

Analyzing the impact of social media on social movements: a computational study on Twitter and the Occupy Wall Street movement

Li Tan¹ Suma Ponnam¹ Patrick Gillham² Bob Edwards³ Erik Johnson⁴

¹ School of Electrical Engineering and Computer Science
Washington State University
Richland, WA 99354
{litan, suma.ponnam}@wsu.edu

² Department of Sociology
University of Idaho
Moscow, ID 83844
gillham@uidaho.edu

³ Department of Sociology
East Carolina University
Greenville, NC 27858
edwardsr@ecu.edu

⁴ Department of Sociology
Washington State University
Pullman, WA 99164
johnsone@wsu.edu

Abstract—The extensive use of digital social media by social movement actors is an emerging trend that restructures the communication dynamics of social protest, and it is widely credited with contributing to the successful mobilizations of recent movements (e.g., Arab Spring, Occupy Wall Street). Yet, our understanding of both the roles played by social movement’s use of social media and the extent of its impact is largely derived from anecdotal evidence, news reports, and a thin body of scholarly research on web-based technologies. In this research we explore several computational methods for measuring the impact of social media on a social movement. Inspired by methodologies originally developed for analyzing computer networks and other dynamic systems, these methods measure various static and dynamic aspects of social networks, and their relations to an underlying social movement. We demonstrated the feasibility and benefits of these measurement methods in the context of Twitter and the Occupying Wall Street movement (OWS). By analyzing tweets related to OWS, we demonstrated the link between the *vitality* of the movement and the volume of the related tweets over time. We show that there is a positive correlation between the dynamic of tweets and the short-term trend of OWS. The correlation makes it possible to forecast the short-term trend of a social movement using social media data. By ranking users based on the number of their OWS-related tweets and the durations of their tweeting, we are able to identify “buzz makers”. Using a strategy similar to the page-rank algorithm, we define the influence of a user by the number of re-tweets that his/her original tweets incite. By tracing where OWS-related tweets are generated, we measure the geographic diffusion of OWS. By analyzing

the percentage of OWS tweets generated from different sources, we show that smart phones and applications such as tweet deck had been used extensively for tweeting in the OWS movement. This indicates the involvement of a younger and more technology-inclined generation in OWS.

I. INTRODUCTION

The advent of new age social media has the potential to alter the mobilization, strategies, tactics, and outcomes of social movements. By some accounts the social media like Twitter, Facebook, and Youtube played a pivotal role in the extraordinary wave of popular protests that swept the Arab world in 2011. The movement, popularly termed Arab Spring [1], is credited by many as the first social-media-enabled movement. Nevertheless, the contribution of social media to a social movement is often inferred by anecdotal evidence and news coverage [2]. What is needed is a systematical approach for analyzing the impact of social media on social movement, and eventually measuring the impact *quantitatively*. This research represents an important step towards this goal. Using Twitter and the Occupy Wall Street movement (OWS) as an example, we developed and studied several computational methodologies to examine the relation between Twitter and OWS. Our results show that there

is a close interaction between online networks in social media and offline social movement activities in the street.

Twitter started in March 2006 as an online social networking and micro blogging service. A distinctive feature of Twitter is that it limits the length of texts exchanged in its system to only 140 characters. With this limitation, Twitter users have to tweak their texts, known as “tweets”, to be concise and effective [3]. Twitter also provides a search API through which one may search and download posts with some limitations. As of February 2013, Twitter has 200 million active users creating over 400 million Tweets each day [4]. Although most messages contain little informational value, the aggregation of millions of messages can generate important knowledge on real-time events.

The Occupy Wall Street movement (OWS) started on September 17, 2011 when activists set up an encampment in Zuccotti Park in the heart of the Manhattan Financial District. Among other things, activists spotlighted greed and corruption within the financial sector, increased social and economic inequality, and the undue influence of corporations on government [5], [6]. Nurses [7], teachers [8], blue collar workers [9], as well as other workers joined in OWS. Many of them had been laid off or had their hours and benefits reduced due to state and local budget cuts. Within two months the movement had diffused across the USA and the globe. Anecdotal evidence and some preliminary research showed that OWS participants used Twitter extensively [10].

The long-term goal of our research is to develop computational metrics and tools for quantitatively measuring the impact of social media on social movements. Twitter and OWS serve as a case study showcasing the analysis methodologies that we are using, and the direction that our research is headed for. Section III describes our analysis methodologies, and our current result of applying these methodologies to study the relation between Twitter and OWS. Specifically, in Section III-A we show that the volume of tweets over time may indicate the vitality of a social movement. In Section III-B we show that by tracing dynamics of related tweets, it is possible to measure the strength of an event and forecast its trend. In Section III-C we demonstrate that we can define and identify the “buzzmakers based on the volumes of tweets that he/she generated. Using a strategy inspired by the page-rank algorithm [11], in Section III-D we explore the influence of a Twitter user involved

in OWS by measuring re-tweets that his/her tweets incite. In Section III-E we analyze the semantics of tweets by identifying popular keywords. In Section III-F we measure the geographical diffusion of OWS by analyzing the geographic origins of OWS-related tweets. Finally in Section III-G we examine the role of technology in a social movement by identifying the platforms from which OWS-related tweets are generated.

II. DATA COLLECTION AND LIMITATION

A total of 431,187 publicly available tweets were collected for this study. 237,175 tweets were aggregated for #occupywallstreet from September 21, 2011 to July 31, 2012 and 194,012 tweets for #occupy from November 23, 2011 to July 31, 2012. The #occupywallstreet hashtag was created on July 13 by the activist group Adbusters [12]. The purpose of this hashtag was to help New York City activists mobilize the September 17 occupation, and to disseminate information about local events occurring in the city once the occupation began [13]. The use of occupy hashtags spread across various cities and countries. All tweets were collected using Twitter Application Programming Interface (API) and a stand-alone desktop version of Archivist, an online Twitter archiving service [14]. Twitter search API allowed us to collect for each individual tweet a unique Twitter ID and profile picture of the tweeter, the message tweeted, and the publication time of the tweet. We also extracted the geographic location from original tweets when available, which account for less than 1% of all the tweets in our sample. By counting the tweets for each unique ID we identified top tweet generators or buzz makers. Retweets were quantified by measuring the tweets with format RT@USER. Content analysis of tweets was accomplished by ranking the most frequent occurring keyword or phrases in our sample of hashtags. Our examination of the structure and content of #occupywallstreet and #occupy tweets provides a partial snapshot of OWS during the time studied.

There are some limitations to our study. First, due to Twitter restrictions and the rate limiting factor of Twitter API our sample of tweets is limited to a maximum of 1,500 per day [15]. Thus, we are prevented from identifying peaks of Twitter activity for our hashtags and from drawing conclusions that the content of sampled tweets can be generalized to the broader population of all tweets for those days. Second, because we started collecting Twitter data after OWS began mobilizing, we missed

the first 2 weeks of mobilization activity recorded on #occupywallstreet, and the first 2 months of tweets from #occupy. Because planning for OWS started in June 2011 [16], and the hashtags for our study were started shortly thereafter, we have no data on that important time period. Third, the tweets that we collected are limited to public Twitter API and do not include nonpublic or protected tweets, neither of which can be accessed for privacy reasons. Fourth, only a limited number, less than 1% in our sample, of publicly available tweets had geographical location explicitly attached to the tweets. Finally, the lack of standard re-tweeting syntax makes quantifying re-tweets accurately difficult. Although RT@USER is a commonly used syntax for retweeting, retweets may utilize several other forms of syntax like RT:@, via@, RR@, retweet@ and thx@. Currently we only use RT@USER for quantifying retweets and exclude non-traditional syntax formats, and hence we might have missed important patterns of retweeting carried out using the other forms of syntax.

III. DATA ANALYSIS AND VISUALIZATION

A. Assessing the vitality of a social movement

The vitality of a social movement is manifested by the ability of organizers and other participants to mobilize resources and convert bystanders into participants [17]. Nevertheless, the number of participants involved in a movement are usually impossible to gauge precisely. Thus, social movement scholars typically use social movement organizations or protest events as their units of analysis. This approximation makes it easier to identify the populations of interest [18]. in the context of social networks built upon Twitter, we use hashtags as a way to define *virtual* organizations within a movement.

We assess the vitality of these *virtual* organizations by measuring the volume of tweets under related hashtags. When we started to collect #occupywallstreet tweets on September 21, 2011, OWS was in the process of quickly mobilizing new occupation groups spreading across the country [19]. The graph in Figure 1 correlates tweet volume over time for #occupywallstreet and #occupy hashtags. From the day when we started to aggregate the tweets for #occupywallstreet, the tweets for the hashtag showed a high volume of communication and remained so until early November, then it fluctuated between peaks and valleys until starting to show a gradual decline after February 2012. We speculate the decline in tweets

corresponded with a parallel decline in the number of protest events occurring in New York City the location for which the tag was originally created. The line related to #occupy in Figure 1 shows the daily volume of tweets for #occupy remained high from the time when we began collecting tweets through most of December 2011. This hashtag was created to discuss and respond to general issues related to the movement at a national and international scale. Two periods in May and June of 2012 showed renewed Twitter activity on the #occupy hashtag. This renewed activity can likely be accounted for by two major events planned by OWS organizers. The first, a May Day celebration resulted in arrests around the country, which likely kept activists tweeting at a high volume for two weeks. The second was a national occupy gathering that occurred at Independence Mall in Philadelphia. Tweets plausibly increased in the days leading up to that event as those in the online domain tracked organizing efforts.

B. Measuring the strength of an event

As news reports, activist accounts and police accounts of events often vary widely and frequently contradict one another [20], it is difficult to determine the size and intensity of an event, even after its occurrence. Using computational methodologies, social media analysis provides a relatively objective and unobtrusive means for predicting the size and intensity of an event before it occurs, and gauging the number of participants afterwards.

As noted previously, the frequency of tweets in general increased on OWS-related hashtags in response to events on ground. Over time the frequency of OWS-related tweets declined as the number of the OWS events declined. Using a qualitative method for our first approximation, we assessed whether the frequency of tweets corresponded with the occurrence of an event. We choose the “re-occupy” event in mid-December, 2011 to demonstrate the linkage between the trend of frequency of tweets and an upcoming event. On December 17 New York activists tried to reoccupy a space outside Zuccotti park to celebrate the 3rd month anniversary of the OWS movement. During the event about 50 protesters were arrested by police [21]. The graph in Figure 2 illustrates the volume of all the tweets on and around December 16, 2011 for both hashtags used in our study. It shows a high volume of tweets in the several days before the re-occupation event and a surge in the number of tweets on December 16. The relatively high volume leading up

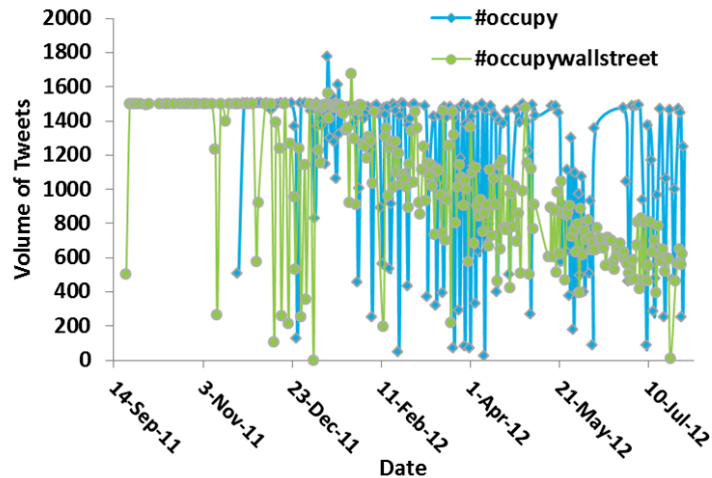


Fig. 1. Correlation of volume of tweets over time for #occupy and #occupywallstreet. A total of 194,012 tweets for #occupy and 237,175 tweets for #occupywallstreet were collected to perform the analysis.



Fig. 2. Volume of tweets over time for hashtag #occupy for the period between December 11 and December 18 2011. December 17 marks the 3rd month anniversary for OWS. The graph shows how the number of tweets started picking up from December 16 for the December 17 reoccupy event. Individual events can be anticipated beforehand by analyzing the tweets.

to the event may have reflected efforts to disseminate information about the forthcoming event and mobilize protesters. The spike in tweets on December 16 suggests that one may use the activities on Twitter to forecast the occurrence and intensity of a future event.

C. Identifying buzz makers

We define “buzz makers” as the most productive Twitter users in a group. These tweeters produce a disproportionately high number of tweets within a social media network. In our analysis we ranked users based

on the volume of tweets they tweeted for a selected period. The graph in Figure 3 shows the 25 most active users of the two OWS-related hashtags in our study. Among the 25 most active users only 20% of them actively used both #occupy and #occupywallstreet hashtags. This suggests that these buzz makers may have different preferences over the usage of these hashtags. Interestingly this relatively small group of buzz makers actively using both hashtags generate a large amount of tweets, even among buzz makers themselves.

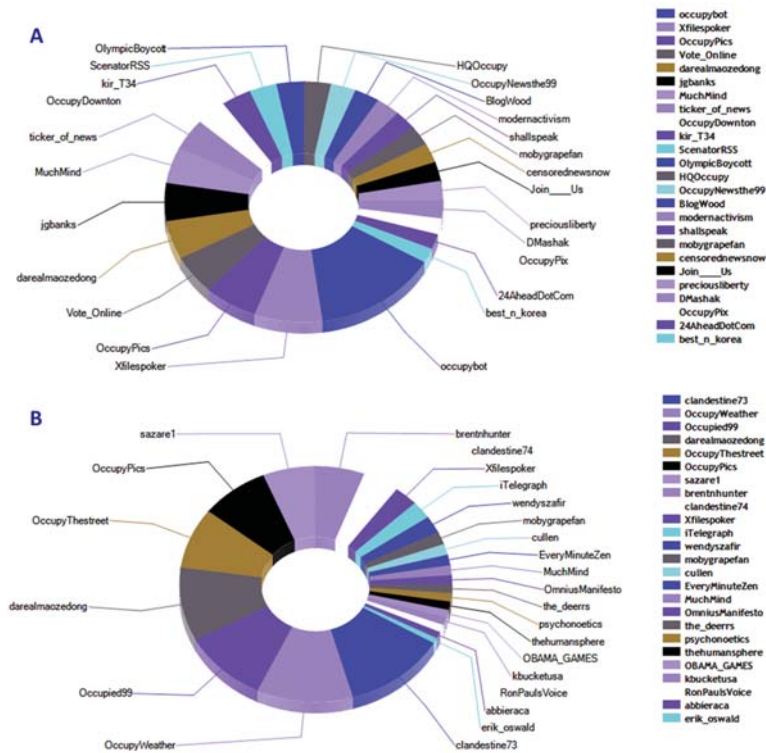


Fig. 3. Buzz Makers or top users by total tweet volume with hashtags.

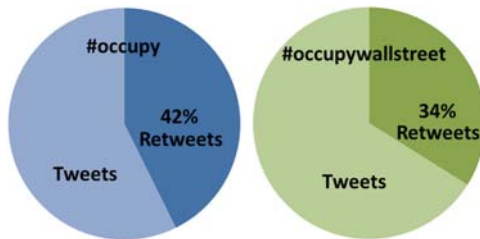


Fig. 4. % of original tweets and the tweets that were retweeted for #occupy and #occupywallstreet. Retweeting shows information flow and group thinking.

D. Measuring influence

As populated by the PageRank algorithm [11], the influence of a subject may be measured by the references made to the subject. in the context of Twitter, the influence of a Twitter user may be measured by analyzing retweets that his/her tweets incite. As seen in Figure 4 retweets account for roughly 30-40% of the total number of tweets. In case of OWS, retweets concentrate on an umbrella hashtag, #occupy. This indicates that

tweeters originated with an umbrella #occupy hashtag may appeal to a wide variety of audience than a location-specific hashtag. Furthermore, analyzing the geographic locations of tweets and retweets under umbrella and location-specific hashtags could help identify the locality of various interests in OWS.

E. Analyzing the content of tweets

An important topic in social movement research is to understand how activists’ endeavors impact the public interpretation of movement grievances, tactics and goals, and how this impact may increase solidarity among activists, resonate with potential participants, and convey a favorable impression to the broader public [22]. Having a computational methodology available to categorize, quantify and describe tweet content provides social scientists a powerful tool for identifying cultural elements of movement strategy. We hope that the keyword analysis can provide a snapshot of tweet content.

As the length of a Twitter message is restricted to 140 words, this brevity increases the weight of individual

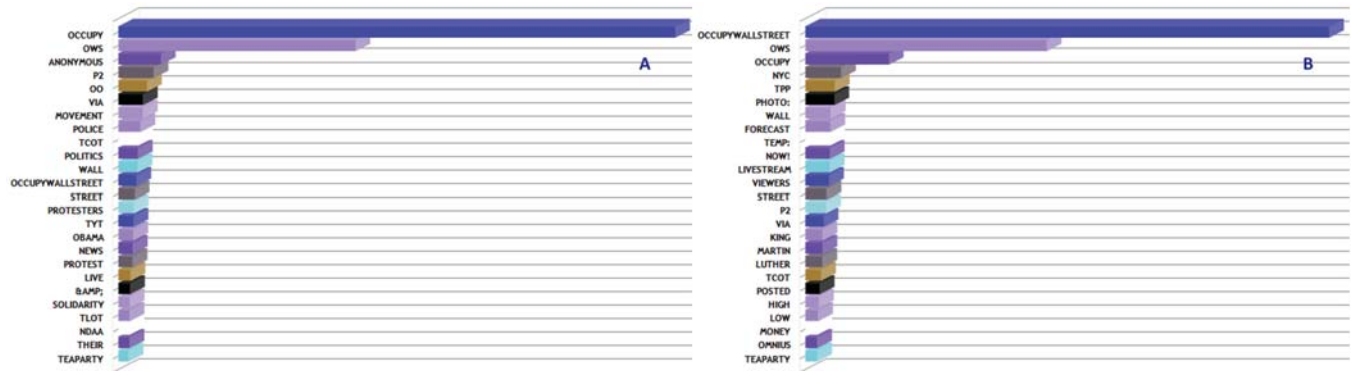


Fig. 5. The graphs show top 25 keywords used along with #occupy (A) and #occupywallstreet (B) and the number of times these keywords were used.

words in a message. One way for analyzing the context of a Twitter message is through assessing the meanings of individual words in a message. We defined the key words as verbs, nouns, and hashtags used in the messages. Figure 5 shows the top 25 keywords used in messages related to our two OWS-related hashtags, ranked by their frequencies in these messages. Five (20%) of these keywords are shared by both hashtags. Popular keywords used in key tweets with #occupy include “news”, “live”, “protesters”, “protest” and “police”. The prevalence of these key words suggests that the followers of #occupy are more interested in being updated on OWS events and the consequences of these events. In other words, tweets on #occupy functioned as a news source. In comparison, the keyword analysis on the key tweets for #occupywallstreet indicates that the followers of this hashtag are more interested in organizing events and campaigns in support of OWS.

The meanings of key words convey much information on the semantics of underlying tweets. Key words such as “forecast”, “high”, and “low” suggest that their underlying Twitter users were concerned about whether the weather was conducive to draw participants to protest events. Key words like “photo”, “livestream” and “posted” may reflect the use of online photography sites as a tactical tool, for example, taking photos or livestreams of events, and then posting them online. Keyword sequence like “King”, “Martin”, and “Luther” could reference efforts by activities to portray the movement in positive light by identifying with nonviolent protest. Further research is needed to develop a reliable tool for systematically analyzing the patterns of the content of tweets and their impact on social movements.

Analyzing the content of tweets also help assess the psychological aspect of a social movement. Human emotions play a substantial and important role in social movements and protest [23]. In 2010, Danforth *et. al* showed that the emotional states expressed in song lyrics, blogs, and presidential speeches can be quantified and analyzed. By using the Affective Norms for English Words (ANEW) [24], the emotional valence of over 1,000 words can be coded and integrated into systematic analyses. Using this method the authors graded emotional content of short texts. Currently, we are working to apply such techniques to keywords from the collected tweets. This will allow us to identify the emotional landscape of OWS tweets from happiness to sadness, joy to anger. Due to the space limitation, the result from the content analysis on emotional states will be included in the extended version of this paper.

F. Analyzing geographical diffusion

Tracking the geographical location of the tweets can help us better understand dynamics around the diffusion of social movements. Social movement scholars have longstanding theoretical interests in the diffusion and spread of social movement frames, tactics, and events [25]. Spatial and ideological proximity, as well as pre-existing networks between activists, are important elements in scholarly narratives on movement diffusion. At the same time, digital social media represent uniquely new types of networks that some envisage may free social movement activity from geographic constraints.

In this work we examine the geographic distribution of the followers of #occupy and #occupywallstreet. We



Fig. 6. Geographic distribution of the followers of #occupy(A) and #occupywallstreet

used Twitter API to aggregate the geographical location of the followers of the two OWS-related hashtags used in our study. As shown in Figure 6, the followers of #occupy and #occupywallstreet were not constrained to a specific location but spreaded widely throughout the world. Nevertheless, the geographic location does impact the use of digital social media in the occupy movement. For example, the followers of both hashtags are more likely located at a place near where ground events take place. This pattern is more apparent for the location-specific #occupywallstreet feed.

G. Identifying prevailing technology

Tweets may be posted using a variety of platforms, including mobile devices. Analyzing the source of tweets indicate how technology plays a role in a social movement. It may also be used to analyze the demographics of OWS followers on Twitter. Mobile devices tend to be favored by a younger and more technology-inclined audience, whereas traditional microblog-style tweeting is more likely used by a mature audience. The table in Figure 7 show the top 10 sources of the tweets aggregated using #occupy and #occupywallstreet. Interestingly only 31-35% of the tweets were tweeted using the Twitter website. Applications like tweet deck and tweet button are finding increased usage. Smartphone usage through iPhone, Android and Black Berry applications was much higher for tweets aggregated using #occupy.

| Rank | #Occupy | #Occupywallstreet |
|------|---------------------------|---------------------|
| 1 | Web 35% | Web 31% |
| 2 | Tweet Deck 10% | Social Oomph 8% |
| 3 | Tweet Button 8% | SAM Broadcaster 8% |
| 4 | Twitter for iPhone 7% | Tweet Button 5% |
| 5 | Best of Internet 6% | Tweet Deck 5% |
| 6 | Hoot Suite 5% | dlvr.it 5% |
| 7 | Twitter for Android 4% | Occupy Weather 5% |
| 8 | Twitter Feed 4% | Twitter Feed 5% |
| 9 | Twitter for Blackberry 2% | Livestream Stats 5% |
| 10 | dlvr.it 2% | cl-twit-repl 5% |

Fig. 7. The table shows the top 10 sources of tweets for #occupy and #occupywallstreet. Interestingly the tweets from web are only about 31-35%.

IV. CONCLUSIONS

Anecdotal evidence and news coverage indicate an increasing use of social media in social movements. Using Twitter and the Occupy Wall Street movement as an example, we applied several analysis methods to characterize and measure the influence of social media on a social movement. We showed that we can track the trend of a social movement by tracking its related tweets. Our study suggested that the volume of tweets may be used to forecast an upcoming event on ground. By ranking the volume of tweets generated by a Twitter user, we are able to identify “buzz makers”. We also performed keyword analysis on key tweets generated for two hashtags. The popularity analysis of key words provides us a glimpse on the content of and the emotion behind tweets. Our analysis of the geographic origins of tweets indicate that OWS was becoming a true global phenomenon. Finally, the analysis of the source of the tweets suggested the interest of a younger and more technology-inclined generation in the social movement. This research presents an important step towards our long term goal of building an automated and reliable tool for measuring the impact of social media on social movement. We are developing computational metrics that provide quantitative measurements of characteristics of social networks in context of a social movement.

Future Directions We plan to extend this research in several directions. One direction is to harness the synergy among the methods that we proposed. For example, we want to analyze the content of a buzz maker’ tweets, which will help us clarify the role of a buzz maker within a movement. We are also working on defining a quality metric for measuring the significance of a tweet

in the context of a social movement. The quality metric will enable us to identify significant messages from less important ones. By focusing on these significant messages, we may more efficiently and effectively identify key events that changed the course of a social movement.

REFERENCES

- [1] G. Lotan, E. Graeff, M. Ananny, D. Gaffney, I. Pearce, and D. Boyd, "The Arab Spring] the revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions," *International Journal of Communication*, vol. 5, 2011.
- [2] J. Earl, H. M. Hurwitz, A. M. Mesinas, M. Tolan, and A. Arlotti, "This protest will be tweeted: Twitter and protest policing during the Pittsburgh G20," *Information, Communication and Society*, 2013.
- [3] Twitter, "What is Twitter?" 2011. [Online]. Available: <http://twitter.com/about>
- [4] Twitter, February 2012. [Online]. Available: <http://blog.twitter.com/2013/03/celebrating-twitter7.html/>
- [5] Adbusters, "Occupy Wall Street," 2012. [Online]. Available: <http://www.adbusters.org/campaigns/occupywallstreet>
- [6] P. Gillham, B. Edwards, and J. Noakes, "Strategic incapacitation and the policing of Occupy Wall Street protests in New York City, 2011," *Policing and Society*, vol. 23, pp. 81–102, 2013.
- [7] OWS Occupy Nurses, "Occupy nurses." [Online]. Available: <http://www.facebook.com/pages/Occupy-Nurses>
- [8] California Teachers Association, "Teachers for occupy wall street." [Online]. Available: <http://www.cta.org/Issues-and-Action/School-Funding/Tax-Fairness/Occupy-Wall-Street.aspx>
- [9] OccupyWallSt-Unions.org., "Occupy wall street—unions." [Online]. Available: <http://www.occupywallst-unions.org/>
- [10] M. Orcutt, "How Occupy Wall Street occupied Twitter, too," *MIT Technology Review*, November 2011.
- [11] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks and ISDN Systems*, vol. 30, pp. 107 – 117, 1998.
- [12] Adbusters, "Occupy Wall Street: a shift in revolutionary tactics," 2011. [Online]. Available: <http://www.adbusters.org/blogs/adbusters-blog/occupywallstreet.html>
- [13] Reuters, "From a single hashtag, a protest circled the world," 2011. [Online]. Available: <http://www.reuters.com/article/2011/10/17/us-wallstreet-protests-social-idUSTRE79G6E4201110171>
- [14] Tweet Archivist, 2011. [Online]. Available: <http://archivist.visitmix.com>
- [15] Twitter, "Twitter API," <https://support.twitter.com/articles/15364-about-twitter-limits-update-api-dm-and-following>, 2012.
- [16] B. Brucato, "The crisis and a way forward: What we can learn from Occupy Wall Street," *Humanity & Society*, vol. 36, no. 1, pp. 76–84, 2012.
- [17] J. D. McCarthy and M. Zald, "Resource mobilization and social movements: A partial theory," *American Journal of Sociology*, vol. 82, pp. 1212–41, 1979.
- [18] B. Klandermans and S. Staggengborg, Eds., *Methods of Social Movement Research*, ser. Social Movements, Protest, and Contention. University of Minnesota Press, 2002, vol. 16.
- [19] N. Caren and S. Gaby, "Occupy online: Facebook and the spread of Occupy Wall Street," *Social Science Research Network*, October 2011.
- [20] J. Earl, S. Soule, and J. McCarthy, "Protest under fire? explaining the policing of protest," *American Sociological Review*, vol. 68, pp. 581–606, 2003.
- [21] D. D. Dodley and J. Solomon, "50 arrested in Occupy Wall Street demonstration," *CNN*, December 18 2011.
- [22] D. Snow, L. Zurcher, and S. Eklund-Olson, "Social networks and social movements: A microstructural approach to differential recruitment," *American Sociological Review* 45 : 787-801, 1980.
- [23] J. Jasper, "Emotions and social movements: Twenty years of theory and research," *Annual Review of Sociology*, vol. 37, pp. 285–304, 2011.
- [24] M. Bradley and P. Lang, "Affective norms for english words (anew): Instruction manual and affective ratings," The Center for Research in Psychophysiology, University of Florida, Technical Report C-1, 1999.
- [25] R. K. Givan, K. M. Roberts, and S. A. Soule, *The Diffusion of Social Movements*. Cambridge University Press, 2010.