and 0.17, respectively.

A direct comparison of the modularity values is however problematic because of the different size and overall connectivity of the two networks. We need a way to compare the 'goodness' of cluster assignments across different graphs. To this end we generate, for both retweet and mention graphs, $N = 1000$ *shuffled* versions of the graph that preserve the original *degree sequence*.

Each randomized network is clustered with the method described above for the original graphs and associated with the resulting modularity value. We use the distribution of these values as a baseline against which to compare the quality of the clusters in the original graph. The intuition behind this approach is that the degree to which the actual graphs are more modular than the shuffled graphs tells us how amenable each is to being split into two clusters — a measure of segregation. The modularities of the shuffled graphs can be viewed as observed values of a random variable. We can use these values to compute z-scores for the modularities of the original networks; they are $z_r = 11.02$ and $z_m = 2.06$ for the retweet and mention networks, respectively. We conclude that the community structure found in the retweet network is significantly more segregated than that found in the mention network.

4.1.7 Community Composition

Given the communities of the retweet network identified in § 4.1.6 and the findings of previous studies, it is natural to investigate whether the clusters in the retweet network correspond to groups of users of similar political alignment.

To accomplish this in a systematic, reproducible way we used a set of techniques from the social sciences known as *qualitative content analysis* [56, 55]. Similar to assigning class labels to training data in supervised machine learning, content analysis defines a set of practices that enable social scientists to define reproducible categories for qualitative features of text. Next we outline our annotation categories, and then explain the procedures used to establish the rigor of these category definitions.

Our coding goals were simple: for a given user we wanted to identify whether his tweets express a 'left' or 'right' political identity, or if his identity is 'undecidable.' The groups primarily associated with a 'left' political identity are democrats and progressives; those primarily associated with a 'right' political identity are republicans, conservatives, libertarians and the Tea Party. A user coded as 'undecidable' may be taking part in a political dialogue, but from the content of her tweets it is difficult to make a clear determination about political alignment. Irrelevant non-English and spam accounts constitute less than 3% of the total corpus and were excluded from this analysis. We experimented with more detailed categorization rubrics but the simple definitions described above yielded the highest inter-annotator agreement in early trials of the coding process.

Using this coding scheme one author first annotated 1,000 random users who appeared in both the retweet and mention networks. Annotations were determined solely on the basis of the tweets present in the six week sample. In line with the standards of the field, we had a non-author judge with a broad knowledge of politics annotate 200 random users from the set of 1,000 to establish the reproducibility of this annotation scheme. The judge was provided a brief overview of the study and introduced to the coding guidelines described above, but did not have any other interaction with the authors during the coding process.

The statistic typically used in the social sciences to measure the extent to which a coders' annotations agree with an objective judge is Cohen's Kappa, defined as

$$\kappa = \frac{P(\alpha) - P(\epsilon)}{1 - P(\epsilon)} \tag{4.2}$$

where $P(\alpha)$ is the observed rate of agreement between annotators, and $P(\epsilon)$ is the expected rate of random agreement given the relative frequency of each class label [56, 55]. For agreement between the 'left' and 'right' categories we report $\kappa = 0.80$ and $\kappa = 0.82$ respectively, both of which fall in the "nearly perfect agreement" range [62]. For the undecidable category we found "fair to moderate" agreement ($\kappa = 0.42$), indicating that there are users for whom a political identity might be discernible in the context of specific domain knowledge. To address this issue of context-sensitive ambiguity we had a second author also annotate the entire set of 1,000 users. This allowed us to assign a label to a user when either author was able to determine a political alignment, resolving ambiguity in 15.4% of users.

For completeness we also report binomial p-values for observed agreement, treating annotation pairs as observations from a series of Bernoulli trials. Similar to the Kappa statistic results, inter-annotator agreement for the 'left' and 'right' categories is very high ($p < 10^{-12}$). Agreement on the 'undecidable' category is again lower ($p = 0.18$).

Based on this analysis it is clear that a majority of politically active users on Twitter express a political identity in their tweets. Both annotators were unable to determine a political identity in only 8% of users. A more conservative approach to label assignment does not change this story much; if we assign a political identity only to users for whom both annotators agree, we report unambiguous political valences for more than 75% of users.Using these annotations we can infer the expected political makeup of the network communities identified in § 4.1.6. As shown in Table 4.4, the network of political retweets exhibits a highly partisan community structure with two homogenous clusters of users who tend to share the same political identity. Surprisingly, the mention network does not exhibit a clear partisan community structure. Instead we find that it is dominated by a politically heterogeneous cluster accounting for more than 97% of the users, suggests that politically active Twitter users may be exposed to views with which they do not agree in the form of cross-ideological mentions.

Increasing the number of target communities in the mention network does not reveal a more fine-grained ideological structure, but instead results in smaller yet politically heterogeneous clusters. Similarly, the retweet network communities are maximally homogenous in the case of two clusters.

4.1.8 Cross-Ideological Interactions

The strong segregation evident in the retweet network and the fact that the two clusters correspond to political ideologies suggest that, when engaging in political discourse, users often retweet just other users with whom they agree politically. The dominance of the mention network by a single heterogeneous cluster of users, however, suggests that individuals of different political alignments may interact with one another much more frequently using mentions. Let us test these conjectures, and propose an explanation based on selective hashtag use by politically motivated individuals.

To investigate cross-ideological mentions, we compare the observed number of links between *manually-annotated* users with the value we would expect in a graph where users connect to one another without any knowledge of political alignment. The intuition for the expected number of links is as follows: for a set of users with k directed edges among them, we preserve the source of each edge and assign the target vertex to a random user in the graph, simulating a scenario in which users connected irrespective of political ideology. For example, if there are a total of k_R links originating from right-leaning users, and the numbers of left-leaning and right-leaning users are U_L

and U_R respectively, then the expected number of edges going from right-leaning to left-leaning users is given by:

$$E[R \to L] = k_R \cdot \frac{U_L}{U_L + U_R}. \tag{4.3}$$

We compute the other expected numbers of edges ($R \to R$, $L \to R$, $L \to L$) in the same way.

In Table 4.5 we report the ratio between the observed and expected numbers of links between users of each political alignment. We see that for both means of communication, users are more likely to engage people with whom they agree. This effect, however, is far less pronounced in the mention network, where we observe significant amounts of cross-ideological interaction.

4.1.9 Content Injection

Any Twitter user can select arbitrary hashtags to annotate his or her tweets. We observe that users frequently produce tweets containing hashtags that target multiple politically opposed audiences, and we propose that this phenomenon may be responsible in part for the network structures described in this study.

As a thought experiment, consider an individual who prefers to read tweets produced by users from the political left. This user would frequently see the popular hashtag `#p2` ("Progressives 2.0") in the body of tweets produced by other left-leaning users, as shown in Table 4.6. However, if this user clicked on the `#p2` hashtag hyperlink in one of these tweets, or searched for it explicitly, she would be exposed to content from users on both sides of the political spectrum. In fact, because of the disproportionate number of tweets produced by left- and right-leaning users, nearly 30% of the tweets in the `#p2` search feed would originate from right-leaning users.

A natural question is why a user would annotate tweets with hashtags strongly associated with ideologically opposed users. One explanation might be that he seeks to expose those users to information that reinforces his political views. Consider the following tweets:

```
User A: Please follow @Username for an outstanding progressive voice!  #p2 #dems #prog
#democrats #tcot
User B: Couple Aborts Twin Boys For Being Wrong Gender..http://bit.ly/xyz #tcot #hhrs
#christian #tlot #teaparty #sgp #p2 #prolife
```

These tweets were selected from the first page of the realtime search results for the `#tcot` ("Top Conservatives on Twitter") and `#p2` hashtags, respectively, and messages in this style make up a substantial portion of the results.

This behavior does not go unnoticed by users, as underscored by the emergence of the left-leaning hashtag `#p21`. According to a crowdsourced hashtag definition site (www.tagdef.com), `#p21` is a hashtag for "Progressives sans RWNJs" and "Political progressives w/o all the RWNJ spam that `#p2` has," where RWNJ is an acronym for "Right Wing NutJob." This tag appears to have emerged in response to the efforts by right-leaning users to inject messages into the high-profile `#p2` content stream, and ostensibly serves as a place where progressives can once again be exposed only to content aligned with their views.

We propose that when a user is exposed to ideologically opposed content in this way, she will be unlikely to rebroadcast it, but may choose to respond directly to the originator in the form of a mention. Consequently, the network of retweets would exhibit ideologically segregated community structure, while the network of mentions would not.

4.1.10 Political Valence

To explore the content injection phenomenon in more detail let us introduce the notion of *political valence*, a measure that encodes the relative prominence of a tag among left- and right-leaning users. Let $N(t, L)$ and $N(t, R)$ be the numbers of occurrences of tag t in tweets produced by left- and right-leaning users, respectively. Then define the valence of t as

$$V(t) = 2\,\frac{N(t,R)/N(R)}{[N(t,L)/N(L)] + [N(t,R)/N(R)]} - 1 \tag{4.4}$$

where $N(R) = \sum_t N(t, R)$ is the total number of occurrences of all tags in tweets by right-leaning users and $N(L)$ is defined analogously for left-leaning users. The translation and scaling constants serve to bound the measure between -1 for a tag only used by the left, and $+1$ for a tag only used by the right. Table 4.7 illustrates the usefulness of this measure by listing hashtags sampled from valence quintiles ranging from the far left to the far right, where valence is computed only for hashtags produced by manually-annotated users.

If hashtag-based content injection is related to the comparatively high levels of cross-ideological communication observed in the mention network, we expect users who use hashtags in this way to receive proportionally more mentions from users with opposing political views. Using community identities in the retweet network as a proxy for political alignment, we plot in Figure 4.5 the average proportions of mentions users receive from and direct toward members of the other community versus the mean valence of all tags produced by those users. A key finding of this study, these results indicate that users contributing to a politically balanced combination of content streams on average receive and produce more inter-ideological communication than those who use mostly partisan hashtags. Moreover, Table 4.6 shows that the most popular hashtags do not have neutral valence, ruling out that neutral-valence users are simply using the most popular hashtags.

4.2 Political Knowledge Discovery

Armed with the insights into the structure of domestic political communication established in the previous section, we now focus on the problem of making inferences about individual's political orientations based on their behavioral attributes. Industrial applications for this line of research are clear, as political advertising expenditures are estimated to have reached four billion US dollars during the 2010 U.S. congressional midterm elections [29]. The 'Citizens United' Supreme Court ruling, which removed restrictions on corporate spending in political campaigns has only accelerated this trend. As a result, political campaigns are placing more emphasis on social media tools as a low-cost platform for connecting with voters and promoting engagement among users in their political base.

Of particular interest to political campaigns is how the scale of the Twitter platform creates the potential to monitor political opinions in real time. For example, imagine a campaign interested in tracking voter opinion relating to a specific piece of legislation. One could easily envision applying sentiment analysis tools to the set of tweets containing keyword relating to the bill. However, without the ability to distinguish between users with different political affiliations, aggregation over conflicting partisan signals would likely obscure the nuances most relevant to political strategy.

In this section we explore several different approaches to the problem of discriminating between users with left- and right-leaning political alignment using the thousand user political communication dataset described in Section 4.2.1. While it's clear that the network-based approach to partisanship prediction should perform well (95% accuracy), this technique has several limitations that warrant the development of alternative text-based approached. For one, the clusterization technique described in Section 4.1.6 relies on node membership in the largest connected component of the retweet network, thus excluding individuals in a multitude of disconnected subgraphs. Moreover, the network-based classifier does not generalize well to new individuals, as the network clusterization step must be re-run if we wish to make a determination about a novel actor under the current protocol. To address these concerns, we develop text-based machinery that achieves 91% overall accuracy when tasked with predicting whether a user's tweets express a 'left' or 'right' political alignment.

Additionally, using latent semantic analysis we identify hidden sources of structural variation in user-generated metadata that are strongly associated with individuals' political alignment.

We conclude with a proof of concept application based on these classifications, identifying the websites most frequently tweeted by left- and right-leaning users. We show that domain popularity among politically active Twitter users is not strongly correlated with overall traffic to a site, a finding that could allow campaigns to increase returns on advertising investments by targeting lower-traffic sites that are very popular among politically active social media users.

4.2.1 Training Data

For a training corpus we rely on the manually-annotated data describing 1,000 politically-active Twitter users developed in Section 4.2.1. As noted in that section, inter-annotator agreement is quite high for the 'left' and 'right' categories, but quite marginal for the 'ambiguous' category. As a consequence, there exist several users for whom one annotator had the domain knowledge required to infer a political alignment while the other did not. To address this issue we assigned a label to a user when either annotator detected information suggesting a political alignment in the content of a user's tweets. This mechanism was used to resolve ambiguity in 16% of users. Among the 956 relevant users in the sample there were 45 for whom the annotators explicitly disagreed about political alignment ('left' vs. 'right'). These individuals were included in the 'ambiguous' category.

After this resolution procedure, 373 users were labeled by the human annotators as expressing a 'left' political alignment, 506 users were labeled as 'right', and 77 were placed in the 'ambiguous' category, for a total of 956 users (Table 4.9). Ambiguous classifications are a typical result of scarce data at the individual level, but for completeness we report worst-case bounds on accuracy for the scenario in which all of these users are classified incorrectly.

4.2.2 Classification

For this classification task we examine several features from two broad categories: user-level features based on content and network-level features based on the relationships between users. Each feature set is represented in terms of a feature-user matrix M, where M_{ij} encodes the value for feature i with respect to user j.

For content-based classifications we use linear support vector machines (SVMs) to discriminate between users in the 'left' and 'right' classes. In the simple case of binary classification, an SVM works by embedding data in a high-dimensional space and attempting to find the hyperplane that best separates the two classes [51]. Support vector machines are widely used for document classification because they are well-suited to classification tasks based on sparse, high-dimensional data, such as those commonly associated with text corpora [50].

To quantify performance for different feature sets we report the confusion matrix for each classifier, as well as *accuracy* scores based on 10-fold cross-validation. For a confusion matrix containing true left (tl), true right (tr), false left (fl) and false right (fr), the accuracy of a classifier is defined by:

$$accuracy = \frac{tl + tr}{tl + tr + fl + fr} \tag{4.5}$$

where tl is the number of left-leaning users who are correctly classified, and so on.

4.2.2.1 Full-Text

To establish a performance baseline, we train a support vector machine on a feature-user matrix corresponding to the TFIDF-weighted terms (unigrams) contained in each user's tweets [88]. In addition to common stopwords we remove hashtags, mentions, and URLs from the set of terms produced by all users, a step we take to facilitate comparison with other feature sets. Additionally, we exclude terms that occur only once in the entire corpus because they carry no generalizable information and increase memory usage. After these preprocessing steps, the resulting corpus contains 13,080 features, each representing a single term.

The classification accuracy for this representation of the data is 79%, and its confusion matrix is shown in Table 4.10. The lower accuracy bound for this approach, assuming that all ambiguous users are incorrectly classified, is 72.6%.

4.2.2.2 Hashtags

Hashtags emerged organically within the Twitter user community as a way of annotating topics and threads of discussion. Since these tokens are intended to mark the content of discussion, we might expect that they contain substantial information about a user's political leaning.

In this experiment we populate the feature-user matrix with values corresponding to the relative frequency with which user j used a hashtag i. This value is equivalent to the TF measure from Equation 2.1, but described in terms of hashtags rather than unigrams. We note that weighting by IDF did not improve performance. Eliminating hashtags used by only one user we are left with 4,701 features. For this classification task we report an accuracy of 90.8%; see Table 4.10 for the confusion matrix. The lower bound on this approach, assuming that all ambiguous users were misclassified, is 83.5%.

As evidenced by its higher accuracy score, a classifier that uses hashtag metadata outperforms one trained on the unigram baseline data. Analogous findings are observed in biomedical document classification, where classifiers trained on abstracts outperform those trained on the articles' full text [4]. The reasoning underlying this improvement is that abstracts are necessarily brief and information rich. In the same way, Twitter users must condense substantial semantic content into hashtags, reducing noise and simplifying the classification task.

4.2.2.3 Latent Semantic Analysis of Hashtags

We apply latent semantic analysis to the hashtag-user matrix in an attempt to identify latent factors corresponding to political alignment. The coefficients of the linear combination of hashtags most strongly associated with the second left singular vector, shown in Table 4.11, suggest that one is present in the data. Hashtags with extreme coefficients for this dimension include `#dadt` for 'Dont Ask Don't Tell', `#p2` for Progressives 2.0, `#tcot` for Top Conservatives on Twitter, and `#ocra` for 'Organized Conservative Resistance Alliance.' The hashtag `#whyimvotingdemocrat` originally became a trending topic among left-leaning users, but was subsequently hijacked by right-leaning users to express sarcastic reasons they might vote for a Democratic candidate. A consequence of these coefficients is that users who use many left-leaning hashtags will have negative magnitude with respect to this dimension, and users who use many right-leaning hashtags will have positive magnitude in this dimension. Figure 4.6 shows clear separation between left- and right-leaning users in terms of the first and second right singular vectors.

A support vector machine trained on features describing users in terms of the first two right singular hashtag vectors does not improve accuracy compared to hashtag TF scores alone. Expanding the feature space to the first three LSA dimensions improves performance by an insignificant amount (about 0.1%), and the addition of subsequent features only degrades performance.

4.2.3 Network Analysis

The strong association between retweet cluster membership and political alignment developed in Section 4.2.1 suggests a simple classifier. This classifier would accept the cluster label of a user (either A or B), and assign that user to the 'left' if she was in cluster B, and the 'right' otherwise. The accuracy of this method is 95%; its confusion matrix is show in Table 4.10. Assuming that all 'ambiguous' users are incorrectly classified yields a lower bound of 87.3% on the accuracy. This finding emphasizes the importance of the structure relative to content with respect to the loci of information about actors engaged in digitally-mediated communication.

We also experimented with combining topological information with the content data introduced earlier, resulting in a feature set comprised of cluster assignments and 19 hashtag features selected by Hall's feature selection algorithm [40]. Using a support vector machine, this method performed no better than using the network cluster label alone (Table 4.10).

4.2.4 Applications

The machine learning apparatus outlined above, paired with sentiment analysis techniques, could supplement traditional phone-based opinion surveys by allowing political campaigns to monitor public opinion regarding specific candidates and issues among users in their voting base. Similarly, burst-detection mechanisms could be employed to detect a rise in prominence of a specific candidate or issue, allowing campaigns to shape marketing and messaging efforts in response to emerging topics and trends [54]. Critically, analyses of this nature depend on the ability to disambiguate users of different political identities, lest conflicting signals from users of opposing ideology cancel one another out at the aggregate level.

As an application of these data mining techniques consider buying decisions for web-based political advertising. Here we produce ranked lists of the domains most frequently tweeted by users of each political alignment, based on the predictions of the network classification method. Many Twitter users rely on URL shortening services to hash hyperlinks into a more compact format, and here we focus on links encoded using the popular *bit.ly* platform.

Ordered lists of the most popular domain names among left- and right-leaning users are presented in Table 4.12, and predictably tend to correspond to left- and right-leaning media. One exception is feedproxy.google.com, which is popular in both communities but is not politically aligned; it is the domain used for RSS feeds by Google Reader. Given these results, we emphasize that the domains most popular among left- and right-leaning Twitter users are not simply those with high traffic volume generally. Using the Alexa Web Information Service (aws.amazon.com/awis/) we obtained traffic statistics for each of the 200 most popular domains among users in each community. Alexa reports popularity in terms of pageviews per million impressions among users who have downloaded the Alexa toolbar plugin. Figure 4.7 suggests a weak correlation between the popularity among politically active Twitter users and global traffic volume. The Kendall's correlation coefficient between site popularity on Twitter and site popularity as measured by Alexa is $\tau = 0.12$ for sites popular among left-leaning users and $\tau = 0.14$ for right-leaning users.[2] These values confirm that

[2]Since the data are broadly distributed, the assumption of normality required for computing the Pearson correlation coefficient does not hold. Therefore, we turn to a non-parametric test of dependence and use Kendall's τ to measure *rank* correlation.

the correlation is weak for both groups. Consequently, marketing efforts targeted at users of a specific alignment (for example, calls for campaign contributions and issue-specific ads targeted at mobilizing a political base) may achieve a higher return on investment by purchasing advertising on sites that are popular among social media users but have lower traffic from the internet population at large.

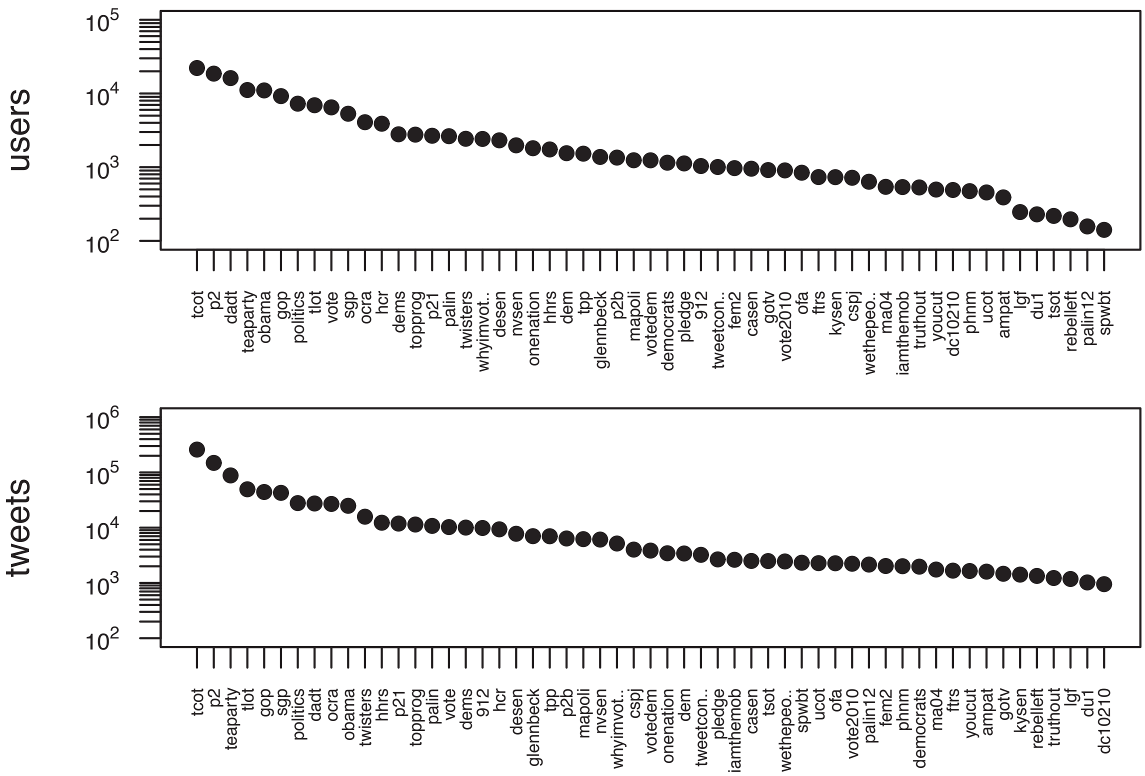

Figure 4.1: **Hashtag popularity decay in terms of total number of tweets and users associated with each tag.** On the horizontal axis tags have been ordered according to one of the two popularity measures: number of tweets (bottom) and users (top). The roughly exponential decay indicates that the inclusion of additional hashtags is unlikely to result in a substantial increase in the size of the corpus.

4.3 Conclusions

In this section we demonstrated that the two major mechanisms for public political interaction on Twitter — mentions and retweets — induce distinct network topologies. The retweet network is highly polarized, while the mention network is not. To explain these observations we highlight the role of hashtags in exposing users to content they would not likely choose in advance. Specifically, users who apply hashtags with neutral or mixed valence are more likely to engage in communication

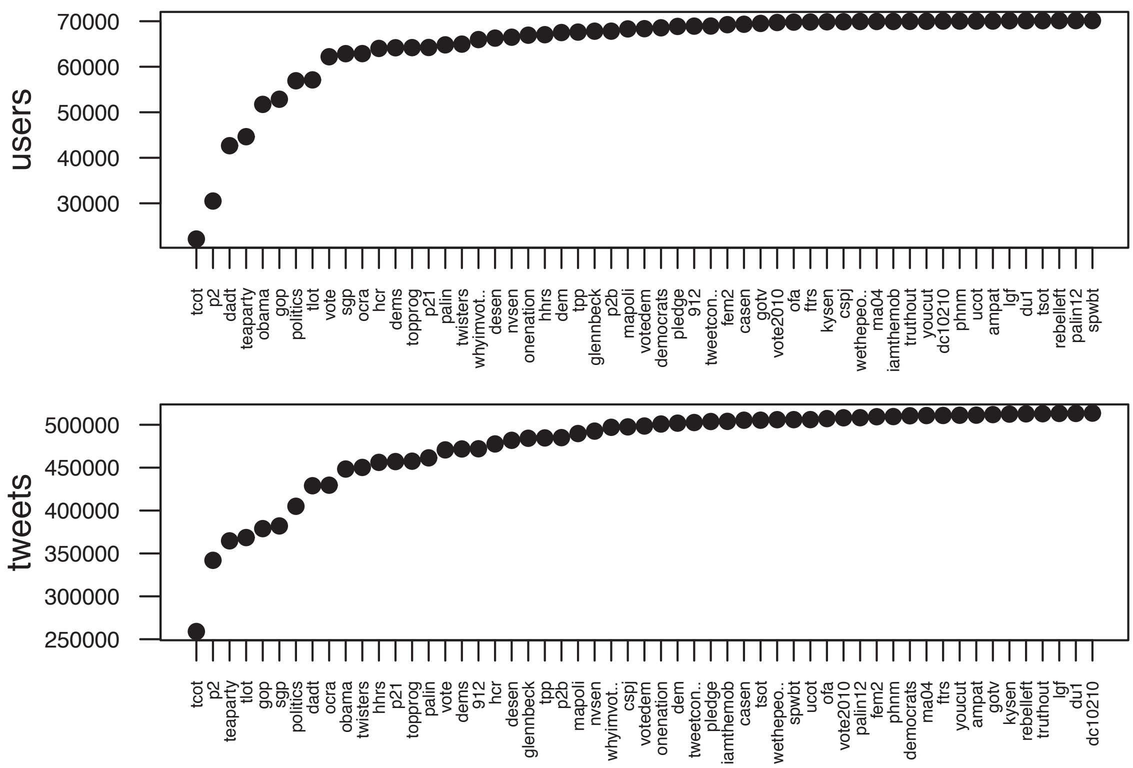

Figure 4.2: **Size of the set of unique users and tweets resulting from the inclusion of additional hashtags.** Axes are ordered according to the total number of tweets (top) and users (bottom) associated with each tag.

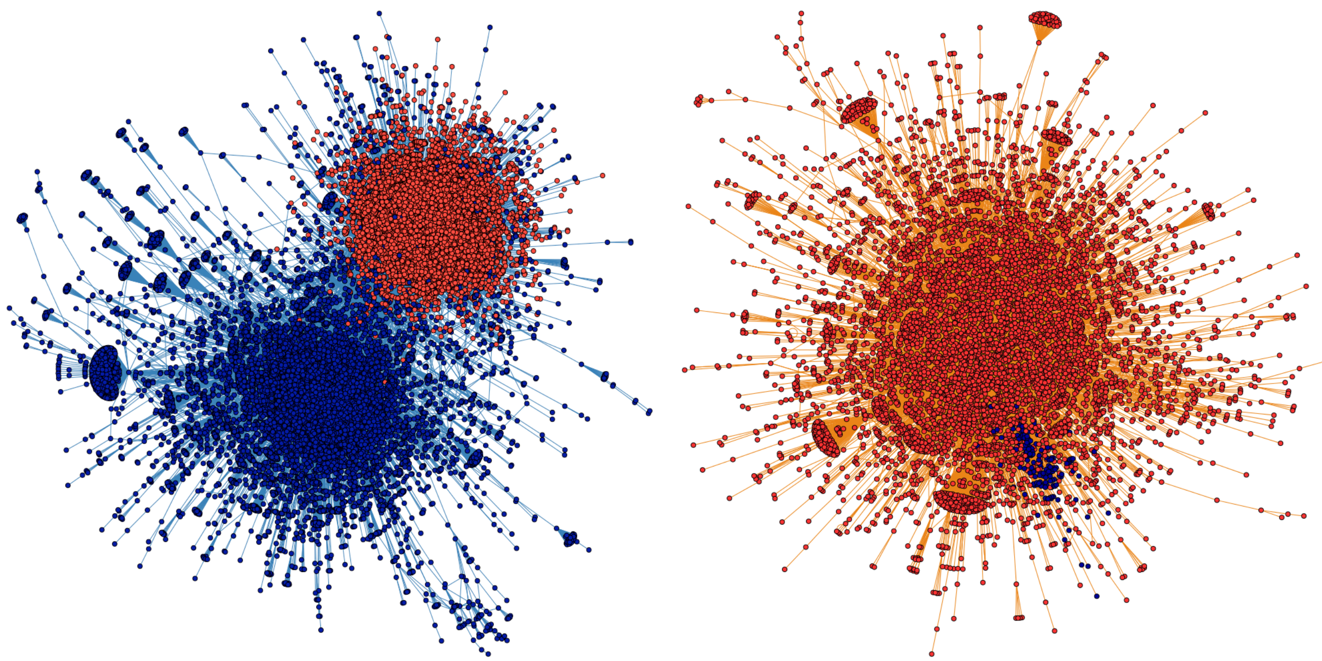

Figure 4.3: **The political retweet (left) and mention (right) networks, laid out using a force-directed algorithm.** Node colors reflect cluster assignments (see § 4.1.6). Community structure is evident in the retweet network, but less so in the mention network. We show in § 4.2.1 that in the retweet network, the red cluster A is made of 93% right-leaning users, while the blue cluster B is made of 80% left-leaning users.

with opposing communities. In addition, we demonstrate that politically-active Twitters users generate text- and network-based information that can be used to effectively predict the political alignment of large numbers of individuals.

Although our findings on polarization could be interpreted as encouraging evidence of cross-ideological political discourse, we emphasize that these interactions are almost certainly not a panacea for the problem of political polarization. While we know for certain that ideologically-opposed users interact with one another, either through mentions or content injection, they very rarely share information from across the divide with other members of their community. It is possible that these users are unswayed by opposing arguments and facts, or that the social pressures that lead to group polarization are too strong for most users to overcome [93]. Whatever the case, political segregation, as manifested in the topology of the retweet network, persists in spite of substantial cross-ideological interaction.

Qualitatively speaking, our experience with this body of data suggests that the content of political discourse on Twitter remains highly partisan. Many messages contain sentiments more extreme than you would expect to encounter in face-to-face interactions, and the content is frequently disparaging of the identities and views associated with users across the partisan divide. If Yardi and boyd are

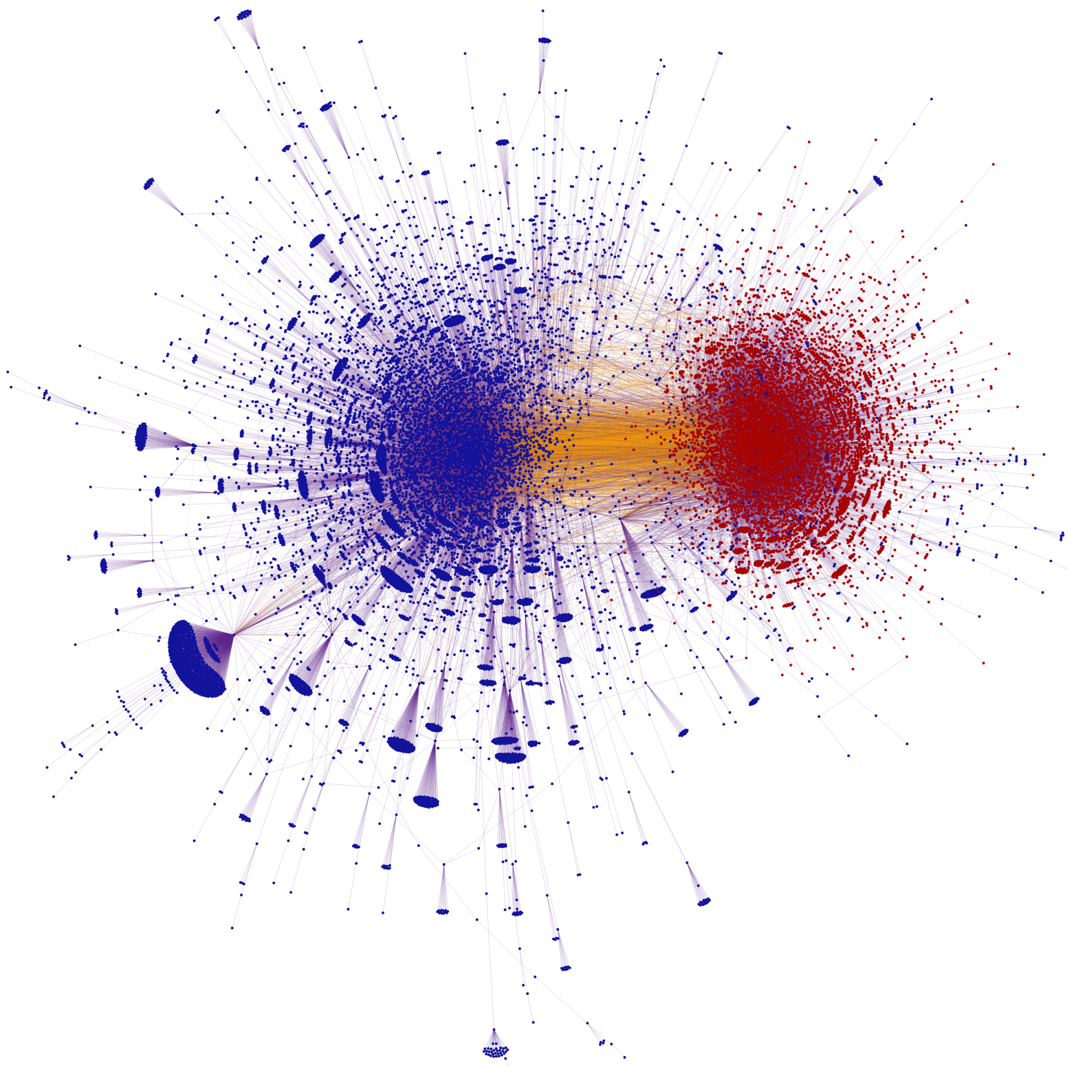

Figure 4.4: **The political retweet (left) and mention (right) networks, laid out using a force-directed algorithm.** Node colors reflect cluster assignments (see § 4.1.6). Community structure is evident in the retweet network, but less so in the mention network. We show in § 4.2.1 that in the retweet network, the red cluster A is made of 93% right-leaning users, while the blue cluster B is made of 80% left-leaning users.

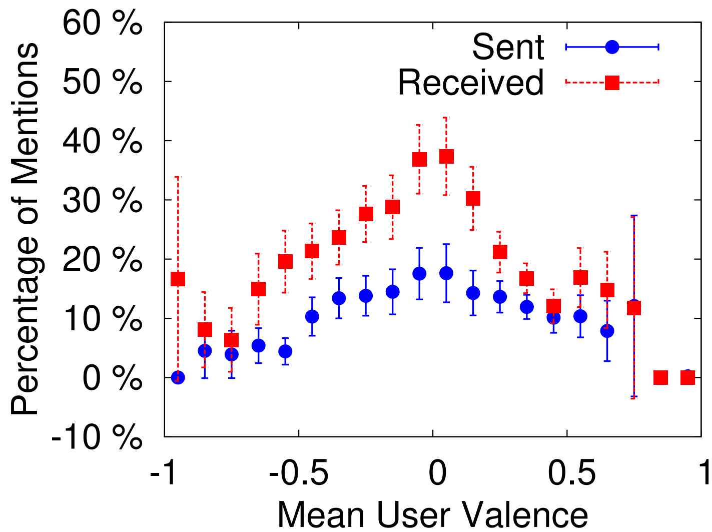

Figure 4.5: **Proportion of mentions a user sends and receives to and from ideologically-opposed users relative to her valence.** Points represent binned averages. Error bars denote 95% confidence intervals.

correct, and our experience suggests this may be the case, these interactions might actually serve to exacerbate the problem of polarization by reinforcing preexisting political biases. Further study of the content of inter-ideological communication, including sentiment analysis, as well as studies of network topology that include the follower network, could help to illuminate this issue.

With respect to the automated prediction of individuals partisan leanings, we detailed several approaches based on both content and network analysis. Techniques based on the statistical analysis of political communication networks provide the highest accuracy, thanks to the strong signal present in the partisan structure of the retweet network. However, we have shown that information-rich hashtag features are almost as effective at capturing political alignment, and have the benefit of generalizing without the need to recluster the network to accommodate new users.

Finally, as a proof of concept we illustrate the utility of this prediction capability by identifying the websites most popular among Twitter users from the political left and right, respectively. This approach reveals novel information about the popularity of different media outlets that can be leveraged to improve web-based advertising purchasing decisions.

We see much potential in the techniques described herein, which together represent a critical component in the real-time analysis of public opinion at the scale of tens of thousands of individual political actors. Looking forward, interesting open questions remain with respect to the generalizability of these approaches to international political discourse, multi-party systems, and the overall representativeness of communication on social media platforms.

Table 4.1: **Hashtags related to #p2, #tcot, or both.** Tweets containing any of these were included in our sample.

Just #p2	#casen #dadt #dc10210 #democrats #du1 #fem2 #gotv #kysen #lgf #ofa #onenation #p2b #pledge #rebelleft #truthout #vote #vote2010 #whyimvotingdemocrat #youcut
Both	#cspj #dem #dems #desen #gop #hcr #nvsen #obama #ocra #p2 #p21 #phnm #politics #sgp #tcot #teaparty #tlot #topprog #tpp #twisters #votedem
Just #tcot	#912 #ampat #ftrs #glennbeck #hhrs #iamthemob #ma04 #mapoli #palin #palin12 #spwbt #tsot #tweetcongress #ucot #wethepeople

Table 4.2: **Hashtags excluded from the analysis due to ambiguous or overly broad meaning.**

Excl. from #p2	#economy #gay #glbt #us #wc #lgbt
Excl. from both	#israel #rs
Excl. from #tcot	#news #qsn #politicalhumor

Table 4.3: **Minimum, maximum, and average ARI similarities between 4,950 pairs of cluster assignments computed by label propagation on the mention and retweet networks.**

Network	Min	Max	Mean
Mention	0.80	1.0	0.89
Retweet	0.94	0.98	0.96

Table 4.4: **Partisan composition and size of network clusters as determined by manual inspection of 1,000 random user profiles.**

Network	Clust.	Left	Right	Undec.	Nodes
Retweet	A	1.19%	93.4%	5.36%	7,115
	B	80.1%	8.71%	11.1%	11,355
Mention	A	39.5%	52.2%	8.18%	7,021
	B	9.52%	85.7%	4.76%	154

Table 4.5: **Ratios between observed and expected number of links between users of different political alignments in the mention and retweet networks.**

	Mention		Retweet	
	$\to$ Left	$\to$ Right	$\to$ Left	$\to$ Right
Left	1.23	0.68	1.70	0.05
Right	0.77	1.31	0.03	2.32

Table 4.6: **The ten most popular hashtags produced by left- and right-leaning users in the manually annotated set of users, including frequency of use in the two retweet communities and ideological valence.**

Rank	Hashtag	Left	Right	Valence
1	`#tcot`	2,949	13,574	0.384
2	`#p2`	6,269	3,153	-0.605
3	`#teaparty`	1,261	5,368	0.350
4	`#tlot`	725	2,156	0.184
5	`#gop`	736	1,951	0.128
6	`#sgp`	226	2,563	0.694
7	`#ocra`	434	1,649	0.323
8	`#dems`	953	194	-0.818
9	`#twisters`	41	990	0.843
10	`#palin`	200	838	0.343
	Total	26,341	53,880	

Table 4.7: **Hashtags in tweets by users across the political spectrum, grouped by valence quintiles.**

Far Left	Moderate Left	Center	Moderate Right	Far Right
`#healthcare`	`#aarp #women`	`#democrats #social`	`#rangel #waste`	`#912project`
`#judaism #hollywood`	`#citizensunited`	`#seniors #dnc`	`#saveamerica`	`#twisters #gop2112`
`#2010elections`	`#democratic`	`#budget #political`	`#american #gold`	`#israel #foxnews`
`#capitalism #recession`	`#banksters #energy`	`#goproud`	`#repeal #mexico`	`#mediabias`
`#security #dreamact`	`#sarahpalin`	`#christian #media`	`#terrorism`	`#constitution`
`#publicoption`	`#progressives`	`#nobel`	`#gopleader #palin12`	`#patriots #rednov`
`#topprogs`	`#stopbeck #iraq`			`#abortion`

Table 4.8: **Contingency table of inter-annotator agreement on manual classifications.**

	Left	Ambiguous	Right
Left	303	51	23
Ambiguous	19	32	24
Right	22	59	423

Table 4.9: **Final class assignments based on resolution procedures described in text.**

Left	Ambiguous	Right
373	77	506

Table 4.10: **Summary of confusion matrices and accuracy scores for various classification features, with the sections in which they are discussed.**

Features	Conf. matrix	Accuracy	Section
Full-Text	$\begin{bmatrix} 266 & 107 \\ 75 & 431 \end{bmatrix}$	79.2%	§ 4.2.2.1
Hashtags	$\begin{bmatrix} 331 & 42 \\ 41 & 465 \end{bmatrix}$	90.8%	§ 4.2.2.2
Clusters	$\begin{bmatrix} 367 & 6 \\ 38 & 468 \end{bmatrix}$	94.9%	§ 4.2.3
Clusters + Tags	$\begin{bmatrix} 366 & 7 \\ 38 & 468 \end{bmatrix}$	94.9%	§ 4.2.3

Table 4.11: **Most extreme hashtag coefficients for second left singular vector. This linear combination of hashtags appears to capture variance associated with political alignment.**

Hashtag	Coeff.	Hashtag	Coeff.
#tcot	0.380	#p2	-0.914
#sgp	0.030	#dadt	-0.071
#ocra	0.020	#p21	-0.042
#hhrs	0.013	#votedem	-0.039
#twisters	0.012	#lgbt	-0.038
#tlot	0.011	#p2b	-0.032
#whyimvotingdemocrat	0.009	#topprog	-0.027
#rs	0.005	#onenation	-0.025
#ftrs	0.004	#dems	-0.023
#ma04	0.004	#gop	-0.021
#tpp	0.003	#hcr	-0.017

Table 4.12: **Websites most frequently tweeted by left- and right-leaning users, ranked by popularity.**

Popular Left	Popular Right
feedproxy.google.com	feedproxy.google.com
mediamatters.org	hotair.com
politicalwind.com	gop2112.com
youtube.com	youtube.com
dailykos.com	redstate.com
truthy-out.org	firstthings.com
msnbc.msn.com	americanthinker.com
thinkprogress.org	google.com
harryreid.com	survivalstation.org
realclearpolitics.com	newsbusters.org
www.google.com	biggovernment.com
twitwall.com	realclearpolitics.com
thedailybeast.com	conservatives4palin.com
feeds.dailykos.com	newsmax.com
crooksandliars.com	nationalreview.com

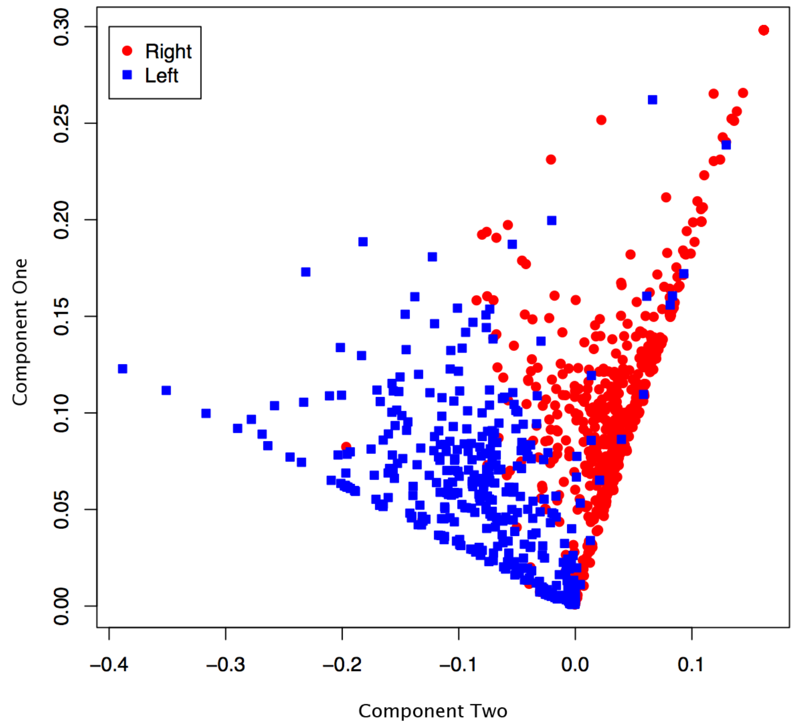

Figure 4.6: **Users plotted in the latent semantic space of the first and second right singular vectors.** Colors correspond to class labels.

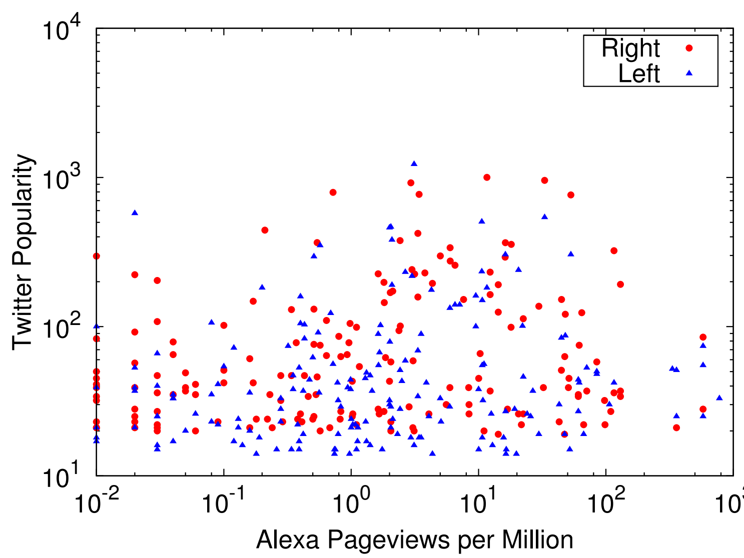

Figure 4.7: **Scatter plot of popularity of the 200 most frequently tweeted domains for members of left- and right-leaning network clusters versus global traffic among users of the Alexa toolbar.**

CHAPTER V

Partisan Asymmetries

"The political technology of the Industrial age is no longer appropriate technology for the new civilization taking form around us. Our politics are obsolete."

Alvin Toffler

Motivated by the connection between the widely reported advantage in on-line mobilization and the result of the 2008 presidential election [20, 29, 43], here we leverage partisanship predictions from Section 4.2 to understand how two communities of users, both afforded the same advantages of low cost networked information production, exhibit distinct outcomes with respect to their use of the Twitter platform.

Survey data from the Pew Research Center showed that, along the seven dimensions used to measure online political activity, Obama voters were substantially more likely to use the Internet as an outlet for political activity [77]. In particular, Obama voters were more likely than McCain voters to create and share political content, and to engage politically on an online social network [77]. Moreover, a 2009 Edelman report found that in addition to a thirteen million member e-mail list, the Obama campaign enjoyed twice as much web traffic, had four times as many YouTube viewers and five times more Facebook friends compared to the McCain campaign [65]. While the direct effect of any one media strategy on the success of a campaign is difficult to assess and quantify, the data show that Obama campaign had a clear advantage in terms of online voter engagement. We work toward this goal by examining partisan differences in the behavior, communication patterns and social interactions of more than $18,000$ politically-active users of Twitter, a social networking platform that allows individuals to create and share brief 140-character messages.

Having established the large-scale structure of these communication networks, in this chapter we employ a variety of methods to provide a more detailed picture of domestic political communication on Twitter. We characterize a wide range of differences in the behavior, communication, geography and social connectivity of thousands of politically left- and right-leaning users. Specifically, we demonstrate that right-leaning Twitter users exhibit greater levels of political activity, tighter social bonds, and a communication network topology that facilitates the rapid and broad dissemination of political information, a finding that stands in stark contrast to the online political dynamics of the 2008 campaign.

With respect to individual-level behaviors, we find that right-leaning Twitter users produce more than 50% more total political content and devote a greater proportion of their time to political discourse. Right-leaning users are also more likely to use hyperlinks to share and refer to external content, and are almost twice as likely than left-leaning users to self-identify their political alignment in their profile biographies. At the individual level, these behavioral factors paint a picture of a right-leaning constituency comprised of highly-active, politically-engaged social media users, a trend we see reflected in the communication and social networks in which these individuals participate.

Regarding connectivity patterns among users in these two communities we report findings related to three different networks, described by the set of explicitly declared follower/followee relationships, mentions, and retweets. Casting the declared follower network as the social substrate over which political information is most likely to spread, we find that right-leaning users exhibit a greater propensity for mutually-affirmed social ties, and that right-leaning users tend to form connections with a greater number of individuals in total compared to those on the left. With respect to the way in which information actually propagates over this substrate in the form of retweets, right-leaning users enjoy a network structure that is more likely to facilitate the rapid and broad dissemination of political information. Additionally, right-leaning users exhibit a higher probability to rebroadcast content from and to be rebroadcast by a large number of users, and are more likely to be members of high-order retweet network k-cores and k-cliques, structural features that are associated with the efficient spreading of information and adoption of political behavior and opinions. Pointing definitively to a vocal, socially engaged, densely interconnected constituency of right-leaning users, these topological and behavioral features provide a significantly more nuanced perspective on political communication on this important social media platform.

5.1 Data

The analysis described in this chapter relies on data collected from the Twitter 'gardenhose' streaming API between September 1^{st} and January 7^{th}, 2011 — the eighteen week period surrounding the November 4^{th} United States congressional midterm elections, inclusive of the study period described in Chapter IV. From this eighteen week period we collected data on $6,747$ right-leaning users and $10,741$ left-leaning users, responsible for producing a total of $1,390,528$ and $2,420,370$ tweets, respectively. It's useful to note that we evaluate all gardenhose tweets associated with each user, rather than just those containing political hashtags, in order to facilitate comparisons between the two groups in terms of relative proportions of attention allocated to political communication.

5.2 Methodology

In order to examine differences in the behavior and connectivity of left- and right-leaning Twitter users we rely on the political hashtags (Section 4.1.2) and partisan cluster membership labels (Section 4.2.1) established in Chapter IV.

In the previous chapter we used the set of political tweets from the six weeks preceding the 2010 midterm election to build a network representing political retweet interactions among Twitter users. In this network an edge runs from a node representing user A to a node representing user B if B retweets content originally broadcast by A, indicating that information has propagated from A to B. This network consists of $23,766$ non-isolate nodes among a total of $45,365$, with $18,470$ nodes in its largest connected component and 102 nodes in the next-largest component. We describe the construction of an analogous network of political mentions in Section 5.4.3.

Based on manual content analysis from Section 4.2.1, we determined that the retweet network communities are highly politically homogeneous, consisting of 80.1% left- and 93.4% right-leaning users, respectively. In this chapter we use network community membership as a proxy for the political identities of all $18,470$ users in the largest connected component of the retweet network, and hereafter focus on the behavior of these users. Based on the relative proportions of right- and left-leaning users identified during the qualitative content analysis stage, this mechanism results in correct predictions for 87.3% of users in the largest connected component of the retweet network [27].

In the following sections we leverage these data to explore, in detail, how users from the political left and right utilize this important social media platform for political activity in different ways.

5.3 Behavior: Individual-level Political Activity

Before examining structural differences in the social and communication networks of left- and right-leaning Twitter users, we first focus on political activity at the individual level. In this section we compare users in the left- and right-leaning communities in terms of their relative rates of content production, the amount of attention they allot to political communication, their respective rates of political self-identification, and their propensity for sharing information resources in the form of hyperlinks.

Right-leaning users are substantially more active and politically engaged with this social media platform. Specifically, our analysis shows that left-leaning users produce less total political content, allocate proportionally less time to creating political content, are less likely to reveal their political ideology in their profile biography, and are less likely to share resources in the form of hyperlinks. All of these findings stand in stark contrast to survey data and media reportage of the 2008 online political dynamics, and provide evidence in support of the notion that right-leaning voters are becoming more politically engaged online.

5.3.1 Political Communication

From the perspective of leveraging social media for political organization, the baseline level of activity among a constituency is one of the most important characteristics of a population. Figure 5.1 shows that while left- and right-leaning users produce approximately the same number of tweets per user, right-leaning individuals actually produce 54% more total political content despite comprising fewer users altogether. This trend is the result of divergent priorities among left- and right-leaning users, as right-leaning users devote a substantially larger portion of their activity on Twitter to political communication. In fact, right-leaning users were almost twice as likely to create political content, with 22% of all tweets produced by right-leaning users containing one or more of the political hashtags under study, compared to only 12% for left-leaning users.

5.3.2 Partisan Self-Identification

In addition to devoting a larger proportion of tweets to political content, right-leaning users are much more likely to use their 140-character profile 'biography' to explicitly self-identify their political alignment. A survey of the biographies of 400 random users from the set of individuals selected for qualitative content analysis (Section 5.2) reveals that 38.7% of right-leaning users included reference to their political alignment in this valuable space, as compared with only 24.6% of users in the left-leaning community. Taken together, this analysis demonstrates that right-leaning users are much more likely to use Twitter as an outlet for political communication, and are substantially more inclined to view the Twitter platform as an explicitly political space.

5.3.3 Resource Sharing

One of the key functions of the Twitter platform is to serve as a medium for sharing information in the form of hyperlinks to external content. Given the constraints of the 140-character format, hyperlinking activity is especially important to the dissemination of detailed political information among members of a constituency.

With respect to this aspect of online political engagement, too, we see that right-leaning users are more active then those individuals in the left-leaning community. Among all tweets produced by users in the right-leaning community, 43.4% contained a hyperlink, compared with 36.5% of all tweets from left-leaning users. This trend is even more pronounced if we consider only resource sharing within the set of political tweets, with left-leaning users including a hyperlink in 50.8% of political tweets, as compared to right-leaning users, who include hyperlinks 62.5% of the time. From these observations we conclude that right-leaning users are more inclined to treat Twitter as a platform for aggregating and sharing links to web-based resources, an activity crucial to the efficient spread of political information on the Twitter platform.

5.4 Connectivity: Global-level Political Activity

Next, we turn our attention to structural differences in social interaction and communication networks of left- and right-leaning users.

5.4.1 Follower Network

We begin with an analysis of the network defined by the follower/followee relationships shared among members of these two groups (Figure 5.2). Encoding the fact that a user subscribes to the content produced by another, the follower network is best understood as describing the social substrate over which information is likely to flow between political actors on Twitter. Specifically, though not all connections in the follower network encode equally meaningful social relationships, content is broadcast equally along all edges in this network.

We examine the differences in the follower subgraphs induced by considering only connections between users of the same political affiliation. For the purposes of this analysis, a directed edge is drawn from user A to user B if A is a follower of B. Basic statistics about these two subgraphs, including average degree, undirected clustering coefficient, and proportion of reciprocal links are presented in Table 5.1. We see that along all dimensions, users in the right-leaning community are much more tightly interconnected, with a substantially higher average clustering coefficient and greater average degree. Additionally, we observe a higher proportion of reciprocal links between right-leaning users, indicating the presence of stronger, mutually-affirmed interest among individuals in this community. All of these factors indicate that right-leaning users are more tightly interconnected, resulting in a basic structural advantage with respect to the challenge of efficiently spreading political information on the Twitter platform.

Using the Kolmogorov-Smirnov two-sample test to measure the degree of similarity between the in- and out-degree distribution for left- and right-leaning users we find a significant difference between the in-degree distributions of left- and right-leaning users, but only a marginal difference between the corresponding out-degree distributions (Figure 5.3). We interpret this to mean that a right-leaning user is more likely to have a large audience of followers who may potentially rebroadcast his or her call to action or piece of political information. For example, left-leaning users are roughly twice as likely as right leaning users to have in-degree one, while users that are associated with the right are almost four times more likely to have in-degree $1,000$ than users associated with the left. Additionally, users in the left-leaning community are more likely to be only peripherally connected into the network, as evidenced by the distribution of the k-core shell indices of users in each community (Figure 5.4). For a given network, the k-core is the maximal subgraph whose nodes (as members of the subgraph) have at least degree k, or, in other words, have at least k neighbors in the k-core itself. The shell index, c, of a node refers to the coreness (k) of the highest-order k-core of which the node is a member [8].

These observations lead us to conclude that there are substantial structural differences in the fundamental patterns of social connectivity among politically left- and right-leaning Twitter users, a finding supported by the seminal work of Adamic & Glance [1] on the connectivity patterns of high-profile partisan bloggers. Specifically, the right-leaning community is much more densely interconnected, with more users tightly integrated into the right-leaning social network. In contrast, the network of follower/followee relations among left-leaning users exhibits a much more decentralized, loosely-interconnected structure, with far fewer mutually-affirmed social connections.

5.4.2 Retweet Network

Next we consider the structure of the network of political retweets in order to understand how information actually spreads on the social substrate characterized in Section 5.4.1. While each link in the follower network represents a potential pathway along which information may flow, edges in the retweet network correspond to real information propagation events. Specifically, when user A rebroadcasts a tweet produced by user B, she explicitly signifies receipt of the content in question, and thus we draw an edge from user B to user A indicating the direction of information flow. Consequently, the structure of the retweet network reveals much about how information actually spreads within these two communities. Visualized previously in Figure 4.4, basic statistics describing the networks induced by retweets containing at least one political hashtag between users of the same partisan affiliation are show in Table 5.2.

In practice, the tightly-interconnected structure of the retweet network confers communication advantages to the right-leaning community of users. Examining the in- and out-degree distributions for these two communities we find that though the power-law exponents are similar, the difference between them is statistically significant at the 95% level (Figure 5.5). The faster decay in the degree distribution of the left-leaning community implies that right-leaning users are rebroadcast by and rebroadcast content from a larger number of individuals than users on the left. That right-leaning users pay attention to more information sources compared to left-leaning individuals is indicative of a higher degree of engagement with the Twitter platform itself. Similarly, an individual wishing to rapidly reach a wide audience has a natural advantage given the structure of the right-leaning retweet network.

With respect to the number of users in high-order k-cores, too, we see that the right-leaning community enjoys structural advantages, with a greater proportion of highly active users connected to other highly active users (Figure 5.6). This difference could lead to consequences in the spread of information through these networks. Work by Kitsak *et al.* indicates that it is individuals with high shell index, rather than those who are most central or well connected, who are the most effective spreaders of information under a simple SIR-based information diffusion model [53]. Users on the right therefore, are more likely than those on the left to be wired into the political communication network in such a way that they are able to facilitate the broad and rapid dissemination of political information.

We also find that a substantially higher proportion of right-leaning user participate in fully-connected subgraphs of size k, known as k-cliques. This result is especially important in the context of the complex contagion hypothesis, which posits that repeated exposures to controversial behaviors are essential to the adoption of these behaviors. Work by Romero, Meeder and Kleinberg focused specifically on online social networks indicates that this effect is particularly pronounced for political discourse on Twitter [86]. With fewer users in high-order k-cores, individuals in the left-leaning community will be less likely to encounter multiple users discussing the same partisan talking points or calls to action, exactly the kind of contentious content whose propagation is most likely to benefit from repeated exposure.

5.4.3 Mention Network

Mentions are most strongly associated with direct, conversational engagement when the target username appears at the beginning of a tweet, as opposed to appearing in the body text. Among the mentions in our sample, the overwhelming majority (94.5%) take this form, providing strong evidence that connectivity among and between users in these two groups represents actual political discourse rather than simply third-person references. In Table 5.3 we report descriptive statistics on the topology of the left- and right-leaning mention networks, where an edge from A to B is drawn between two users of the same political affiliation if A mentions B in a tweet containing at least one political hashtag. Though the two networks exhibit very similar degree distributions, one important distinction is the fact that a greater proportion of mention relationships in the right-leaning community are reciprocal. Compared to the number of reciprocal mentions observed in degree-preserving reshufflings of the left- and right-leaning mention networks, the right-leaning community exhibits 7.5 times as many reciprocal mention interactions than is expected by chance alone, compared to a 5.6 times as many reciprocal links in the left-leaning community. Reciprocal interactions suggest the presence of more meaningful social connections, manifest in conversational dialogue, rather than, for example, unidirectional commentary on the content of another user's tweets. Here too, we find that users on the political right are more engaged with one another on Twitter, indicating that they are likely to benefit from a richer dialogue and hence more opportunities for frame-making and consensus building with respect to political topics.

5.5 Political Geography

In addition to characterizing differences in behavior and connectivity, we can also examine the geographic distribution of individuals in these two communities. Here we present a cartogram in which the color of each state has been scaled to correspond to the degree to which, in that state, the observed number of tweets originating from the left-leaning community exceeds what we should expect by chance alone.

Because fewer than one percent of Twitter users provide precise geolocation data, we instead rely on the self-declared 'location' field of each user's profile to enable geographic analysis of data at the scale of this study. As a free-text field, users are able to enter in arbitrary data, and non-location responses such as 'the moon' do appear in the results. Complicating this analysis further, some users do not report any location data, though we do not report a partisan bias in terms of non-entries. Despite these caveats, a large number of users do report actual locations, and using the Yahoo Maps

Web Service API (http://developer.yahoo.com/maps/rest/V1/geocode.html), we are able to make a best-guess estimate about the state with which a user most strongly identifies.

Thus, for each state in which we observe N total tweets, and the relative proportion of tweets originating from left-leaning users (P_l), we can treat the arrival of partisan tweets as a Bernoulli process, and compute the number of tweets we should expect to see from left-leaning users as NP_l. Likewise, we can compute the extent to which the observed number of tweets associated with left-leaning users (T_l) is above or below the expected number, measured in terms of standard deviations, as

$$\frac{T_l - NP_l}{\sqrt{NP_l \cdot (1 - P_l)}}. \tag{5.1}$$

Figure 5.7 uses color to encode these deviations for each state, with states in which the volume of activity far exceeds what should be expected by chance shown in deep red, and those in which the observed volume is far below what should be expected by chance shown in light yellow.

Initial inspection of this figure reveals that the geographic distribution of individuals from the left-leaning network community corresponds strongly to the traditional political geography of the United States. We see that left-leaning individuals feature prominently on the coasts and North East, and tend to be underrepresented in the midwest and plains states.

Looking more closely, however, we find that there are some places in which the partisan makeup of tweets is quite different from what might be hypothesized intuitively. For example, Utah, a traditionally conservative state which at the time of this writing had two Republican senators, exhibits a dramatically higher volume of left-leaning content than should be expected by chance alone. One possible explanation for this observation could be that individuals in some states with an ideologically homogeneous population turn to social media as an outlet for political expression. While this is but one possible explanation among many, and a more rigorous analysis is required to support any definitive claim, this example illustrates the ways in which novel hypotheses can derive from data-driven analyses of political and sociological phenomena.

5.6 Conclusion

These analyses indicate a shifting landscape with respect to partisan asymmetries in online political engagement. We find that, in contrast to what might be expected given the online political dynamics of the 2008 campaign, right-leaning Twitter users exhibit greater levels of political activity, tighter social bonds, and a communication network topology that facilitates the rapid and broad dissemination of political information.

In terms of individual behavior, politically right-leaning Twitter users not only produce more political content and devote a greater proportion of their time to political discourse, but are also more likely to view the Twitter platform as an explicitly political space and identify their political leanings in their profiles. With respect to social interactions, the right-leaning community exhibits a higher proportion of reciprocal social and mention relationships, are more likely to rebroadcast content from a large number of sources, and are more likely to be members of high-order retweet network k-cores and k-cliques. Such structural features are directly associated with the efficient spreading of information and adoption of political behavior. Taken together, these features are indicative of a highly-active, densely-interconnected constituency of right-leaning users using this important social media platform to further their political views.

Finally, these results also demonstrate that, despite being afforded the same opportunities to produce and consume large volumes of content for very low cost, these two communities of users exhibit very different usage patterns of the Twitter platform.

Table 5.1: **Follower network statistics for the subgraphs induced by the set of edges among users of the same political affiliation.** Reciprocity is defined as $\frac{D_R}{D}$, where D_R is the number of dyads with an edge in each direction and D is the total number of dyads with at least one edge. Follower data was only available for a subset of the study population, owing to private or deleted accounts.

Community	Nodes	Edges	Avg. Degree	Clust. Coeff.	Reciprocity
Left	9,941	803,329	80.80	0.134	42.8%
Right	6,426	1,503,417	233.95	0.221	64.8%

Table 5.2: **Retweet network statistics for the subgraphs induced by the set of edges among users of the same political affiliation.**

Community	Nodes	Edges	Avg. Degree	Clust. Coeff.	Reciprocity
Left	11,353	32,772	2.88	0.032	13.5%
Right	7,115	39,713	5.58	0.045	12.1%

Table 5.3: **Mention network statistics for the subgraphs induced by the set of edges among users of the same political affiliation.**

Community	Nodes	Edges	Avg. Degree	Clust. Coeff.	Reciprocity
Left	11,353	50,273	4.42	0.053	20.8%
Right	7,115	64,993	9.13	0.078	24.5%

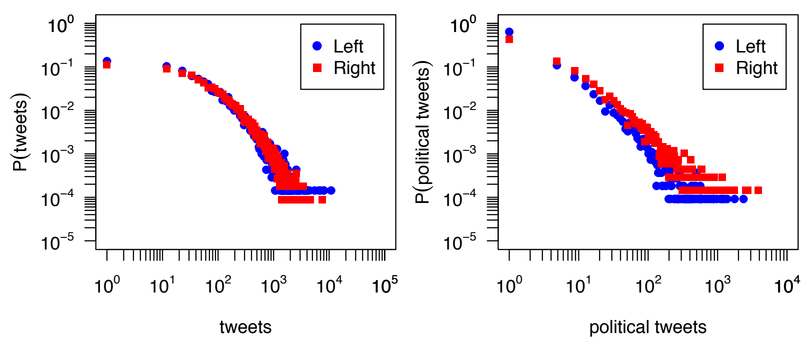

Figure 5.1: **Total number of tweets produced by right- and left-leaning users (left) compared to the total number of political tweets produced by users in each group.** While both groups produce a comparable amount of content in general, right-leaning users produce a much larger number of political tweets despite comprising fewer users in total. We observe that users' behavior tends to be broadly distributed, with many individuals creating relatively few tweets, while a few individuals produce substantially larger volumes of content. Note, however, that this sample includes only users who produced at least one political hashtag, rather than a random sample among all Twitter users, a feature likely responsible for the low number of users who produce few total tweets.

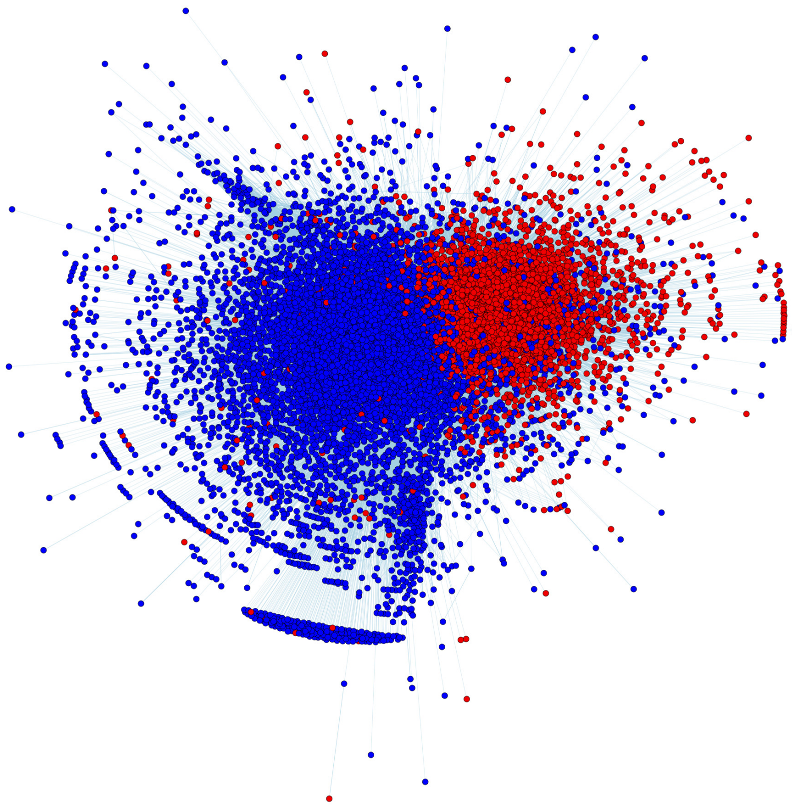

Figure 5.2: **Force-directed layout of the follow relationships among politically-active Twitter users.** Nodes are colored according to political identity, Connections to users who did not engage political communication on Twitter are not included.

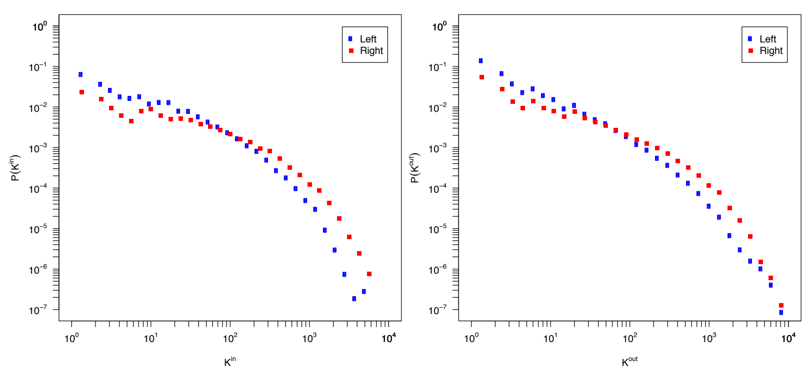

Figure 5.3: **Log binned in- and out-degree distributions of the internal follower network at left, and right, respectively.** As a result of considering only follower relationships among politically-active users we observe strong cutoffs in both distributions that make curve-fitting unreliable. However, comparing the two distributions it's clear that the right-leaning community has a much greater proportion of users with many followers (Kolmogorov-Smirnov $p < 10^{-3}$), despite being comprised of fewer users in total. Understood as an information diffusion substrate, the proliferation of high-profile hubs gives a natural advantage to the right-leaning community.

Figure 5.4: **Linearly binned core distribution of the internal follower network.** The difference between these two distributions is highly significant (Kolmogorov-Smirnov $p < 10^{-3}$).

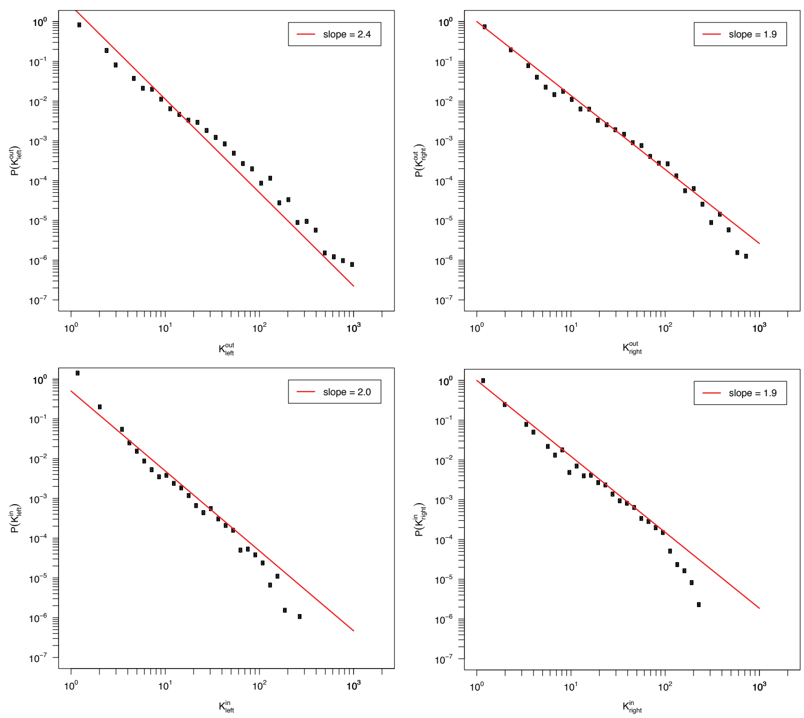

Figure 5.5: **Log binned in- and out-degree distributions for the left- and right-leaning retweet network communities.** Slopes and standard errors were inferred using the maximum likelihood estimation method described by Clauset, Shalizi & Newman [25]. The rapid decay of the left-leaning degree distribution indicates that right-leaning users are retweeted by and retweet content from a larger number of users than those on the left.

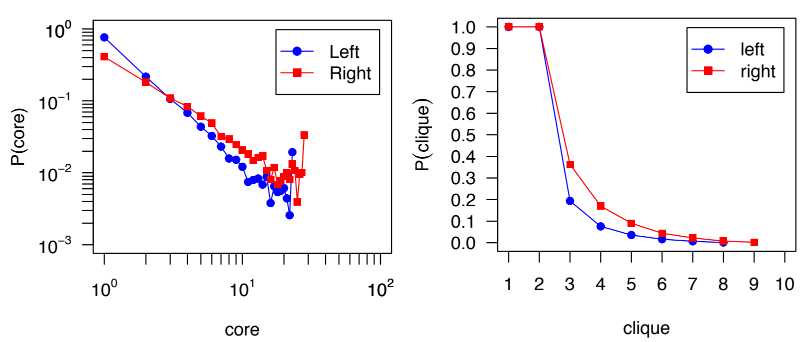

Figure 5.6: **Proportion of users with a given k-core shell index (left) and membership in a k-clique (right) for the retweet network.**

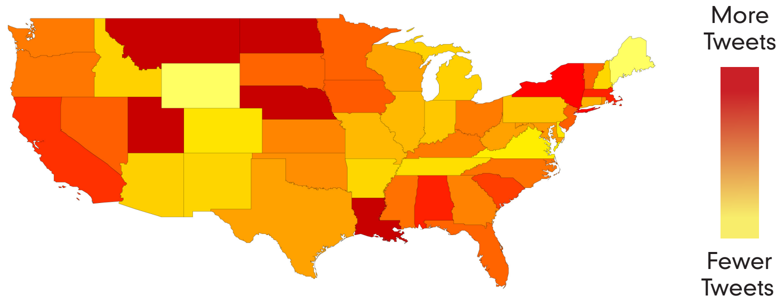

Figure 5.7: **Deviation in volume of left-leaning political communication compared to expected baseline.** Each state is filled with a color corresponding to the extent to which the observed number of tweets is above or below what should be expected in the case where each state has traffic volume proportional to that observed across all Twitter traffic.

CHAPTER VI

Social Movement Communication

"The revolution is not an apple that falls when it is ripe. You have to make it fall."

Che Guevara

One of the most prominent American political movements of the past thirty years, Occupy Wall Street ('Occupy') is remarkable in the extent to which social media played a central role in its development and organization [23, 19]. In this study, we examine how the needs and constraints of social movements are reflected in the geospatial characteristics and temporal dynamics of the information sharing practices of Twitter users engaged in communication about the Occupy movement. With respect to the first issue, we focus on the geographic distribution of these users and the ways in which the relationships among them diverge from those of users contributing to the two most popular streams for stable political discourse in the United States, 'Top Conservatives on Twitter' and 'Progressives 2.0.' With respect to the movement's temporal evolution, we investigate changes in Occupy participant engagement, interests, and social connectivity over a fifteen month period starting three months prior to the movements first protest action.

The organizing forces underlying successful social movements have been studied extensively by sociologists and political scientists. From this body of work common themes have emerged, include the problems of resource mobilization and collective framing, which together constitute two of the core issues any social movement must address in order to effect social or political change. Resource mobilization refers to the process through which a social movement must marshal the financial, material, and human resources required to sustain its activities [66]. Collective framing is a process whereby the constituents of a social movement, through formal or informal processes, come to establish the narratives, language, and imagery that capture the essential features of the movement's purpose and struggle [9]. Effective framing helps to foster a sense of community and engagement, and can be a powerful response to countervailing social pressures from establishment organizations. [34].

Here we study Occupy Wall Street, a social movement focused on issues relating to the uneven distribution of wealth, social inequality, corporate greed, and the regulation of major financial institutions. Since the first protest on September 17^{th}, 2011, a major feature of the movement has been the long-term physical occupation of high-visibility encampments, often found in parks, banks, libraries and foreclosed homes. As a result, the Occupy movement requires substantial supporting infrastructure, including housing and sanitation facilities, as well as access to communication technologies. In spite of this, Occupy has sustained a lasting presence in American cities including New York City, Oakland, Washington, D.C., and Boston, which also represent key loci of decision making and protest activity [23, 19]. Under the Occupy model, proposals are brought to a vote before a general assembly, a form of direct democracy in which any participant is free to comment or vote on any proposal under consideration. The most prominent among these organizational structures is the New York City General Assembly, which has been responsible for producing policy and key narrative frames such as the popular protest slogan, "We are the 99%," which references the disproportionate concentration of wealth among the top 1% of the world's population.

Social media have played a prominent role in facilitating communication and coordination throughout the development of the Occupy Wall Street movement. For example, the first call to action in the Canadian anticapitalist magazine 'AdBusters' used the Twitter 'hashtag' `#occupywallstreet` as one of just ten words featured in a full-page ad. Ever since, the Twitter platform has been used extensively by movement participants [19], with `#ows` being one of the hundred most popular hashtags on Twitter for the year 2011.

The use of information communications technologies by social movement organizations has been attributed to lowered barriers to participation, increased ease with which small-scale acts can be aggregated, the rapid propagation of logistical information and narrative frames, and a heightened sense of community and collective identity [72, 99, 103, 13, 12]. With respect to Twitter in particular, one study of tweets related to the Spanish social and economic 'Indignados' protests found that Twitter played a role in the recruitment of new individuals and the dissemination of information related to mass mobilization [35]. Another found that users tweeting about the 2011 Egyptian revolution were broadly distributed both inside and outside of of Egypt [24].

In the first portion of this chapter, we seek to understand the relationship between the geospatial dimensions of social movement communication networks and the organizational pressures facing such movements. To accomplish our analysis of the movement's geospatial properties, we use a state-of-the-art location inference technique to model relationships among users as a weighted directed network of communication flows between states, in which the weight of each edge corresponds to the volume of traffic between pairs of locations. Using this framework we investigate three distinct relationships: attention allocation and proximity to on-the-ground events, resource mobilization and localized information sharing, and the role of collective framing in long-distance communication.

With respect to the issue of attention allocation, we find that compared to stable domestic political communication the Occupy Wall Street movement exhibits very high levels of geographic concentration, with users in New York, California, and Washington D.C. producing more than half of all retweeted content. Consequently, the Occupy communication network exhibits a distinctive hub and spoke topology, with just a few high profile locations serving as the principle sources of widely-rebroadcast information.

Additionally, we report that with the exception of the largest hubs, the appeal of content relating to Occupy Wall Street has a disproportionately local audience. Specifically, we find that information is three and a half times more likely to be produced and consumed by users in the same state when compared to the network of stable domestic political communication. With extended, high profile encampments and large-scale protest action playing central roles in the Occupy movement, we propose that this structural feature reflects the importance of mobilizing human resources at the local level.

Finally, we report on evidence indicating that the content of communication at the national level is distinct from the content of communication among users in the same state. Comparing intrastate versus interstate communication, we find that the terms most overrepresented in interstate communication relate to the movement's core framing language and the news media, while the terms most overrepresented in local communication reference physical places, protest action, and specific times. These results support the hypothesis that local-level communication activity is driven by the challenge of resource mobilization, while long-distance communication is more strongly associated with collective framing processes.

In the second piece of this analysis, we study the total amount of Occupy-related traffic on the platform from September 2011 through September 2012. With respect to this measure of activity, we find that Occupy traffic has diminished by orders of magnitude relative to peak activity volumes in late 2011. This effect is evident even in concerted attempts to revive the movement's flagging levels of engagement, with activity returning to baseline within a week of May 1st, 2012 reoccupation efforts.

Finding little evidence of sustained activity, we turn our attention to Occupy participants themselves, in hopes of understanding how these users were changed as a result of engaging with the movement online. Using a random sample of 25,000 Occupy users, we study changes in behavior at the individual level with respect to attention allocation and social connectivity. From this analysis we are left to conclude that, on Twitter, Occupy evoked interest from a highly-interconnected community of users with pre-existing interest in domestic politics and foreign social movements. Though we find statistically significant changes in political interests and social connectivity over the study period, the magnitude of these changes pales in comparison to the amount of attention these individuals allocated to the Occupy Wall Street cause.

6.1 Materials and Methods

6.1.1 Data

The analysis described in this article relies on several related dataset collected from the Twitter 'gardenhose' streaming API. To identify Occupy-related content, we deem relevant any tweet containing a hashtag matching either `#ows` or `#occupy*`, where `*` represents a wildcard character. This set includes high-profile tags such as `#occupy` as well as location-specific tokens such as `#occupyoakland` and `#occupyseattle`. To provide a baseline against which to compare our observations, we also extracted content originating from the two most popular communication channels associated with stable domestic political communication, `#tcot` (Top Conservatives on Twitter) and `#p2` (Progressives 2.0). While this approach does not allow us to study content that does not contain an Occupy-specific hashtag, we argue that it is appropriate for two reasons. As outlined above, hashtags allow a user to reach an audience beyond his or her immediate followers, and it is this kind of expressly public engagement in which we are primarily interested. Moreover, while topic modeling techniques may allow for the analysis of untagged tweets, their use would introduce noise that could cloud the interpretation of any analytical results.

To characterize the geospatial structure of the Occupy communication network we rely on a corpus of tweets collected between July 3^{rd}, 2011 and March 12^{th}, 2012. As this analysis is concerned primarily with information spreading processes we consider only retweet events from this corpus, resulting in 676,369 retweets among 257,657 users associated with Occupy Wall Street, and 259,703 retweets among 68,049 users associated with stable domestic political communication. Additionally, we produce a corpus of all sampled tweets containing at least one of these hashtags from the year-long period between September 1st, 2011 to August 31st, 2012. Referred to hereafter as the *Occupy corpus*, this dataset contains approximately 1.82 million tweets produced by 447,241 distinct accounts.

In addition to changes in activity explicitly related to the Occupy movement, we are also interested in changes to the behavior of individual users over time. To this end, we identified a random sample of 25,000 random users who produced at least one tweet in the Occupy corpus. We then produced a second corpus containing any tweet, regardless of content, generated by each account in this sample during the 15-month period spanning June 1st, 2011 through August 31st, 2012. Including tweets from the three-month period preceding the start of the Occupy Wall Street movement allows us to study the behavior of these users before, during, and after the movement's primary period of activity. Referred to hereafter as the *random sample*, this dataset contains approximately 7.74 million tweets produced by 25,000 unique users.

To facilitate analysis relating to the attention allocation habits of these individuals, we rely on three non-overlapping sets of hashtags: those related to Occupy Wall Street (defined above), a second set relating to foreign social movements, and a third relating to domestic political communication. As we are interested exclusively in the attention allocation habits of Occupy users, we identified the set of hashtags relating to domestic political communication and foreign social movements by manually inspecting the 300 hashtags most frequently used by individuals in the random sample. Table 6.2 lists the hashtags associated with each topic. While not exhaustive due to a long-tail use distribution, the 300 most popular hashtags account for 70.8% of all tagging activity, with the 300th most popular tag constituting just 0.027% of all tags. We therefore believe that the inclusion of additional tags in our topic lists is not likely to affect the results of this study.

6.1.2 Time Series

Many of the analyses in this article rely on time series describing changes to measured quantities over the course of the study period. Each time series is produced by computing a single statistic on disjoint sets of tweets partitioned into adjacent, temporally non-overlapping bins of k hours. For all of these analyses we use one of three temporal resolutions to reveal different characteristics of the signal under study: 12 hours, 24 hours, or one week.

At various times over the course of the study period, our system experienced service outages that affected our ability to collect data from the Twitter API. Amounting to 15 days in total, these periods are: September 29 to October 4, 2011; October 11–12, 2011; December 28–30, 2011; February 11–13, 2012; February 16–17, 2012; and May 28–31, 2012. Owing to the fact that the measures we employ reflect relative composition of the stream rather than its absolute volume, these outages do not unduly influence the statistical character of our results.

6.1.3 Geocoding

To facilitate a geospatial analysis of communication activity associated with these content streams we require a high quality method to infer individual users' locations. To accomplish this, we rely on self-reported location strings and the services of a commercial geocoding API. This technique, popularized in work by Onnela et al. [76], has been shown to produce high-resolution, high-quality geolocation data in the presence of geographically meaningful input.

A caveat to this technique, however, is that it relies on raw text generated by a broad swath of the Twitter population, and so we find geographically meaningless location descriptors included in the dataset. To address this issue we rely on an extensive hand-curated blacklist of popular non-geographical responses such as 'everywhere' and 'the dance floor'. To produce this list we sorted all location strings by popularity and reviewed the thousand most popular strings manually, blacklisting those that did not correspond to geographically meaningful entities. Drawn from a long tailed distribution, 53% of all tweets in the data set are associated with a location among the 1,000 most popular responses, with 27% of all tweets containing one of the top hundred location strings. From this set of one thousand we blacklisted 161 non-location strings, corresponding to 6% of the tweets associated with the 1,000 most popular responses.

To improve recall in the presence of novel input, we used a modified version of the Ratcliff-Obershelp algorithm [84] to detect fuzzy matches between free text location strings and the blacklist of popular non-location responses. As a result, because 'the dance floor' is in the set of blacklist responses, strings taking a slightly modified form, such as 'on the dance floor,' will also be classified as invalid input. The hand-coded blacklist combined with the Ratcliff-Obershelp fuzzy matching technique resulted in 9% of the free-text location strings being classified as non-location input.

From among the remaining responses we submitted location strings to the Bing.com geocoding API, which returns a best-guess estimate for the corresponding physical coordinates. This output is hierarchically formatted to describe the finest level of geographic resolution available. For example, if a user reports 'Logan Square, Chicago' as his or her location, the Bing API will return information about the likely zip code, city, state and country associated with that location. However, if the user reports only 'USA,' the information provided by the API describes only a country-level guess as to the user's location. Owing to decreased coverage at the city-level and the proportionately few users associated with each individual city, we utilize the state-level location estimates for the geospatial components of this analysis.

In total, 68.4% of Occupy Wall Street users reported location strings, and from these we were able to obtain geolocation estimates for 55.7% of these accounts. Among this set of users, 60% of the resulting geolocation estimates included state-level metadata. Response rates were somewhat diminished for users associated with the stream of domestic political communication, with 36% of individuals reporting free-text location strings. Using the procedure described above, we were able to obtain geolocation estimates for 29.3% of all users in the domestic political communication stream, 82.4% of which contained state-level metadata.

6.1.4 Geographic Profile

One of our goals is to establish a coarse-grained geographic profile for communication activity associated with the Occupy Wall Street movement. Formally, for each stream we define an activity distribution across states as, $A_i = \frac{T_i}{|T|}$, where T_i is the total number of retweets originating from state i and $|T|$ is the total number of retweets originating from all states. As outlined above, we focus on retweets as they correspond to attention allocation rather than total content production volume.

In addition to the distribution of activity across individual states we examine the information sharing relationships among users in different locations. To accomplish this, we rely on a network representation to characterize the flow of information on Twitter. Taking users as nodes, we define a weighted directed network in which an edge with weight w is drawn from node U_1 to U_2 in the event that user U_2 retweets user U_1 w times. The intuition underlying this approach is that each retweet provides evidence suggesting that information produced by user A was evaluated and acted upon by user B.

Combining the user-level geocode metadata described in Section 6.1.3 with the network representation defined here we can induce another network describing the volume of communication between users in each state. In this network, nodes represent states, and weighted directed edges are drawn among them. The weight of the edge from S_1 to S_2 is defined as the sum of the weights among all edges originating from users in state S_1 and terminating in state S_2. We note, however, that this induced network must have geolocation labels for each node in a dyad. In the Occupy Wall Street stream we identify 143,437 tweets for which both the source and target have state-level geolocation data and 78,467 likewise restricted tweets in the stream of stable domestic political communication.

6.1.5 Textual Content

Finally, we wish to investigate whether the content of tweets with different geospatial properties serve distinct communication functions. To accomplish this, we segregate Occupy Wall Street tweets into two classes: *interstate* tweets connect pairs of users in different states, and *intrastate* tweets connect users in the same state. We compute the probability of observing a token, t, in a tweet from a given class, x, as $P(t|x)$. Comparing these probabilities yields a ratio, $\frac{P(t|intrastate)}{P(t|interstate)}$, a value which is large when a token is more common in intrastate traffic than interstate traffic and small under the opposite conditions.

6.2 Results

6.2.1 Geographic Concentration

Figure 6.1, in which states are ordered according to the proportion of stream activity, shows that content in the Occupy stream is substantially more geographically concentrated in a few key states compared to domestic political communication. For example, New York accounts for 30% of the total retweet activity in the Occupy stream, while the most popular source for stable domestic political communication, Washington D.C., accounts for only 10.7% of the stream's total volume. As these plots make clear, the primary locations for on-the-ground Occupy activity are those places responsible for the majority of widely rebroadcast Occupy content, with California, New York and Washington D.C. acting as the source of 53.8% of total retweets. Figure 6.2 maps the states where the proportion of activity associated with the Occupy stream deviates the most from that associated with the stream of domestic political communication.

We also study the ratio of content production to content consumption by locale. Figure 6.3 shows this ratio, defined as the total number of retweets originating from users in that state divided by the total number of tweets retweeted by users in that state. This value serves to highlight the extent to which users in a given location are functioning as content producers or content consumers. Inspecting this plot, we find that in the Occupy stream users from just five states produced more content than they consumed. This stands in contrast to the stream of stable domestic political communication, in which fourteen states exhibit a ratio greater than one.

To highlight the effect of this geospatial concentration on communication flows between states it is instructive to visualize the structure of these networks. However, owing to the geographic aggregation process outlined in Section 6.1.4 both networks are highly dense, with edges spanning most pairs of states. To address this issue we utilize a technique known as multiscale backbone extraction [89], which is useful for identifying statistically significant edges in weighted networks, regardless of the absolute value associated with the weight of that edge. This technique selects for edges with weights significantly above the expectation given by an analytically defined probability distribution that models a random allocation of each node's strength among its adjacent edges. Parameterized by a confidence level factor, α, this technique allows for the selection of statistically significant edges across all weight scales, a feature that is especially valuable when working with networks with heterogeneous weight distributions such as those associated with communication or human mobility.

Applying this technique to both networks reveals a communication backbone for the Occupy network that exhibits the highly concentrated hub and spoke structure described above. Figure 6.4 shows that the Occupy Wall Street network is characterized by minimal state-to-state connectivity, with the majority of statistically significant traffic flowing to and from New York, California and Washington D.C. This is in contrast to the communication backbone for the network of domestic political communication, in which we observe extensive interactions among many pairs of states.

6.2.2 Localization

In Figure 6.5 we present interstate connectivity for each communication network as a matrix in which the weight of an edge is mapped to a grayscale hue ranging from white for weak relationships to black for the strongest relationships. Inspecting these plots, one of the most striking ways in which the topology of the Occupy Wall Street communication network departs from that of the domestic political communication network is the high degree of localization. This is evidenced by the presence of a strong diagonal in the Occupy Wall Street connectivity matrix, as well as the significant off-diagonal mass in the domestic political communication matrix. We find that 40% of Occupy retweets originate and terminate with users in the same state. In contrast, 11% of retweets from the domestic political stream exhibit this type of locality, an increase of more than 350%.

6.2.3 Textual Analysis

To study the relationship between geography, resource mobilization, and collective framing, we focus on the content of tweets flowing within and between states. Restricting our analysis to tokens that account for at least 0.1% of both the intrastate and interstate tweet text, Table 6.1 presents the ten tokens most overrepresented in both intrastate communication as well as interstate communication.

6.2.4 Activity Volume

We next focus to the total number of tweets in the Occupy corpus over the course of a year-long period between September 1st, 2011 and August 31st, 2012. As evidenced by Figure 6.6, in general, Occupy traffic closely mirrors on the ground activity, and is characterized by an increase to peak levels during the month-long period following the movement's initial protests, with significantly diminished activity levels over the following eleven months. In terms of relative change, average levels of Occupy traffic in the second half of the period from September 17th, 2011 to August 31st, 2012 decreased 80.8% relative to the first half of the same period.

6.2.5 Attention Allocation

In light of these diminishing activity levels we wish to gain insights into the character of the individuals from which Occupy drew its support. We begin by studying how Occupy user interests changed in time, examining the frequency with which 25,000 random individuals produced content relating to one of three topics: Occupy Wall Street, foreign social movements, and domestic politics. Based on the random sample described in Section 5.1, the results of this analysis describe activity from June 1, 2011 to August 31, 2012, a period including the three months prior to the initial protest action.

As we are interested in the behavior of individuals who were active on Twitter at a given time, we identify the set of users U_i from whom we observe at least one tweet at time step i, regardless of its content. Within this set we isolate, at each timestep, the set of users U_{it} from whom we observe, in any of their tweets, at least one hashtag relating to topic t. The *engaged user ratio* $|U_{it}|/|U_i|$ describes the extent to which individuals chose to engage in communication relating to each of the three topic areas.

Among the set of users engaged with a topic, we next examine the extent to which that topic tends to dominate their content production activity. To accomplish this, let us consider, for each user $u \in U_{it}$, the collection H_{iu} of hashtags contained in his or her tweets at time step i. From this we compute the proportion of each user's tagging activity that is associated with a given topic, $|H_{iut}|/|H_{iu}|$, where H_{iut} is the set of tags from topic t produced by u at time step i. Averaging this value across all engaged users provides a lens on the behavior of these individuals as a whole, and is reported as the *engaged user attention ratio*. Figure 6.7 presents this value alongside the engaged user ratio to show how the amount of attention allocated to the three topics changed over time.

As expected, a large fraction of users produced Occupy related content during the period of peak activity, with more than 40% of sampled users allocating on average 64% of their attention to the topic during the third week following the initial protests. However, this intense focus on the subject is not sustained over the course of the following year, with the engaged user ratio decaying to less than 5% in the last three months of the study period. Moreover, comparing the engaged user attention ratios from the first half of the period following the initial Occupy protests ($\mu = .439$) to those from the second half ($\mu = .318$), we find that individuals who continue to produce Occupy content do so with significantly lower frequency ($p < 10^{-3}$), suggesting diminished enthusiasm even among the most persistent individuals.

With respect to foreign social movements and domestic political communication, we observe that users who would go on to engage with the Occupy movement online tended to exhibit interest in these topics before the initial protest activity in September, 2011. Comparing the engaged user ratios in the first 12 weeks of the study period with those observed during the last 12 weeks of the study period, we find a significant ($p < 10^{-3}$) but small increase in domestic political communication activity (from $\mu = 0.066$ to $\mu = 0.077$). With respect to interest in foreign social movements, we observe a significant ($p < 0.05$) but small decrease in engagement for the same periods (from $\mu = 0.074$ to $\mu = 0.057$). These differences suggest that the changes in individual behaviors in response to the Occupy Wall Street movement were limited.

Finally, let us examine the extent to which Occupy users tended to interact with one another over the course of the study period. To this end we focus on the proportion of retweets and mentions produced by active users in the random sample that involved another user who produced at least one Occupy-related tweet during the year following the movement's inception. This proportion is computed with respect to all of a user's retweets and mentions, regardless of content, rather than just those related to Occupy Wall Street. Figure 6.8 shows that we observe a statistically significant increase ($p < 10^{-6}$) in in-group retweet and mention activity during the peak period of Occupy activity, followed by a gradual decay to values approaching pre-Occupy levels. Comparing the period before the inception of the movement to the one at the end of the study period, we report a small but significant increase in both in-group retweets ($p < 10^{-6}$) and mentions ($p < 10^{-3}$), with the mean connectivity increasing 5.1% for retweets and 3.2% for mentions. Although these changes are statistically significant, they can hardly be interpreted as evidence that this community's long-term social connectivity has been dramatically altered in response to participation in the Occupy Wall Street movement.

6.3 Conclusion

The Twitter platform represents a powerful organizing tool for social movement actors, enabling participants to communicate with one another and share information in a public, networked environment. Utilizing hashtags, members of the Occupy Wall Street movement created high-visibility communication channels that made it easy for users to engage with a community of activists well beyond their immediate network of followers. These streams rapidly grew to be among the most popular on the Twitter platform, with `#occupywallstreet` ranking among the 100 most popular hashtags in 2011. These high levels of activity and visibility were to the benefit of the movement, generating news coverage and potentially facilitating recruitment of new participants by way of social reinforcement processes. However, such favorable conditions did not last, and over the course of the following year the Occupy streams decayed to activity levels orders of magnitude lower than those they once exhibited.

Among our main findings we determine that, relative to stable domestic political communication, the Occupy network has a highly localized geospatial structure, with a disproportionately large amount of traffic being produced and consumed by users in the same state. We propose that this phenomenon may be related to the issue of resource mobilization, that is, the process whereby any social movement must marshall resources such as money, infrastructure and human capital to further the goals of the movement. In the case of Occupy Wall Street, such resources are often quite tangible, and include not only tents and food, but also the participants required to facilitate large-scale protest action and extended encampments in cities across the country. In this light, it is easy to understand why such a disproportionately large fraction of attention is allocated to communication at the local level.

With respect to the finding that the majority of widely rebroadcast content is produced by users in a small number of high profile locations, we observe that these states represent sites of major encampment and decision making activity. Despite the fact that all users can contribute equally to the Occupy stream, it appears that proximity to events on the ground plays a major role in determining which content receives the most attention. This is in contrast to the stream of domestic political communication, in which content from users across the United States is allocated a significant share of attention. Where the stream of domestic political communication looks more like a conversation taking place at the national level, the structure of the Occupy stream is more akin to a broadcast, with just a few locations playing the role of net content producers.

Finally, we propose that interstate communication plays a significant role in the propagation of narrative imagery associated with collective framing processes, and that intrastate communication is driven more predominantly by the pressures of resource mobilization. Looking to the lists of tokens most overrepresented in each type of traffic (Table 6.1), we find that those more common in interstate communication include references to core framing language and the news media. This finding suggests that when users engage in communication across state boundaries they allocate proportionately higher levels of attention to speech associated with collective framing processes. In contrast, the tokens more common in intrastate traffic relate to protest action and specific times and places. From this we conclude that the content of intrastate tweets deals much more frequently with rallying the movement's participants, a core function of resource mobilization.

Regarding the movement's temporal dynamics, while interest and activity relating to the Occupy movement has substantially diminished, one could envision that increased levels of engagement with the political process online might constitute a positive outcome for the movement's participants. Along these lines, however, Occupy users remain barely changed, exhibiting a slight increase in attention paid to domestic politics and a slight decrease in attention paid to foreign social movements. Relative to the dramatic behavioral changes these users exhibited in the early stages of the movement, and the magnitude of Occupy-related communication in general, these changes constitute a somewhat underwhelming long-term effect.

Similarly, a supporter of the movement might take as a promising outcome increased levels of interaction among Occupy users. Such a scenario could indicate that these individuals formed a more tight-knit community over the course of the year, creating social and communication bonds that may help to facilitate the efficient spread of information, potentially even reinforcing individual propensity for offline activity [35]. The data, however, provide little evidence to indicate that Occupy precipitated a dramatic rewiring of these users' information sharing networks. While we observe significant increases in the proportion of in-group retweet and mention activity during the movement's peak, the trend suggests that these values are slowly returning to those observed before the movement's birth. What's more, in the months preceding the initial protests we find evidence indicating that these users were already highly interconnected, with more than a quarter of their directed communication (either retweeting or mentioning) involving another individual who would go on to create Occupy related content.

Taken together, these data suggest that, on Twitter, the Occupy movement tended to elicit participation from a set of highly interconnected users with pre-existing interests in domestic politics and foreign social movements. These same users, while highly vocal in the months immediately following the movement's birth, appear to have lost interest in Occupy-related communication over the remainder of the study period, and have exhibited only marginal changes in their attention allocation habits and social connectivity as a result of their participation.

These findings should not be taken to suggest that the Occupy movement itself has failed, as an argument can be made that the movement played a role in increasing the prominence of social and economic inequality in the public discourse. Though it would be unreasonable to argue that users could have maintained the frenetic pace of Occupy's earliest days, it is doubtless that supporters may have hoped for a more sustained discourse than is evident from the near-complete abandonment of these once high-profile communication channels.

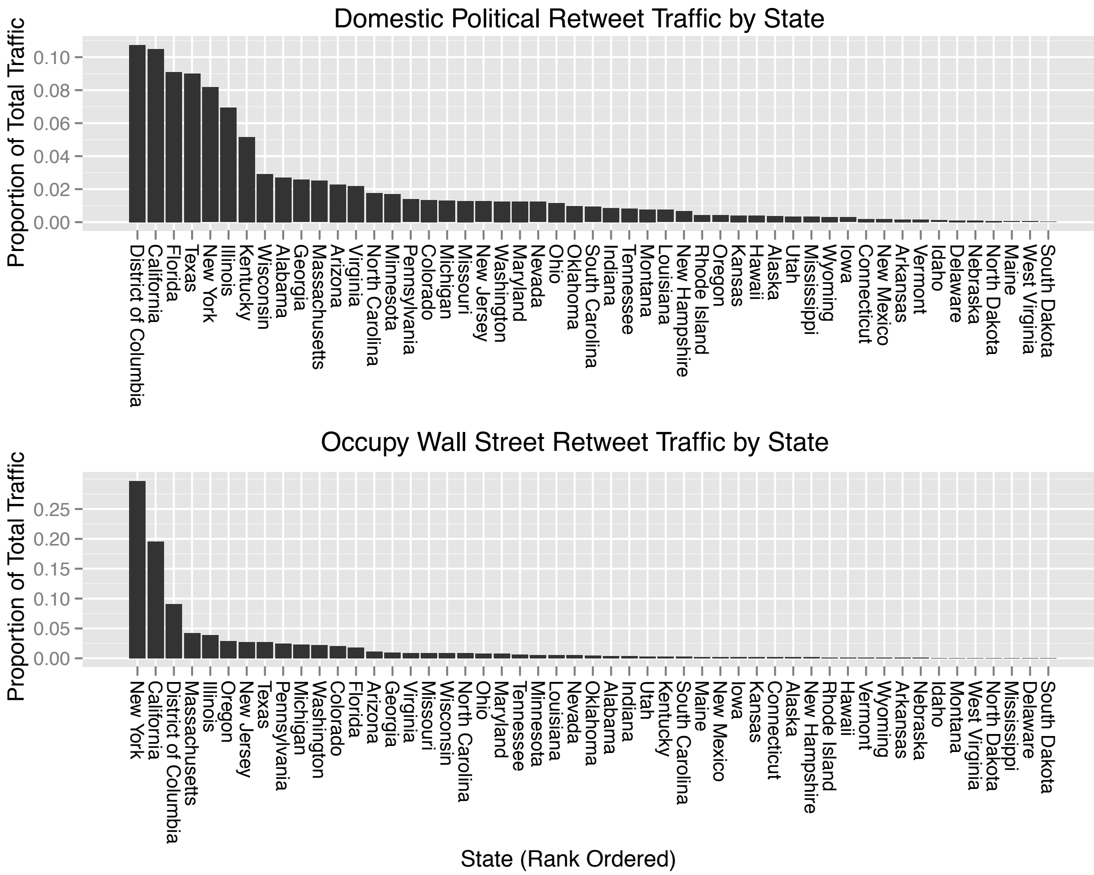

Figure 6.1: **Proportion of retweet traffic associated with each state, for each content stream.** Ordered by the amount of traffic associated with each state, it is clear that a few high-profile locations serve as the dominant sources of content in the Occupy stream. This concentration stands in contrast to the more heterogeneous activity profile for the stream of domestic political communication.

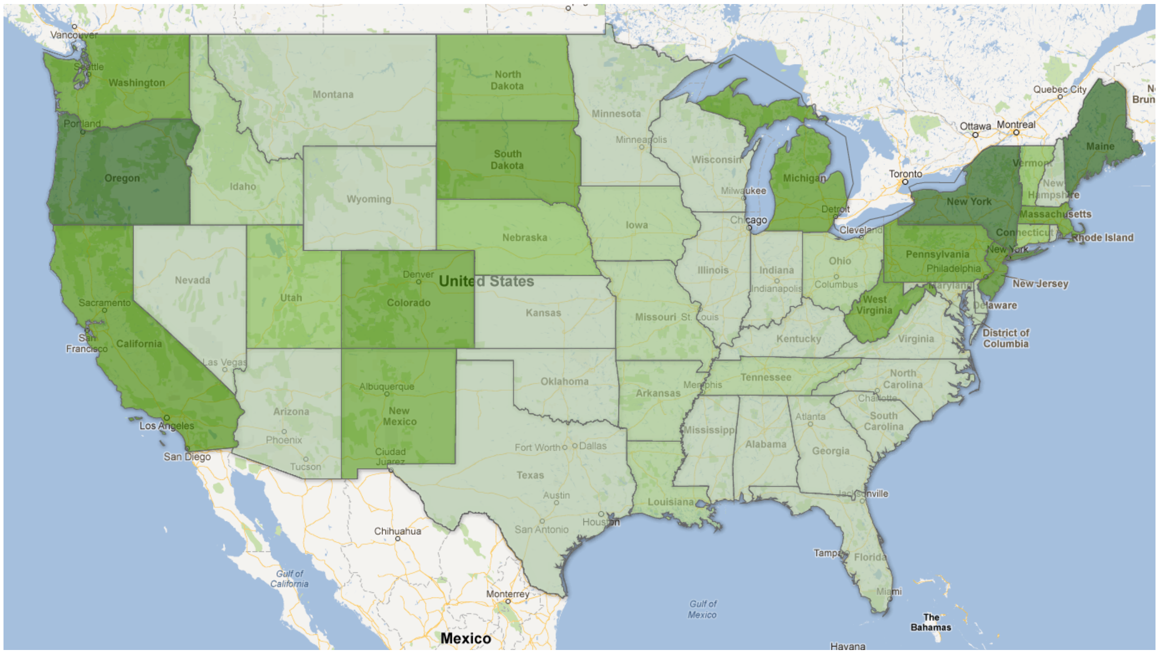

Figure 6.2: **Divergences in geographic distribution of users.** This cartogram uses color to represent the extent to which the number of Occupy Wall Street tweets in each state deviates from the domestic political communication baseline, computed as: $\frac{Occupy - Domestic}{Domestic}$. Whiter colors indicate that proportionally less Occupy content originated from the associated state, while greener colors indicate the opposite. Map data copyright Google, INEGI, 2012.

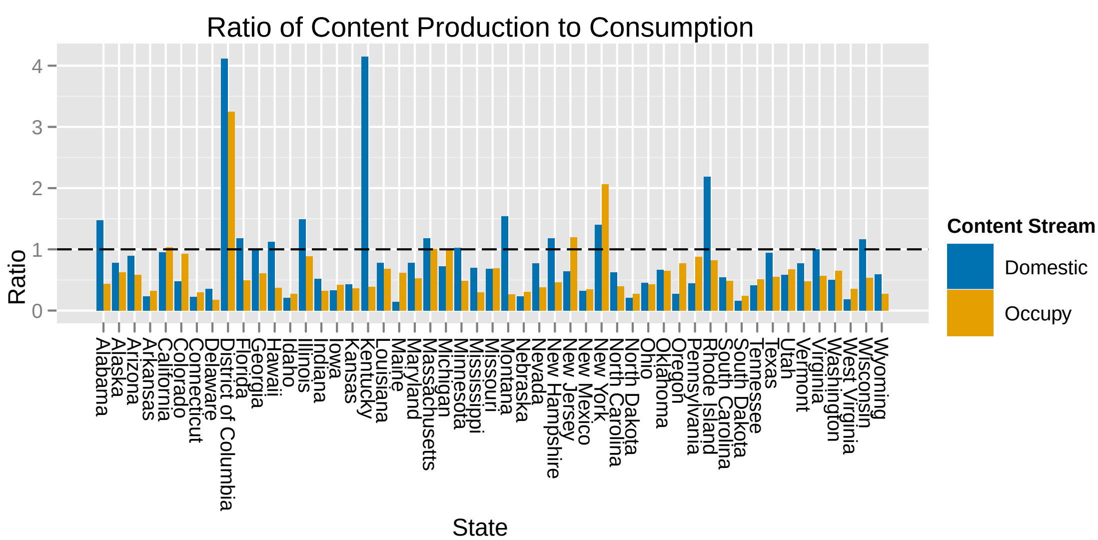

Figure 6.3: **Ratio of content production versus content consumption, by stream.** Occupy Wall Street users, by state, exhibit a lower content production to consumption ratio relative to users in the domestic political communication stream.

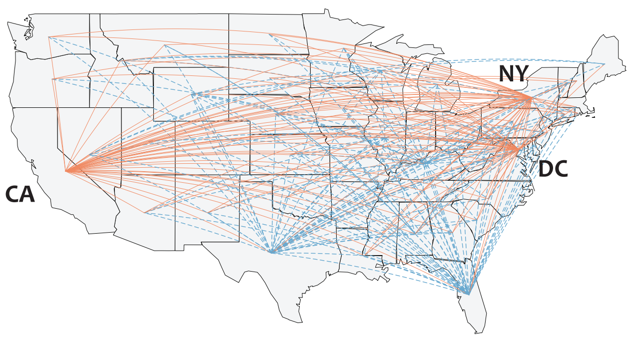

Figure 6.4: **Multiscale backbone (α = .15) of the continental interstate communication networks.** Stable domestic political communication is shown at left, Occupy Wall Street at right. Edges adjacent to New York, California, and Washington D.C. are shown in red. Note that Occupy Wall Street exhibits a clear hub-and-spoke pattern, with these locations accounting for 58% of interstate communication compared to just 28% for domestic political communication. These values are robust to different parameterizations of the multiscale backbone extraction algorithm, ranging, for $\alpha \in [0.05, 0.1, 0.15, 0.2]$, between 22.7% and 29.9% for domestic political communication and 52.7% and 61.8% for Occupy Wall Street.

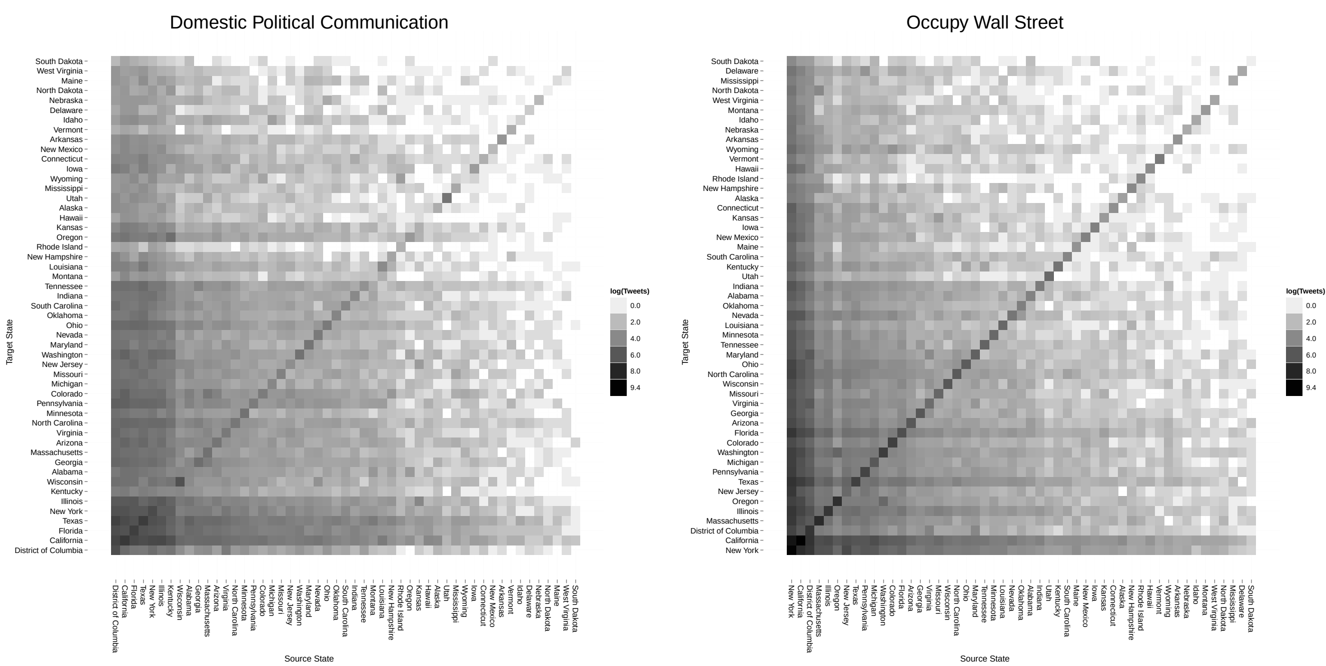

Figure 6.5: **Connectivity matrices describing directed interstate communication volume.** The edge weight corresponding to each cell is mapped to a color hue on a logarithmic scale ranging from white for edges with the least weight to black for edges with the most weight. The strong diaonalization and limited off-diagonal mass apparent in the Occupy Wall Street matrix is indicative of highly localized communication activity.

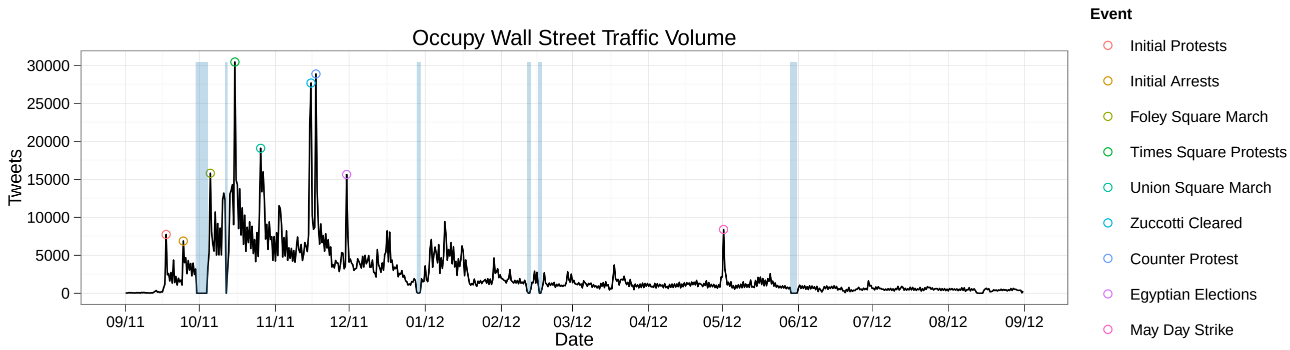

Figure 6.6: **Total number of Occupy Wall Street related tweets between September 2011 and September 2012.** Each timestep represents a twelve hour period, with vertical blue bars overlaid on periods during which access to the Twitter streaming API was interrupted. Large bursts in activity tend to correspond to on the ground protest or police action, demarcated with circles. From left to right, the events are: initial Occupy Wall Street protest in Zuccotti Park; initial NYPD arrests of protestors; march from Foley Square to Zuccotti Park; protest at U.S. Armed Forces recruiting station in Times Square; protest in support of Iraq veteran injured by police-fired projectile; NYPD action to clear Zuccotti Park; protest against eviction from Zuccotti Park; first round of Egyptian elections; 'May Day' general strike and planned reoccupation of former encampments.

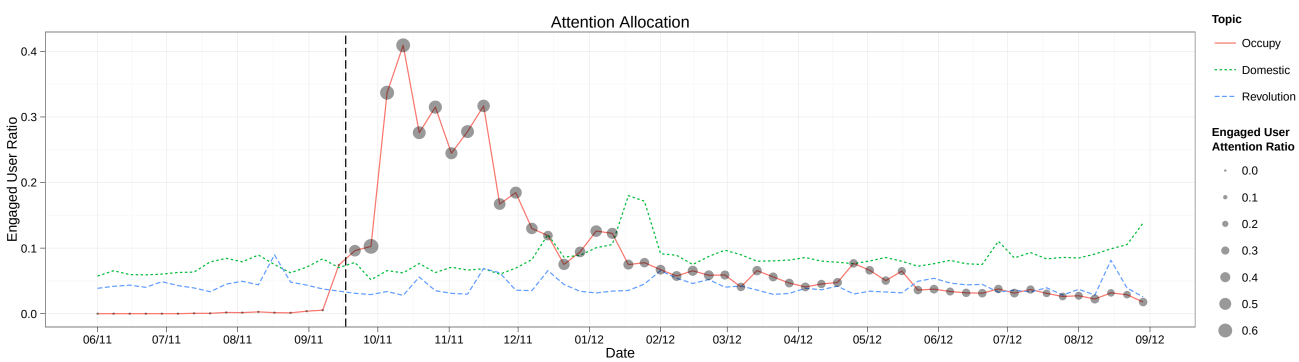

Figure 6.7: **Attention allocation of 25,000 randomly selected Occupy users to each of three topics: Occupy Wall Street, domestic politics, and revolutionary social movements.** Engaged User Ratio describes the proportion of active users in each timestep who produced at least one topically-relevant tweet. Engaged User Attention Ratio describes, among these users, the share of average attention allocated to each topic. The Engaged User Attention Ratio did not exhibit meaningful trends for either domestic politics or foreign social movements, and so it is omitted from the figure for sake of visual clarity. Refer to Section **??** for the full derivation of these measures. The dashed vertical line corresponds to the date of the first Occupy protest.

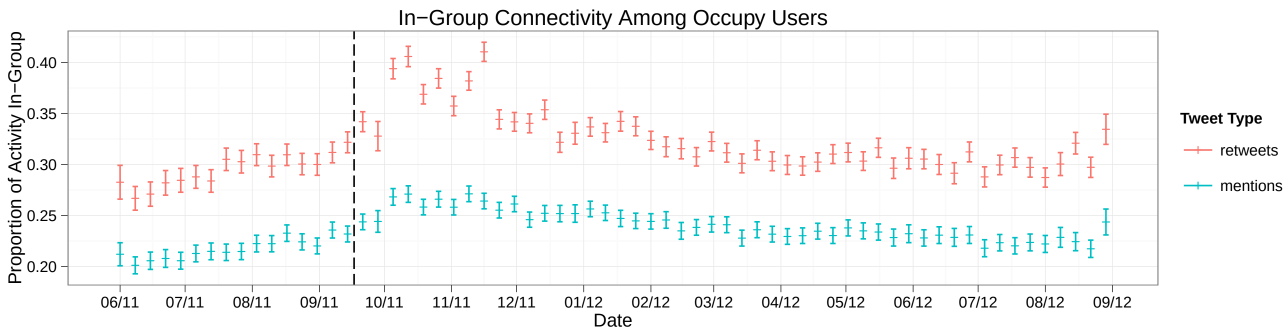

Figure 6.8: **Proportion of all retweet and mention traffic, regardless of content, from 25,000 randomly selected Occupy users involving another individual who produced at least one Occupy-related tweet.** Shown are means and 95% confidence intervals for each time step. The dashed vertical line corresponds to the date of the first Occupy protest.

Table 6.1: **Lists of tokens most overrepresented in intrastate and interstate communication.** 'Ratio', defined as $\frac{P(Token|Intrastate)}{P(Token|Interstate)}$, is small when a token is more common in intrastate traffic and large when a token is more common in interstate traffic. Terms relating to rallying supporters are more predominant in intrastate communication, while interstate traffic tends to favor terms such as protest slogans and references to the media.

Interstate		**Intrastate**	
Token	Ratio	Token	Ratio
wall	.590	city	2.254
nyc	.600	tonight	1.737
street	.699	march	1.669
news	.718	join	1.494
99%	.756	solidarity	1.387
bank	.763	day	1.354
don't	.782	square	1.333
media	.837	please	1.243
peaceful	.845	park	1.220
nypd	.847	now	1.179

Table 6.2: **Lists of topic-specific hashtags.** Hashtags were manually selected from among the 300 most frequently used by individuals in the 25,000-person random sample of Occupy users.

Domestic Politics	**Social Movements**
#tcot	#syria
#p2	#bahrain
#teaparty	#egypt
#gop	#yemen
#anonymous	#libya
#obama	#tahrir
#tlot	#wiunion
#jobs	#iranelection
#ronpaul	#assange
#romney	#wikileaks
#sopa	#jan25
#ndaa	#14feb
#obama2012	#assad
#ocra	#greece
#twisters	#damascus
#sgp	#gaddafi
#politics	#feb14
#solidarity	#scaf
#gop2012	#antisec
#p21	#arabspring
#topprog	#tunisia
#obamacare	#noscaf
#mapoli	#syrian
#acta	
#sotu	
#newt	
#santorum	
#mittromney	
#gopdebate	
#dem	

CHAPTER VII

Conclusion

"The future belongs to those who believe in the beauty of their dreams."

Eleanor Roosevelt

Questions such as those addressed in this thesis may have once been the exclusive domain of scholars in the social and political sciences. Today, however, high resolution behavioral trace data and sophisticated statistical modeling frameworks allow us to investigate these issues through a quantitative lens, clarifying, in some small measure, the once murky waters of human behavior. In this work, we have shed light on ways in which technology shapes the character of the public sphere, finding that instead of simply creating an environment characterized by increased communication volume, technology exerts unique pressures on political communication that are reflected in its large scale statistical structure.

We find that decentralized propaganda campaigns, although insidious, are betrayed by characteristic network structures that share little in common with the organic spreading processes more commonly found in the Twitter stream. Specifically, artificially constructed communication networks tend to exhibit a sparse topology, with many distinct individuals producing highly similar content. Such insights are not unambiguously to the benefit of the democratic process, however. While the machine learning techniques we employ to detect such campaigns present a promising mechanism for combating the spread of deceptive content, such approaches raise significant concerns about the algorithmic regulation of contentious political speech.

At the intersection of technology and political polarization we have staked a claim on fertile ground. Like others before, we report that digital political communication exhibits a highly partisan community structure. Leveraging this insight, we revealed that although new media savvy has been historically attributed to the political left, it is in fact right-leaning individuals who are most engaged with the Twitter platform. What's more, unlike previously studied digital media, we find that Twitter's unique design affordances lead to a network structure with unexpectedly high levels of cross-ideological interaction. Such counterintuitive conclusions underscore the potential for techniques such as those we employ to yield novel insights into human behavior at a societal scale.

Finally, we have demonstrated that challenges facing social movements, long documented in the social sciences literature, are reflected in the communication patterns of individuals engaged in domestic political action. We find that content related to narrative framing processes tends to be overrepresented in long-range communication, while content relating to resource mobilization is more prominent at the local level. This observation, derived from statistical insights into the structure and content of the Occupy Wall Street communication network, is not, so far as we can tell, documented in the social movement literature. Looking forward, it's easy to envision that such findings could provide grist for the sociological mill, leading to new theory and rigorous qualitative studies of related phenomena. In spite of this promise, however, we conclude that low-cost communication is not in and of itself sufficient for sustaining political protest. Although social media have played an important role in the development of protest movements the world over, clearly they are neither necessary nor sufficient for catalyzing social change.

Before drawing this work to a close it is pertinent to touch upon important and sometimes methodologically significant limitations of these analyses. Chief among such limitations is the fact that we have studied just one, albeit major, communication medium. There is little reason to take on faith, given the idiosyncracies of the Twitter platform, that our findings generalize cleanly to all digital political communication. To make such claims would require isomorphic studies of other communication platforms, a laborious and methodologically complex proposition. Even were such studies undertaken, one of this work's principle findings is that design affordances can affect the structure of political communication, an insight that all but ensures these observations do not hold for social communication platforms in general. Despite this fact, it's clear that Twitter, for the time being, plays a significant role in the modern political process, and this in its own right justifies the usefulness and relevance of the research findings presented herein.

Issues of causality are also pertinent in a discussion of the limitations of this work. Though we have reason to believe that political communication on Twitter is polarized, it's not clear whether this simply reflects preexisting ideological divides or whether this is the result of a dynamic process in which the widespread availability of politically homogeneous content leads to increasingly insular, polarized communities. While longitudinal studies of political communication on Twitter could provide some evidence in support of or in contradiction to this notion, establishing conclusively the presence of a casual mechanism may be a methodologically intractable proposition. Specifically, without the ability to perform controlled experimentation on the population of Twitter users it will remain difficult to distinguish results attributable to the technology from those attributable to larger-scale sociological processes.

Finally, we acknowledge that in our analysis of social movement communication we have not conclusively established the presence or absence of fine-grained ideological communities among Occupy Wall Street participants. Though we have provided compelling evidence to support the notion that these individuals tended to interact at high levels before, during, and after the movement's development, we have not excluded the possibility that there existed multiple inwardly-focused communities of individuals who began to interact with one another only after participating in Occupy Wall Street-related communication. This scenario, though unlikely, is not ruled out by the analyses presented herein, but could be interrogated through temporal analyses of the network's modularity and the topical foci of users in different network clusters.

In spite of these challenges, even the tentative, first-order discoveries made at the dawn of this new scientific era gleam with potential, revealing stores of information with the potential transform society itself. However, for all the insights into the structure and dynamics of digital political communication this work has produced, it is perhaps most satisfying on an aethsthetic and philosophical level. The most inspiring and personally rewarding moments in the study of complexity are those which betray glimpses of hidden order – otherworldy structure shimmering just beneath the surface of all that we touch. In the graceful swells of flocking birds, and the glinting, synchronized movements of schooling fish we are party to a beauty that is more than the sum of its component parts. And, in the study of human dynamics at a societal scale, are we not also party to such beauty? How miraculous it is that, in our interactions with technology, we encode digital traces that enable us to take another perspective, that of a distant observer, reflecting on the elegance, grace, and ineffable organization that characterizes the systems of the world.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Lada Adamic and Natalie Glance. The political blogosphere and the 2004 U.S. election: Divided they blog. In *Proc. 3rd Intl. Workshop on Link Discovery (LinkKDD)*, pages 36–43, 2005.

[2] E. Adar and L.A. Adamic. Tracking information epidemics in blogspace. In *Web Intelligence, 2005. Proc.. The 2005 IEEE/WIC/ACM International Conf. on*, pages 207 – 214, 2005.

[3] Sean Aday, Henry Farrel, Marc Lynch, John Sides, John Kelly, and Ethan Zuckerman. Blogs and bullets: New media in contentious politics. Technical report, U.S. Institute of Peace, 2010.

[4] L. Anália, C. Michael, W. Andrew, N. Azadeh, P. Fengxia, and S. Hagit. A linear classifier based on entity recognition tools and a statistical approach to method extraction in the protein-protein interaction literature. *BMC Bioinformatics*, 12.

[5] Sitaram Asur and Bernardo A. Huberman. Predicting the future with social media. Technical Report arXiv:1003.5699, CoRR, 2010.

[6] R. Axelrod. The dissemination of culture a model with local convergence and global polarization. *Journal of conflict resolution*, 41(2):203–226, 1997.

[7] E. Bakshy, J.M. Hofman, W.A. Mason, and D.J. Watts. Everyone's an influencer: quantifying influence on Twitter. In *Proceedings of the fourth ACM International Conference on Web search and data mining*, pages 65–74. ACM, 2011.

[8] A. Barrat, M. Barthelemy, and A. Vespignani. *Dynamical Processes on Complex Networks.* Cambridge University Press, 2008.

[9] R.D. Benford and D.A. Snow. Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, pages 611–639, 2000.

[10] Yochai Benkler. *The Wealth of Networks: How Social Production Transforms Markets and Freedom.* Yale University Press, 2006.

[11] Lance Bennett. New media power: The Internet and global activism. In Nick Couldry and James Curran, editors, *Contesting media power: Alternative media in a networked world*, pages 17–37. Rowman and Littlefield, 2003.

[12] W. Bennett. Communicating global activism: Strength and vulnerabilities of networked politics. *Information, Communication & Society*, 6(2):143–168, 2003.

[13] W.L. Bennett. Changing citizenship in the digital age. *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning*, pages 1–24, 2007.

[14] L. Bettencourt, A. Cintrón-Arias, D.I. Kaiser, and C. Castillo-Chávez. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications*, 364:513–536, 2006.

[15] Johan Bollen, Huina Mao, and Alberto Pepe. Determining the public mood state by analysis of microblogging posts. In *Proc. of the Alife XII Conf.* MIT Press, 2010.

[16] Johan Bollen, Huina Mao, and Xiaojun Zeng. Twitter mood predicts the stock market. *J. of Computational Science*, In Press, 2011.

[17] R.M. Bond, C.J. Fariss, J.J. Jones, A.D.I. Kramer, C. Marlow, J.E. Settle, and J.H. Fowler. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295–298, 2012.

[18] d. boyd, S. Golder, and G. Lotan. Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. In *Proc. 43rd Hawaii Intl. Conf. on System Sciences*, pages 1–10, 2010.

[19] J. Byrne. Occupy the media: Journalism for (and by) the 99 percent. In Janet Bryne, editor, *The Occupy Handbook*, pages 256–264. Little, Brown, 2012.

[20] David Carr. How Obama tapped into social networks' power. *The New York Times*, November 2008.

[21] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Rev. Mod. Phys.*, 81(2):591–646, 2009.

[22] M. Cha, H. Haddadi, F. Benevenuto, and K.P. Gummadi. Measuring user influence in Twitter: The million follower fallacy. In *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, volume 14, page 8, 2010.

[23] N. Chomsky. *Occupy*. Zuccotti Park Press, 2012.

[24] A. Choudhary, W. Hendrix, K. Lee, D. Palsetia, and W.K. Liao. Social media evolution of the egyptian revolution. *Communications of the ACM*, 55(5):74–80, 2012.

[25] A. Clauset, C.R. Shalizi, and M.E.J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51:661, 2009.

[26] M. Conover, J. Ratkiewicz, M. Francisco, B Gonçalves, A. Flammini, and F. Menczer. Political polarization on twitter. In *Fifth International AAAI Conference on Weblogs and Social Media*, page 89, 2011.

[27] M. Conover, J. Ratkiewicz, B Gonçalves, J. Haff, A. Flammini, and F. Menczer. Predicting the political alignment of twitter users. In *Proceedings of 3rd IEEE Conference on Social Computing (SocialCom)*, 2011.

[28] W.J. Corvey, S. Vieweg, T. Rood, and M. Palmer. Twitter in mass emergency: what nlp techniques can contribute. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media*, pages 23–24. Association for Computational Linguistics, 2010.

[29] Matthew Creamer. Obama Wins! ... Ad Age's Marketer of the Year. *Advertising Age*, May 2008.

[30] K.T. Durant and M.D. Smith. Mining sentiment classification from political web logs. In *Proceedings of Workshop on Web Mining and Web Usage Analysis of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (WebKDD-2006), Philadelphia, PA*, 2006.

[31] M. Efron. The liberal media and right-wing conspiracies: using cocitation information to estimate political orientation in web documents. In *Proceedings of the thirteenth ACM international conference on Information and Knowledge Management*, pages 390–398. ACM, 2004.

[32] Henry Farrell and Daniel Drezner. The power and politics of blogs. *Public Choice*, 134(1):15–30, January 2008.

[33] Thomas M J Fruchterman and Edward M Reingold. Graph drawing by force-directed placement. *Software Practice and Experience*, 21(11):1129–1164, 1991.

[34] R.K. Garrett. Protest in an information society: A review of literature on social movements and new ICTs. *Information, Communication & Society*, 9(2):202–224, 2006.

[35] S. González-Bailón, J. Borge-Holthoefer, A. Rivero, and Y. Moreno. The dynamics of protest recruitment through an online network. *Scientific reports*, 1, 2011.

[36] P.A. Grabowicz, J.J. Ramasco, E. Moro, J.M. Pujol, and V.M. Eguiluz. Social features of online networks: The strength of intermediary ties in online social media. *PLoS One*, 7(1):e29358, 2012.

[37] M. Granovetter. Threshold models of collective behavior. *American journal of sociology*, pages 1420–1443, 1978.

[38] J. Habermas. *The structural transformation of the public sphere: An inquiry into a category of bourgeois society.* mit Press, 1991.

[39] A. Halevy, P. Norvig, and F. Pereira. The unreasonable effectiveness of data. *Intelligent Systems, IEEE*, 24(2):8–12, 2009.

[40] M.A. Hall and L.A. Smith. Practical feature subset selection for machine learning. 1998.

[41] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The WEKA data mining software: An update. *ACM SIGKDD Explorations*, 11(1):10–18, 2009.

[42] E.S. Herman and N. Chomsky. *Manufacturing consent: The political economy of the mass media*. Pantheon, 2002.

[43] Catherine Holahan. John McCain is way behind online. *Bloomberg Businessweek*, June 2008.

[44] Bernardo A. Huberman, Daniel M. Romero, and Fang Wu. Social networks that matter: Twitter under the microscope. Technical Report arXiv:0812.1045, CoRR, 2008.

[45] Lawrence Hubert and Phipps Arabie. Comparing partitions. *Journal of Classification*, 2(1):193–218, 1985.

[46] A.L. Hughes and L. Palen. Twitter adoption and use in mass convergence and emergency events. *International Journal of Emergency Management*, 6(3):248–260, 2009.

[47] Tom Jagatic, Nathaniel Johnson, Markus Jakobsson, and Filippo Menczer. Social phishing. *Communications of the ACM*, 50(10):94–100, 2007.

[48] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury. Twitter power: Tweets as electronic word of mouth. *J. of the American Society for Information Science*, 60:2169–2188, 2009.

[49] Ashkay Java, X. Song, T. Finin, and B. Tseng. Why we Twitter: understanding microblogging usage and communities. In *Proc. of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, 2007.

[50] T. Joachims. Text categorization with support vector machines: Learning with many relevant features. *Machine learning: ECML-98*, pages 137–142, 1998.

[51] T. Joachims. Making large scale svm learning practical. 1999.

[52] D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.

[53] M. Kitsak, L.K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H.E. Stanley, and H.A. Makse. Identifying influential spreaders in complex networks. *Nature Physics*, 6:888–893, January 2010.

[54] J. Kleinberg. Bursty and hierarchical structure in streams. *Data Mining and Knowledge Discovery*, 7(4):373–397, 2003.

[55] Richard H Kolbe. Content analysis research: An examination of applications with directives for improving research reliability and objectivity. *The Journal of Consumer Research*, 18(2):243–250, 1991.

[56] Klaus Krippendorff, editor. *Content Analysis: An Introduction to Its Methodology.* Sage Publications, 2004.

[57] B. Krishnamurthy, P. Gill, and M. Arlitt. A few chirps about Twitter. In *Proc. of the first workshop on Online Social Networks*, 2008.

[58] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is Twitter, a social network or a news media? In *Proc. 19th Intl. World Wide Web Conf. (WWW)*, pages 591–600, 2010.

[59] M.B. Land. Networked activism. *Harvard Human Rights Journal*, 22:205, 2009.

[60] T.K. Landauer and S.T. Dumais. A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review; Psychological Review*, 104(2):211, 1997.

[61] T.K. Landauer, P.W. Foltz, and D. Laham. An introduction to latent semantic analysis. *Discourse processes*, 25(2):259–284, 1998.

[62] JR Landis and GG Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):154–165, 1977.

[63] J. Leskovec, L. A. Adamic, and B. A. Huberman. Dynamics of viral marketing. *ACM Trans. Web*, 1(1):5, 2006.

[64] J. Leskovec, L. Backstrom, and J. Kleinberg. Meme-tracking and the dynamics of the news cycle. In *Proc. 15th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD)*, pages 497–506, 2009.

[65] Monte Lutz. The social pulpit: Barack Obama's social media toolkit. Technical report, Edelman, 2009.

[66] J.D. McCarthy and M.N. Zald. Resource mobilization and social movements: A partial theory. *American Journal of Sociology*, pages 1212–1241, 1977.

[67] M. McPherson, L. Smith-Lovin, and J.M. Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444, 2001.

[68] M. Mendoza, B. Poblete, and C. Castillo. Twitter under crisis: Can we trust what we rt? In *Proceedings of the First Workshop on Social Media Analytics*, pages 71–79. ACM, 2010.

[69] A. Mislove, S. Lehmann, Y.Y. Ahn, J.P. Onnela, and J.N. Rosenquist. Understanding the demographics of Twitter users. In *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*, 2011.

[70] Stephen Morris. Contagion. *Rev. Economic Studies*, 67(1):57–78, 2000.

[71] E. Mustafaraj and P. Metaxas. From obscurity to prominence in minutes: Political speech and real-time search. In *Proc. Web Science: Extending the Frontiers of Society On-Line (WebSci)*, page 317, 2010.

[72] D.J. Myers. Communication technology and social movements: Contributions of computer networks to activism. *Social Science Computer Review*, 12(2):250–260, 1994.

[73] Mark E. J. Newman. Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, 74(3):036104, Sep 2006.

[74] Mark E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):26113, 2004.

[75] Brendan O'Connor, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. From tweets to polls: Linking text sentiment to public opinion time series. In *Proc. 4th Intl. AAAI Conf. on Weblogs and Social Media (ICWSM)*, 2010.

[76] J.P. Onnela, S. Arbesman, M.C. González, A.L. Barabási, and N.A. Christakis. Geographic constraints on social network groups. *PLoS One*, 6(4):e16939, 2011.

[77] Pew Internet and American Life Project. Social networking and online videos take off: Internet's broader role in campaign 2008. Technical report, Pew Research Center, 2008.

[78] Pew Internet and American Life Project. The internet and campaign 2010. Technical report, Pew Research Center, 2010.

[79] Pew Internet and American Life Project. The internet and civic engagement. Technical report, Pew Research Center, 2010.

[80] Pew Internet and American Life Project. Social media and politics in 2010 campaign. Technical report, Pew Research Center, 2010.

[81] O. Phelan, K. McCarthy, and B. Smyth. Using twitter to recommend real-time topical news. In *Proceedings of the third ACM conference on Recommender systems*, pages 385–388. ACM, 2009.

[82] K.T. Poole and H. Rosenthal. D-nominate after 10 years: A comparative update to congress: A political-economic history of roll-call voting. *Legislative Studies Quarterly*, pages 5–29, 2001.

[83] Usha Nandini Raghavan, Réka Albert, and Soundar Kumara. Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 76(3):036106, Sep 2007.

[84] J.W. Ratcliff, D. Metzener, et al. Pattern matching: The gestalt approach. *Dr. Dobb's Journal*, 7:46, 1988.

[85] Jacob Ratkiewicz, Michael Conover, Mark Meiss, Bruno Gonçalves, Snehal Patil, Alessandro Flammini, and Filippo Menczer. Truthy: Mapping the spread of astroturf in microblog streams. In *Proc. 20th Intl. World Wide Web Conf. (WWW)*, 2011.

[86] D.M. Romero, B. Meeder, and J. Kleinberg. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *Proceedings of the 20th international conference on World wide web*, pages 695–704. ACM, 2011.

[87] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proc. 19th Intl. World Wide Web Conf. (WWW)*, pages 851–860, 2010.

[88] G. Salton, A. Wong, and C.S. Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18(11):613–620, 1975.

[89] M.Á. Serrano, M. Boguñá, and A. Vespignani. Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences*, 106(16):6483–6488, 2009.

[90] C.R. Shalizi and A.C. Thomas. Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research*, 40(2):211–239, 2011.

[91] Nate Silver. *The Signal and the Noise: Why Most Predictions Fail But Some Don't.* Penguin Press, 2012.

[92] M. Steinbach, G. Karypis, V. Kumar, et al. A comparison of document clustering techniques. In *KDD workshop on text mining*, volume 400, pages 525–526. Boston, 2000.

[93] Cass Sunstein. The law of group polarization. *Journal of Political Philosophy*, 10(2):175–195, 2002.

[94] Cass R. Sunstein. *Republic.com 2.0.* Princeton University Press, August 2007.

[95] The Fox Nation. Stimulus $ for coke monkeys, 3 2010.

[96] M. Thomas, B. Pang, and L. Lee. Get out the vote: Determining support or opposition from congressional floor-debate transcripts. In *Proceedings of the 2006 conference on empirical methods in natural language processing*, pages 327–335. Association for Computational Linguistics, 2006.

[97] C.J. Tolbert and R.S. McNeal. Unraveling the effects of the internet on political participation? *Political Research Quarterly*, 56(2):175, 2003.

[98] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe. Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In *Proc. 4th Intl. AAAI Conf. on Weblogs and Social Media (ICWSM)*, 2010.

[99] J. Van Laer and P. Van Aelst. Cyber-protest and civil society: the internet and action repertoires in social movements. *Handbook on Internet Crime*, pages 230–254, 2009.

[100] Alex Hai Wang. Don't follow me: Twitter spam detection. In *Proc. 5th Intl. Conf. on Security and Cryptography (SECRYPT)*, 2010.

[101] Haixun Wang, Wei Fan, Philip S. Yu, and Jiawei Han. Mining concept-drifting data streams using ensemble classifiers. In *Proc.9th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD)*, pages 226–235, 2003.

[102] D. Watts and S. Strogatz. The small world problem. *Collective Dynamics of Small-World Networks*, 393:440–442, 1998.

[103] S. Wray. On electronic civil disobedience. *Peace Review*, 11(1):107–111, 1999.

[104] Sarita Yardi and danah boyd. Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of Science, Technology and Society*, 20:S1–S8, 2010.

[105] Sarita Yardi, Daniel Romero, Grant Schoenebeck, and danah boyd. Detecting spam in a Twitter network. *First Monday*, 15(1), 2009.