Received 3 December 2015
Revised 18 May 2016
Accepted 18 May 2016

# Hashtags and information virality in networked social movement

## Examining hashtag co-occurrence patterns

Rong Wang and Wenlin Liu
*Department of Communication, University of Southern California, Los Angeles, California, USA, and*
Shuyang Gao
*Department of Computer Science, University of Southern California, Los Angeles, California, USA*



### Abstract

**Purpose** – The purpose of this paper is to conceptualize the use of Twitter hashtag as a strategy to enhance the visibility and symbolic power of social movement-related information. It examined how characteristics of hashtag drove information virality during a networked social movement.

**Design/methodology/approach** – Twitter data from two days during the Occupy Wall Street Movement in 2011 were collected. With network analysis, the authors identified popular hashtag types and examined hashtag co-occurrence patterns during the two contrasting movement days. It also provides a comparative analysis of how major types of viral hashtag may play different roles depending on different movement cycles.

**Findings** – The authors found that the role of hashtag influencing information virality may vary based on the context of the tweets. For example, movement participants applied more strategic hashtag combinations during the unexpected event day to reach different social circles. Consistent patterns were identified in mobilizing influential actors such as public figures. Different use patterns of media outlet hashtag were found across the two days.

**Originality/value** – Implications on how hashtag type and event dynamics may shape hashtag co-occurrence patterns were discussed.

**Keywords** Network analysis, Online information, Networked social movement, Occupy Wall Street, Twitter hashtags, Virality

**Paper type** Research paper

### Introduction

In recent years, social movement and collective action is experiencing a deep technological and organizational transformation (Castells, 2012). We are witnessing the emergence of networked social movements and a paradigm shift from "collective" to "connective" action (Bennett and Segerberg, 2012). Social movements are increasingly facilitated by the use of social media, such as Facebook and Twitter. These networked platforms not only enable social movement organizations to mobilize geographically dispersed publics, but also afford new ways of citizen engagement through personalized information sharing (Bennett and Segerberg, 2012).

The shift to networked social movement, defined as the type of social movement relying on personal networks and networked communication technologies to coordinate action and achieve goals (Castells, 2012; Juris, 2004), has made the concept of virality of particular relevance. Virality is an important characteristic pertinent to the process of social information flow, defined by the speed of information

Online Information Review
Vol. 40 No. 7, 2016
pp. 850-866

1468-4527
DOI 10.1108/OIR-12-2015-0378

spread, the reach in terms of the number of people exposed to the content, and the distance the information travels to bridge multiple networks (Nahon and Hemsley, 2013). Central to the notion of virality are the capacity of individuals and organizations to share information and successfully mobilize collective attention, as well as the ability for messages to connect diverse networks.

This paper analyzes the virality of social movement messages from the lens of strategic hashtag use on Twitter. The literature on hashtag use tends to focus on using hashtag for sampling or analyzing the diffusion of one particular hashtag to uncover the dynamics of a social movement (Lotan *et al.*, 2011; Papacharissi and de Fatima Oliveira, 2012). Less is known regarding what types of hashtags are more like to be used in organizing a social movement. By identifying viral hashtags during Occupy Wall Street (OWS), and comparing the hashtag co-occurrence patterns (i.e. which types of hashtags tend to be used together for tweeting) from selected social movement days, this study first inductively constructs a typology of viral hashtags defined by the daily frequency. It then examines how different characteristics of hashtags may explain their structural virality in the hashtag co-occurrence network, and how the role may manifest itself differently during a social movement.

This paper makes two contributions to the current theory and research on networked social movement and online information dissemination. First, the current study applies the virality framework and network modeling to uncover how information characteristics of hashtag may influence their co-occurrence patterns on Twitter. Moving beyond a descriptive analysis, the approach allows us to systematically examine how OWS participants used viral hashtags. Second, through a comparative analysis, this research demonstrates how the use patterns of viral hashtags may differ depending on event dynamics in a networked social movement. It further provides evidence that the role of viral hashtags in influencing online information sharing may change as a social movement evolves over time.

## The paradigm shift of social movement

With the rapid emergence of internet and digital technologies, literature on social movements has experienced a paradigm shift, particularly with regard to ways of theorizing citizens' movement engagement, as well as criteria of evaluating movement effectiveness. First, contemporary media environment has challenged the necessity of hierarchical and formal ways of organizing collective action, and have placed a greater emphasis on individual contributor's autonomy and self-organizing capability (Bimber *et al.*, 2012; Gamson and Sifry, 2013). Individuals' participation is no longer limited to the pursuit of explicit political goals offline. The spectrum of activities counted as "contributing" is much extended in the digital era.

Bennett and Segerberg's (2012, 2013) notion of "connective action" further articulates this shift. As they maintain, the use of networked digital media has enabled a distinct logic of "connective action," under which joining collective causes can be realized through personalized expression and a self-validating mechanism of information sharing through personal networks online. Although the argument of "slacktivism" has lamented that networked digital media may breed the "feel-good activism" that has little social or political impact (Morozov, 2011), the new logic of connective action operates in a fundamentally different way than traditional social movement (Gerbaudo and Trere, 2015). Earl and Kimport (2011) studied the "e-tactics" of social movements, utilizing online petitions and boycotts to mobilize public opinion. Thorson *et al.* (2010) exemplified a video-based activism facilitated by YouTube. Instead of relying on formal organizing, newer forms of social movements nowadays tend to employ a large scale but loosely connected network, where individuals can easily join or disassociate (Kavada, 2015).

The second paradigm shift lies in the criteria of evaluating movement outcomes. Traditional social movement literature focuses on evaluating political effectiveness, the extent to which a movement successfully mobilizes resources and achieves stated goals (Jenkins, 1983; McCarthy and Zald, 1977). Networked social movements entail a high level of personalized collective action (Bennett and Segerberg, 2012). Following this logic, we argue that rather than emphasizing an social movement's material outcomes in the political sphere, an alternative way is to evaluate its capacity to attain symbolic power, defined as "the capacity to intervene in the course of events and influence the actions of others by means of the production and transmission of symbolic forms" (Thompson, 2005, p. 50). In other words, what matters is how a social movement can effectively disseminate key values. To further illustrate this point, we introduce the virality framework.

## The virality framework

Nahon and Hemsley's (2013) framework of virality provides a valuable theoretical lens to understand the process of obtaining symbolic power for networked social movements. Virality is defined as "a social information flow process where many people simultaneously forward a specific information item, over a short period of time, within their social networks, and where the message spreads beyond their own (social networks) to different, often distant networks, resulting in a sharp acceleration in the number of people who are exposed to the message" (Nahon and Hemsley, 2013, p. 16). This framework distinguishes two mechanisms of viral process: a bottom-up process, during which the viral process is driven by individuals or organizations that intend to spread the content; a top-down process, during which virality is designed by content makers and promoted by powerful gatekeepers such as mainstream media or political elites. In networked social movements, virality is more often driven by the bottom-up mechanism given movements' heavy reliance on self-motivated participants in disseminating movement messages (Castells, 2012).

The bottom-up virality mechanism consists of two major forces (Nahon and Hemsley, 2013). First of all, information characteristics influence whether and how people will share certain content within their social networks, specifically the factor of salience and relevance. The more salient a message, and the more relevant the context, the higher the possibility of sharing. The second factor is the structures of networks, which include the rules, practices, and arrangements that regulate people's behaviors in networks, and how people are connected via social relations.

Networked social media, particularly Twitter, play a crucial role in online information dissemination, reaching a broader spectrum of audience, and ultimately assisting networked social movements to gain virality (Lotan *et al.*, 2011). To examine Twitter's role in promoting virality, we focus on one important technological feature of Twitter, the hashtag. In the following section, we conceptualize the use of Twitter hashtag as a mechanism to mobilize collective attention, and discuss how different types of hashtags are related to virality.

## Twitter hashtag as attention mobilizer and its role in gaining virality

Hashtag, a word or a phrase prefixed with the symbol #, is a technological feature afforded by Twitter. Originally created as the "channel tags" to allow users to join a particular conversation (Messina, 2007), hashtag has been noted for its instrumental role in publicizing social issues (Bruns and Burgess, 2011). More recent literature has conceptualized hashtag as a thematic identifier, which not only signals relevance with specific known issues, but also enables effective dissemination beyond one's "follower-following" network (Ma *et al.*, 2013). With the strategic use of hashtag, the tweeted content can be broadcast to broader audiences. The choice of using a particular hashtag is influenced by two overlapping processes: deliberatively seeking attention from interested users (Bruns and Burgess, 2011), and contagion process driven by the virality of certain hashtags (Romero *et al.*, 2011).

The strategic use of hashtag on Twitter has increasingly been observed in recent social movements. During Arab Spring, hashtags like #egypt #jasminerevolution, and #jan25 have rapidly gained popularity and facilitated the creation of "ad hoc issue public" (Meraz and Papacharissi, 2013, p. 144), a group of concerned actors gathered with the emergence of a certain issue. Using a shared set of hashtags, a distributed community was able to locate, self-organize, and collectively contribute to the information stream about a networked social movement (Bennett and Segerberg, 2012; Bruns and Burgess, 2011; Meraz and Papacharissi, 2013). While the information stream becomes an important mechanism that connects multiple actors, the use of hashtag undoubtedly facilitates the information flow to the targeted audience.

Hashtag also affords alternative ways of enhancing movement visibility. Prior to networked digital media, the visibility of social movement is largely contingent on its "news-making" ability (Andrews and Caren, 2010). Social movements are increasingly observed to rapidly scale up through networked information dissemination. Under this background, we reference Tufekci's (2013) notion of "attention," defined as "the means through which a social movement can introduce and fight for its preferred framing, convince broader publics of its cause, recruit new members […] and mobilize its own adherents" (p. 849). We argue that hashtag use can serve as an important "attention mobilizer" for networked social movement, and the strategic use of hashtag may increase the virality of movement messages.

Using hashtag to achieve virality can be first understood by Twitter's affordance of visibility. Affordance is defined as the ability of a certain technology that emerges from the interaction between the material qualities of the technology, and individuals or organizations' perception of its utility (Majchrzak and Markus, 2012). The visibility affordance means that Twitter allows users to make their behaviors and communication networks visible to others across explicitly differentiated boundaries (Treem and Leonardi, 2013). The use of hashtag enhances the visibility of a message as the tweet becomes more searchable than plain texts alone (Small, 2011). Visibility is also essential for attaining symbolic power, as it facilitates information diffusion in a rapid and large scale (Castells, 2007; Thompson, 2005).

The role of hashtag in affecting virality of an event can also be understood from Twitter's affordance of generative role taking, the ability of allowing users to participate in decentralized conversation (Majchrzak *et al.*, 2013). Twitter users are autonomous in choosing what hashtag to use. With a deliberate absence of formal leadership, the strategic use of hashtag helps to construct the information-sharing network and bridge diverse social groups with a common interest. The evolving nature of the hashtag use networks facilitates the sharing of collective outrage and hope (Allagui and Kuebler, 2011).

While existing studies examined hashtag use in networked social movements (Sharma, 2013), most of the literature focuses on the diffusion of one particular hashtag or the role of same genre of hashtags in mobilizing supporters. For example, Lotan *et al.* (2011) found that Twitter users organized the Arab Spring around #sidibouzid in Tunisia and Egypt. Similarly, Zamir (2014) analyzed how protest information diffused through #Shahbag in Bangladesh. By mapping the network among users of this hashtag, Zamir (2014) was able to identify influential protesters and their information diffusion behaviors. Less is known regarding how hashtag characteristics, such as hashtags of different nature or thematic types, may carry varied levels of capacity to gain virality.

The bottom-up mechanism suggests that information characteristics of hashtags may explain the level of virality (Nahon and Hemsley, 2013). Twitter's 140 characters rule constrains what to post and how to make the content concise. Hashtags are used to tag tweets by topics so that people can follow live and archived conversations (boyd *et al.*, 2010). As a structural marker in various interest networks centering, hashtag plays an important role in the flow of information and it influences the emergence and the eventual magnitude of a viral message. The virality of a hashtag can be measured by the frequency of use. To understand what roles hashtags may play in explaining the heterogeneity of message virality, we propose a general research question:

*RQ1*. What major types of Twitter hashtags are frequently used during a networked social movement?

Recent literature on digital activism has argued that the simple adoption of social media does not provide necessary conditions for the viral diffusion of messages (Gerbaudo, 2015). Measuring the frequency of hashtag use thus may not capture the multidimensionality of virality. To further uncover how hashtag characteristics explain the virality process, we introduce the concept of hashtag co-occurrence network. In the virality framework, the use of social media helps construct an ad hoc interest network, "temporally bound, self-organized networks where membership is based on an interest in the information content or an interest in being included in the interest network of others" (Nahon and Hemsley, 2013, p. 31). On Twitter, the interest network can arise from the co-occurrence of different hashtags in one tweet.

We define "hashtag co-occurrence network" as a network composed of nodes representing different hashtags, and edges representing instances of co-appearance of any two hashtags in one tweet. It is a weighted network, with edge value indicating how often any two hashtags co-occur. The hashtag co-occurrence network is analogous to author co-citation network in bibliometrics (White and Griffith, 1981). Specifically, a hashtag is analogous to an author, and each tweet is analogous to a research study where multiple authors are cited. The connection pattern of a co-citation network can shed light on the intellectual structure of a knowledge domain (Ding *et al.*, 1999). As for hashtag co-occurrence network, we argue that it can help understand the bottom-up effort of diffusing information. Through the combination of different hashtags, movement participants engage in personalized communication and information sharing (Gerbaudo and Trere, 2015; Ma and Li, 2014).

Hashtag helps users contribute to discussions of existing topics and to locate relevant conversations from the vast amount of topics on Twitter (Huang *et al.*, 2010). Analyzing the structure of a "hashtag co-occurrence network" helps to identify relative prominence of individual hashtags in disseminating movement messages and what strategies are taken by movement participants to speed up information diffusion. As each hashtag represents a circle of like-minded users, the strategic combinations of hashtags allow social movement participants to mobilize public attention from different social circles efficiently, thus achieving higher virality (Gleason, 2013).

From a network perspective, we explore another dimension of hashtag virality based on its structural position in the hashtag co-occurrence network, measured by degree centrality (Hanneman and Riddle, 2006). The higher the degree centrality, the higher frequency that a hashtag is used in combination with others. Centrally located hashtags entail higher symbolic salience as they are more likely to be associated with other hashtags in a tweet to reach broader audiences. This dimension of virality distinguishes itself from the prior measure of frequency. We refer this dimension as "structural virality" to reflect the structural feature of a hashtag in the co-occurrence network.

To identify the types of hashtag that are more likely to gain structural virality, we propose the following research question:

*RQ2*. What types of hashtags are more likely to gain structural virality during a networked social movement?

## Case background

This study applies the virality framework to analyze the OWS movement, a networked social movement that spread to over 100 cities in the USA and over 1,500 cities globally. Beginning with the September demonstration on the Wall Street and the protest at the Zuccotti Park, the series of occupy events showed strong evidence of virality. Virality emerges from social processes within which people share content with each other without explicit leadership (Nahon and Hemsley, 2013). Three key elements of OWS make it a good case study for virality. First, the OWS movement spread to a global level over a short period of time (Castells, 2012; Chase-Dunn and Curran-Strange, 2012). Second, OWS reached to a large population through both new media platforms and traditional media (Costanza-Chock, 2012). Online social networking tools such as Twitter were proved to be a powerful tool to rapidly reach a broad range of audience (Gaby and Caren, 2012). Third, OWS was driven by human and social aspects of information sharing, through multiple forms of communication networks (Castells, 2012).

## Method

### *Data*

Data for this study were collected using the following procedures. First, to create the overall tweet corpus, a set of internet data crawling algorithms was developed to detect and archive tweets relevant to OWS that began in September 2011. Specifically, we used the Gnip PowerTrack service, a third-party commercial reseller of Twitter data, to collect tweets in real time. The advantage of using Gnip rather than relying on Twitter's public APIs is that this commercial service enabled the access to a fuller volume of Twitter activity than what was available through the public APIs (Groshek and Al-Rawi, 2013; Thorson *et al.*, 2013). 381 keywords were fed into PowerTrack to identify and collect tweets that matched at least one of the specified keywords (the "OR" logic was used). The entire set of keywords was developed in an evolving fashion. New terms and phrases were constantly added in response to emergent events. The types of keywords included hashtags, @ mentions, and certain words or phrases in the text to ensure the inclusivity of our data set. Some example keywords were: #occupy, #ows, move your money, ows, occupy, occupy movement, occupy together, occupy Wall Street, and we are the 99.

For the current study, only tweets from two days (November 2 and 18 in 2011, Pacific Time) were selected. In total 89,823 tweets were collected from November 2 and 102,104 tweets were collected from the November 18. These two days were chosen with a particular comparison goal in mind. First, November 2 was a regular weekday, which allowed us to see the general trend of hashtag being used during an ordinary OWS day. Second, one viral event, the Pepper Spray incident at the University of California, Davis (UC Davis), took place around 4 p.m., November 18. Thousands of photos and videos of University Police Lieutenant spraying students were shared on social networking sites within seconds and spread to traditional mainstream media too. Focusing on a specific day with an unplanned event enabled us to analyze what hashtags retained over the event day, what new hashtags emerged, and how the hashtags were used in combination to spread breaking news. Finally, the selection of two different days allowed us to draw implications of how the hashtag use patterns may evolve with event dynamics. It is worth noting that this current research is not intended to provide a comprehensive picture of how OWS participants utilized hashtags during the movement. It aims to unpack how event dynamics and information characteristics during different movement days are related to the hashtag use patterns.

Given the focus was on hashtag use, the second step was to select all the tweets that contained hashtags. 61,616 tweets from November 2 and 101,551 tweets from November 18 were kept for further analysis. Third, 8,868 unique hashtags were extracted from the November 2 data set and 10,343 hashtags from the November 18 data set. To focus on viral hashtags, the top 5 percent of hashtags were selected from both days. This generated a list of 505 most frequently used hashtags from November 2, and 510 from November 18, after removing a number of hashtags that contained foreign languages or unidentifiable texts. Finally, all the selected hashtags were rank ordered by the frequency of use in each data set.

*Hashtag typology development*

Upon constructing the final viral hashtag list, the first two authors coded all the hashtags and inductively developed ten categories (Table I), referencing Agarwal and *et al.*'s (2014) study on Twitter and Occupy protest. The final categories include: geolocation, hashtags containing explicit country, city, or other geographic markers, such as #Boston; movement theme, including major themes of the movement, such as #wallstreet and #protest; government organization, including hashtags like #congress or other identifiable governmental bodies; civil society organization, including non-governmental and non-profit organizations; identity claim, reflecting the value or motifs of the movement as seen in public discourse, such as #weare99, #99percent; public figure, including any

**Table I.** Summary of the most frequently used hashtags on both days

| Hashtag category | November 2, 2011 Number of hashtag | November 2, 2011 Total frequency (percentage) | November 18, 2011 Number of hashtag | November 18, 2011 Total frequency (percentage) |
|---|---|---|---|---|
| Civil society | 10 | 2,030 (1.93) | 17 | 3,078 (3.93) |
| Economy | 29 | 10,495 (9.99) | 16 | 983 (1.26) |
| Event | 49 | 21,034 (20.02) | 50 | 5,423 (6.93) |
| Geolocation | 134 | 34,015 (32.37) | 114 | 18,024 (23.03) |
| Government | 22 | 2,353 (2.24) | 25 | 4,657 (5.95) |
| Identity | 32 | 7,583 (7.22) | 37 | 7,922 (10.12) |
| Media outlet | 37 | 3,082 (2.93) | 44 | 4,031 (5.15) |
| Public figure | 15 | 1,655 (1.57) | 14 | 1,226 (1.57) |
| Theme | 73 | 8,261 (7.86) | 65 | 6,925 (8.85) |
| Time | 9 | 1,198 (1.14) | 16 | 9,380 (11.98) |
| Other | 95 | 13,380 (12.73) | 112 | 16,630 (21.24) |

identifiable politicians or celebrities; economy, which was particularly created because a central theme of OWS was to combat economic inequality; time, a category with explicit time marks, such as #nov2; media outlet, including clear reference to media organizations, such as #foxnews; and event, including specific OWS-related activities. All the hashtags were coded in one of these categories, with an acceptable level of intercoder reliability ($\alpha = 0.85$). The two coders discussed differences in their codings and agreed on the final coding which was used in the further data analysis.

*Hashtag co-occurrence network construction*

The hashtag co-occurrence network was constructed to analyze the use patterns of hashtags during the two selected dates. We define "co-occurrence" by the incidence of any pair of hashtags used together in the same tweet. For instance, if "#pepper" and "#ucdavis" appeared in one tweet, a tie between these two hashtags was recorded. The strength of tie was indicated by the total number of unique tweets where each pair of hashtags co-occurred within one selected day. Following this procedure, two valued, non-directional viral hashtag co-occurrence networks were constructed: one with the top 5 percent of hashtags from November 2; and one with the top 5 percent of hashtags from November 18.

*Data analysis*

A descriptive analysis of hashtag types was conducted for both days to answer *RQ1*. To answer *RQ2*, Exponential Random Graph Modeling (ERGM) was conducted to examine what types of hashtags are more likely to co-occur. ERGMs use simulation techniques to make inferences about how the hypothesized parameters may shape the observed network configurations, by comparing the propensity of a network structure in the observed network to the propensity that would occur at random (Lusher *et al.*, 2013). The following parameters were included in the ERGM models: hashtag homophily, to test whether the same type of hashtags is more likely to be used together; frequency, to test whether the popularity of a hashtag may significantly influence the odds of getting paired with; and each of the ten specific hashtag types, to test which type of hashtags is more likely to be used with other types.

Data analysis was done with the ERGM package in R. The model fits the data when all parameters have $t$ less than 0.10 (Snijders *et al.*, 2006), indicating that the standard error is within a tolerable range. Parameters are significant when the $t$ values are within 1.96 standard errors of the parameters (Robins *et al.*, 2007). Gephi was used for network visualizations and descriptive statistics. Given the relative dense structure of the two networks, the ForceAtlas 2 layout algorithm. This force-directed method follows a simple principle: linked nodes attract each other and non-linked nodes are pushed apart. With a variety of settings such as preventing overlap, adjusting gravity, and dissuading hubs, it allows for more readable visualization.

## Results

In this section, we first reported the descriptive of each data set (Table I) to answer *RQ1*, including the number of unique hashtags and the total frequency of each category, which shows the relative popularity across all the categories. Then we reported characteristics of hashtag co-occurrence networks (Table II), along with the ERGM results to answer *RQ2* (Table III).

*RQ1* examined what types of Twitter hashtags were most frequently used during a networked social movement. Results (Table I) showed that on November 2, 2011, geolocation hashtags had the highest total frequency ($n = 34{,}051$), followed by event (21,034), economy ($n = 10{,}459$), and theme ($n = 8{,}261$). On the November 18 data set, geolocation still remained as the most popular ($n = 18{,}024$), followed by time ($n = 9{,}380$), identity claim ($n = 7{,}922$), OWS theme ($n = 6{,}925$), and event ($n = 5{,}423$). There was a significant increase of use for time, government organizations, identity, media outlet, and OWS theme hashtags. However the following categories experienced decreased use: economy, event, and geolocation.

*RQ2* examined what types of hashtags are more likely to gain structural virality. When constructing the co-occurrence network of top hashtags on November 2, five hashtags were identified as isolates which did not appear with other hashtags. Therefore, the network size was 500, with a density of 0.08 and degree centralization of 0.74. The high centralization level suggested that when users chose what hashtags to include, a small segment of hashtags received disproportionately high level of preference. The betweenness centralization of November 2 network was 0.24, closeness centralization was 0.69, and transitivity was 0.27, indicating a relatively sparse and central network, with a moderate transitivity (Figure 1). Top 50 nodes were labeled, of which degree centrality scores ranged from 76 to 408.

The construction of the co-occurrence network among top hashtags on November 18 identified 31 isolates. Therefore, the network size was 479, with a density of 0.04, and

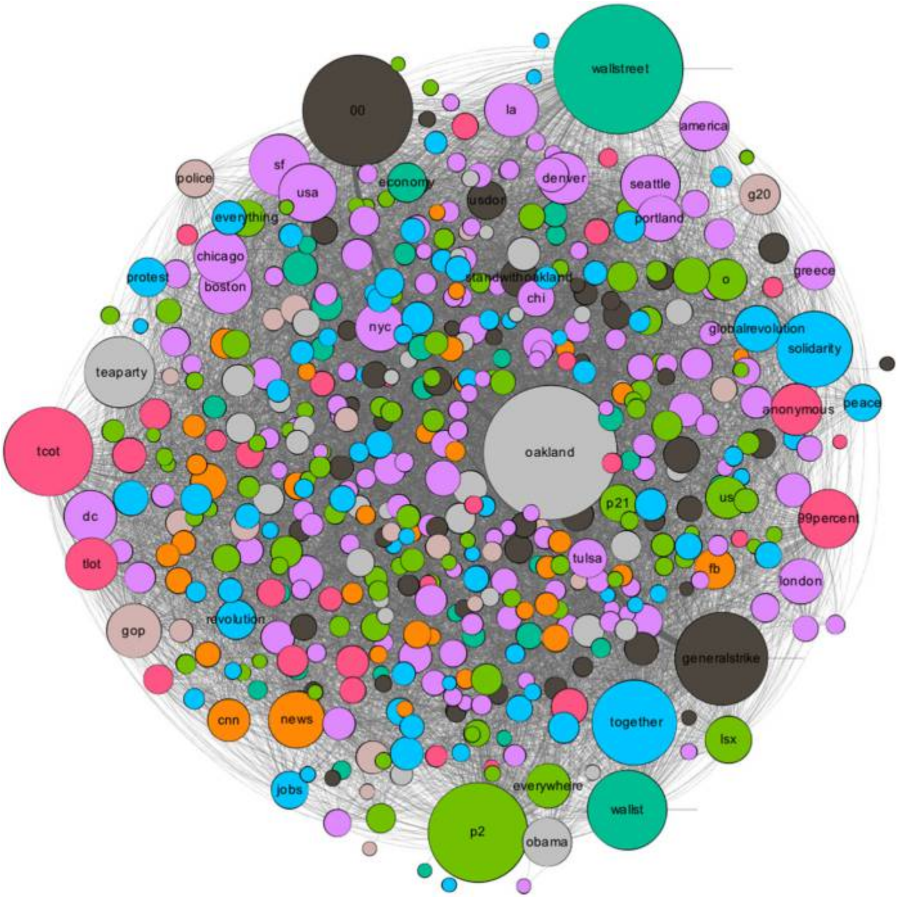
**Table II.** Summary of the network structure on both days

| Network matrices | November 2 | November 18 |
|---|---|---|
| Network size | 500 | 479 |
| Density | 0.08 | 0.04 |
| Degree centralization | 0.74 | 0.34 |
| Betweeness centralization | 0.24 | 0.12 |
| Closeness centralization | 0.69 | 0.37 |
| Transitivity | 0.27 | 0.19 |

**Table III.** Summary of the ERGM on top hashtags co-occurrence

| | November 2, 2011 | | November 18, 2011 | |
|---|---|---|---|---|
| Parameter | Estimates | SE | Estimates | SE |
| Edges | −1.28*** | 0.003 | −4.34*** | 0.06 |
| Frequency | −0.00 | 0.00 | 0.0004*** | 0.00003 |
| Type homophily | 0.09** | 0.03 | −0.01 | 0.05 |
| Theme | −0.44*** | 0.02 | 1.10*** | 0.04 |
| Media outlet | −1.34*** | 0.04 | 0.04 | 0.05 |
| Event | −1.22*** | 0.04 | 0.97*** | 0.04 |
| Identity | −1.70*** | 0.05 | 0.13* | 0.06 |
| Geolocation | −0.63*** | 0.02 | 0.39*** | 0.04 |
| Civil society organization | −0.50*** | 0.05 | 0.82*** | 0.06 |
| Economy | −1.00*** | 0.04 | 0.50*** | 0.07 |
| Public figure | 0.10*** | 0.05 | 0.16* | 0.08 |
| Government related | −1.29*** | 0.05 | 0.43*** | 0.06 |
| Time | −0.08*** | 0.10 | | |
| AIC | 1,69,022 | | 1,57,326 | |
| BIC | 1,69,149 | | 1,57,452 | |

**Notes:** For November 8 model, time parameter showing NA, because of the low number of cases for simulation. $*p < 0.05$; $**p < 0.01$; $***p < 0.001$




**Notes:** Nodes were sized based on degree centrality, color indicates hashtag type, and top 50 nodes with the highest centrality were labeled. Pink color indicates geolocation, light blue indicates theme, black indicates event, orange indicates media outlet, red indicates identity claim, green indicates economy, and beige indicates government organizations. All the other smaller categories of hashtags were coded in gray and the unidentifiable hashtags (the "other" category) were in dark blue

**Figure 1.** Visualization of November 2 hashtag co-occurrence network

degree centralization of 0.34. This network has the betweenness centralization of 0.12, with the closeness centralization of 0.37, and transitivity of 0.19. The network was less sparse and less centralized, with a lower transitivity (Figure 2). Top 50 nodes were labeled, whose degree centrality scores ranged from 40 to 180 which were significantly lower than November 2.

ERGM on the November 2 data set showed that a hashtag's frequency did not significantly influence its chance of getting paired. This suggested that popular hashtags were more likely to be used alone. Meanwhile, hashtags of the same type were more likely to be included in one tweet (estimate = 0.09, SE = 0.03). The results also suggested that theme hashtags were more likely to be used alone, and with hashtags of the same type. The same effect was found on media outlet, event, identity, geolocation, civil society

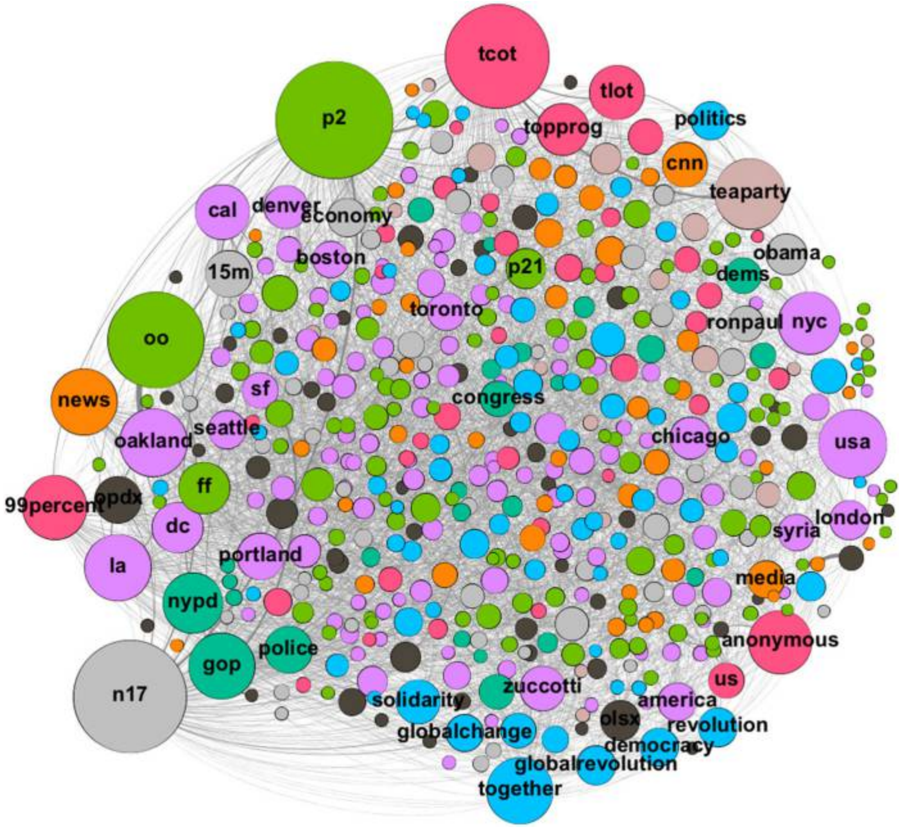



**Figure 2.** Visualization of November 18 hashtag co-occurrence network

**Notes:** Nodes were sized based on degree centrality and top 50 nodes with the highest centrality were labeled. Color coding followed Figure 1

organizations, economy, public figure, government, and time hashtags. On the other hand, public figure hashtags were found to be more likely to co-occur with other types of hashtags, indicating their higher structural virality in driving the network.

November 18 ERGM results showed quite different findings, suggesting the role of hashtags may vary based on the context of the tweets. First, frequency became a significant predictor. The more frequently a hashtag has been used, the more likely it would be used in combination with others. Second, the homophily effect of hashtag type was not significant, which differs from November 2. Third, theme hashtags were found to be more likely to co-occur with other hashtags. The same effect was found on event, identity, geolocation, civil society organization, economy, public figure, and government-related hashtags. It indicates that all these hashtag categories are more likely to gain structural virality on November 18. The use of media outlet hashtag had no significant effect.

## Discussion

To compare the ERGM findings from selected two days, the results showed that the mechanisms driving the hashtag co-occurrence network were distinct. As Table III showed, hashtag frequency had little effect in the November 2 network, while it was a significant predictor of tie formation in the November 18 network. It indicates viral hashtags were more likely to be used in combination with other hashtags when there was an outbreak of events. The homophily effect was significant only in the November 2 network.

Furthermore, the extent to which different types of hashtags were preferred on those two days also differs substantially. Among the top viral hashtags, theme hashtags were more likely to be used alone in the November 2 network. However, during the day of the Pepper Spray event, theme became a popular co-occurring hashtag category. Almost all the other types of hashtags were less likely to be used with other hashtags that day, except public figure hashtags. This indicated that on the November 2 network, the structure of hashtag co-occurrence was mainly driven by homophonous hashtags. To the contrary, the November 18 network was likely to be driven by a more strategic use of hashtag patterns, such as placing different categories of hashtags in one tweet to draw more public attention. Different types of hashtags often help to attract attention from clustered Twitter users. It is through the strategic combination of heterogeneous hashtags that OWS participants were able to distribute information and anger about the Pepper Spray incident to the world.

The effect of media outlet hashtags was negative and significant on November 2. This could indicate that on November 2, top media outlet hashtags were used to report news stories, or simply forward information from new channels. On November 18, media outlet hashtags were not significantly preferred to pair with other hashtags. This could be explained by the fact that the during the Pepper Spray day, more attention was placed on the uncoordinated event that occurred at UC Davis, and Twitter users were more focused on how to better utilize hashtag use patterns to attract more supporters. This difference suggests that on November 18, there was a tendency that self-motivated OWS participants were setting the news agenda by broadcasting and distributing news related to the Pepper Spray incident.

One effect that was consistent was from the public figure hashtags, which were more likely to be used in combination with other hashtags. Given the significant homophily effect on November 2, this could mean that public figure hashtags were more likely to be used with other public figures; while they were more likely to be used with other types of hashtags, such as geolocation or theme hashtags, in the November 18 network. The consistent positive effect from public figure hashtags indicated that Twitter users are strategic in mobilizing public attention, evidenced by how they actively utilized celebrities' social influence to promote their collective agenda.

Another finding worth discussing is the positive effect of identity claim during the UC Davis Pepper Spray event. Identity hashtags tend to be paired up with other hashtags. One explanation is that the outbreak of provocative events may trigger stronger public sentiment, which likely reinforced movement participants' collective identity (Fominaya, 2010). As Taylor and Whittier (1992) argued that collective identity of social movement was often formed in opposition to dominant cultural practices or existing status quo, antagonist events in general and the protest against police brutality in our context, serve to reinforce solidarity and collective identity among movement participants.

The findings suggest that social movement supporters are likely to employ certain strategies to utilize social media to gain public attention toward their collective cause. The significant effect of hashtag types on network formation supports the argument that information characteristics can explain the bottom-up mechanism of achieving virality. The strategic use of hashtags helps self-organized movement supporters to reach other people of the same interest in a timely manner and through different social circles. Self-motivated supporters became the bottom-up social forces in making the movement more visible to the public, ultimately facilitating the achievement of symbolic power. Linking back to the affordance approach about generative role taking, the strategic use of hashtags allows movement participants to take on autonomous roles in constructing a timely information sharing network, so existing and potential contributors can be informed of shared vision and goals.

The different roles hashtag categories played in a networked social movement can also be linked to the context of Twitter use. An unexpected event may shift the strategy of social media use. Distinct hashtag co-occurrence patterns imply that movement communication is effective in facilitating supporters adjust their personalized expression patterns to better address public attention and achieve virality for connective action. By pairing different types of hashtags into their discourse on Twitter, OWS supporters intentionally directed their message to reach a more heterogeneous population in a timely manner.

To look at the general context of networked social movement, there are several important implications we can draw from the current research. First, to succeed in mobilizing public attention, movement participants need to construct effective information flow through strategic combinations of hashtags. Second, the role of viral hashtags in influencing the evolution of a networked social movement varies depending on the event dynamics, suggesting that participants should be aware of the contextual nature of hashtags. Third, the strategic use of viral hashtags indicates the level of autonomy self-motivated movement participants may possess.

## Conclusion

This study examines how different types of hashtags emerged to become viral during the course of a networked social movement, and it further compares how the co-occurrence patterns of viral hashtags differed across two different time points, an ordinary day during the movement vs one with an unexpected public event. By identifying important types of viral hashtags that emerged from both days, the findings shed lights on how hashtags, an important technological feature of Twitter, can be effectively leveraged to help social movement gain visibility and symbolic power.

This paper makes two major contributions. First, the current study conceptualizes hashtag as a mechanism for mobilizing public attention. Through the lens of the virality framework and the network perspective, it uncovered how information characteristics of hashtags could explain the structure of the hashtag co-occurrence network. To the best of our knowledge, few studies have systematically examined the ways in which hashtags are strategically combined to organize collective action. This study fills the gap by empirically testing the bottom-up mechanism of explaining hashtag virality in the context of a networked social movement. Second, through a comparison analysis, this research demonstrates that depending on different event dynamics, certain types of hashtags may be more popular among movement participants. Furthermore, roles of viral hashtags may change over time. It shows that movement dynamics may interact with information characteristics of hashtags to achieve virality.

There are several limitations in the current study. The first deals with the operationalization of virality. In Nahon and Hemsley (2013), virality is essentially defined by three criteria, viral speed, the speed at which a piece of information travels; viral reach by numbers, the size of usage or the total number of users adopting this piece of information; and finally, viral reach by networks, the scope of the user networks (e.g. how heterogeneous, or distant these network clusters are) who adopt information. In the current study, virality was measured in two dimensions: daily frequency of use, and structural centrality in the hashtag co-occurrence network. Future research can take into account the temporal nature and the user composition aspect of virality to capture multiple dimensions, and examine whether there are diverging, or converging sets of hashtags from each viral dimension.

The second limitation lies in the selection of tweets, which may not be representative of the entire movement days. Given the complex event dynamics of OWS, two days' tweets can only provide a limited window of the entire event. However, one motivation of this current research is that there is a bursty nature to certain kinds of information flow across networks. This suggests that analyzing virality should be conducted in combination with the characteristics of the contextual environment. In the OWS movement, selecting two distinct movement days thus opened up an opportunity for us to explore such analysis. Another limitation is that we only focused on top 5 percent of the hashtags to construct the information-sharing network. This filtering method may have affected the representation of OWS hashtag co-occurrence networks. However, this allows us for a closer examination of viral hashtags, defined by frequency of use and degree centrality.

To further capture a more comprehensive picture of how event dynamics, information characteristics are related to information flow during a networked social movement, future study will apply other methodologies such as online experiment, and survey. It will also apply certain computational social science analytics (such as natural language processing and latent semantic analysis) to analyze full-scale text data in a networked social movement. We also call for internet researchers to replicate and test the specific types of viral hashtags identified in this current study, which are by no means exhaustive. By comparing and contrasting a multitude of hashtag co-occurrence networks across social events, one may discover more systematic patterns regarding hashtag use.

Last but not the least, the virality research can combine the hashtags content analysis and Twitter user network analysis. By simultaneously modeling both user and information characteristics, it may uncover more nuanced dynamics. For instance, one may find that certain types of hashtags only gain virality within one group of users, whereas other types of hashtags tend to gain virality universally. Different circles in a network movement may have different social groups as intended audiences. Future studies should examine both content and users networks.

In sum, this research applies the virality framework and a network approach to examine use patterns of hashtags. With the analysis of Twitter data from two comparable days during the OWS movement, this study provides a preliminary test of how information characteristics of a hashtag may influence its use during a networked social movement. It demonstrates that self-organized movement participants have used strategies to leverage social media to better diffuse their message and drive the movement viral.

## References

Agarwal, S.D., Bennett, W.L., Johnson, C.N. and Walker, S. (2014), "A model of crowd enabled organization: theory and methods for understanding the role of Twitter in the occupy protests", *International Journal of Communication*, Vol. 8, pp. 646-672, available at: http://ijoc.org/index.php/ijoc/article/view/2068/1091

Allagui, I. and Kuebler, J. (2011), "The Arab Spring and the role of ICTs: introduction", *International Journal of Communication*, Vol. 5, pp. 1435-1442.

Andrews, K.T. and Caren, N. (2010), "Making the news movement organizations, media attention, and the public agenda", *American Sociological Review*, Vol. 75 No. 6, pp. 841-866.

Bennett, W.L. and Segerberg, A. (2012), "The logic of connective action: digital media and the personalization of contentious politics", *Information, Communication & Society*, Vol. 15 No. 5, pp. 739-768.

Bennett, W.L. and Segerberg, A. (2013), *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*, Cambridge University Press, Cambridge.

Bimber, B., Flanagin, A.J. and Stohl, C. (2012), *Collective Action in Organizations: Interaction and Engagement in an Era of Technological Change*, Cambridge University Press, Chicago, IL.

boyd, D., Golder, S. and Lotan, G. (2010), "Tweet, tweet, retweet: conversational aspects of retweeting on Twitter", *System Sciences (HICSS) 2010 Proceedings of 43rd Hawaii International Conference on System Sciences*, pp. 1-10.

Bruns, A. and Burgess, J.E. (2011), "The use of Twitter hashtags in the formation of ad hoc publics", *Proceedings of 6th European Consortium for Political Research General Conference, University of Iceland*, *Reykjavik*, August 24-27, available at: http://eprints.qut.edu.au/46515/1/The_Use_of_Twitter_Hashtags_in_the_Formation_of_Ad_Hoc_Publics_%28final%29.pdf

Castells, M. (2007), "Communication, power and counter-power in the network society", *International Journal of Communication*, Vol. 1, pp. 238-266.

Castells, M. (2012), *Networks of Outrage and Hope: Social Movements in the Internet Age*, Polity, Cambridge.

Chase-Dunn, C. and Curran-Strange, M. (2012), "The diffusion of the Occupy Movement in California", IROWS Working Paper No. 74, Institute for Research on World-Systems, University of California-Riverside, Riverside, CA.

Costanza-Chock, S. (2012), "Mic check! Media cultures and the Occupy Movement", *Social Movement Studies*, Vol. 11 Nos 3-4, pp. 375-385.

Ding, Y., Chowdhury, G. and Foo, S. (1999), "Mapping the intellectual structure of information retrieval studies: an author co-citation analysis, 1987-1997", *Journal of Information Science*, Vol. 25 No. 1, pp. 67-78.

Earl, J. and Kimport, K. (2011), *Digitally Enabled Social Change: Activism in the Internet Age*, MIT Press, Cambridge, MA.

Fominaya, C.F. (2010), "Collective identity in social movements: central concepts and debates", *Sociology Compass*, Vol. 4 No. 6, pp. 393-404.

Gaby, S. and Caren, N. (2012), "Occupy Online: how cute old men and Malcolm X recruited 400,000 US users to OWS on Facebook", *Social Movement Studies*, Vol. 11 Nos 3-4, pp. 367-374.

Gamson, W.A. and Sifry, M.L. (2013), "The#Occupy Movement: an introduction", *The Sociological Quarterly*, Vol. 54 No. 2, pp. 159-163.

Gerbaudo, P. (2015), "Protest avatars as memetic signifiers: political profile pictures and the construction of collective identity on social media in the 2011 protest wave", *Information, Communication & Society*, Vol. 18 No. 8, pp. 916-929.

Gerbaudo, P. and Treré, E. (2015), "In search of the 'we' of social media activism: introduction to the special issue on social media and protest identities", *Information, Communication & Society*, Vol. 18 No. 8, pp. 865-871.

Gleason, B. (2013), "#Occupy Wall Street: exploring informal learning about a social movement on Twitter", *American Behavioral Scientist*, Vol. 57 No. 7, pp. 966-982.

Groshek, J. and Al-Rawi, A. (2013), "Public sentiment and critical framing in social media content during the 2012 US presidential campaign", *Social Science Computer Review*, Vol. 31 No. 5, pp. 563-576, doi: 0894439313490401.

Hanneman, R.A. and Riddle, M. (2006), *Introduction to Social Network Methods*, University of California, Riverside, CA.

Huang, J., Thornton, K.M. and Efthimiadis, E.N. (2010), "Conversational tagging in Twitter", *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia ACM*, pp. 173-178.

Jenkins, J.C. (1983), "Resource mobilization theory and the study of social movements", *Annual Review of Sociology*, Vol. 9, pp. 527-553.

Juris, J.S. (2004), "Networked social movement: global movements for global justice", in Castells, M. (Ed.), *The Network Society: A Cross-Cultural Perspective*, Edward Elgar Publishing, Cheltenham, pp. 341-362.

Kavada, A. (2015), "Creating the collective: social media, the Occupy Movement and its constitution as a collective actor", *Information, Communication & Society*, Vol. 18 No. 8, pp. 872-886.

Lotan, G., Graeff, E., Ananny, M., Gaffney, D. and Pearce, I. (2011), "The Arab Spring| the revolutions were tweeted: information flows during the 2011 Tunisian and Egyptian revolutions", *International Journal of Communication*, Vol. 5 No. 31, pp. 1375-1405.

Lusher, D., Koskinen, J. and Robins, G. (2013), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*, Cambridge University Press, Cambridge.

McCarthy, J.D. and Zald, M.N. (1977), "Resource mobilization and social movements: a partial theory", *American Journal of Sociology*, Vol. 82 No. 6, pp. 1212-1241.

Ma, F. and Li, Y. (2014), "Utilizing social network analysis to study the characteristics and functions of the co-occurrence network of online tags", *Online Information Review*, Vol. 38 No. 2, pp. 232-247.

Ma, Z., Sun, A. and Cong, G. (2013), "On predicting the popularity of newly emerging hashtags in Twitter", *Journal of the American Society for Information Science and Technology*, Vol. 64 No. 7, pp. 1399-1410.

Majchrzak, A. and Markus, L. (2012), *Technology Affordances and Constraint Theory of MIS*, Sage, Thousand Oaks, CA.

Majchrzak, A., Faraj, S., Kane, G.C. and Azad, B. (2013), "The contradictory influence of social media affordances on online communal knowledge sharing", *Journal of Computer-Mediated Communication*, Vol. 19 No. 1, pp. 38-55.

Meraz, S. and Papacharissi, Z. (2013), "Networked gatekeeping and networked framing on #Egypt", *The International Journal of Press/Politics*, Vol. 18 No. 2, pp. 138-166.

Messina, C. (2007), "Groups for Twitter; or a proposal for Twitter tag channels", FactoryCity (blog), available at: http://factoryjoe.com/blog/2007/08/25/groups-for-twitter-or-a-proposal-for-twitter-tag- channels (accessed December 3, 2015).

Morozov, E. (2011), *The Net Delusion: How Not to Liberate the World*, Penguin, UK.

Nahon, K. and Hemsley, J. (2013), *Going Viral*, Polity, Cambridge.

Papacharissi, Z. and de Fatima Oliveira, M. (2012), "Affective news and networked publics: the rhythms of news storytelling on #Egypt", *Journal of Communication*, Vol. 62 No. 2, pp. 266-282.

Robins, G., Pattison, P., Kalish, Y. and Lusher, D. (2007), "An introduction to exponential random graph (p*) models for social networks", *Social Networks*, Vol. 29 No. 2, pp. 173-191.

Romero, D.M., Meeder, B. and Kleinberg, J. (2011), "Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter", *Proceedings of the 20th International Conference on World Wide Web*, pp. 695-704.

Sharma, S. (2013), "Black Twitter? Racial hashtags, networks and contagion", *New Formations: A Journal of Culture/Theory/Politics*, Vol. 78 No. 1, pp. 46-64.

Small, T.A. (2011), "What the hashtag? A content analysis of Canadian politics on Twitter", *Information, Communication & Society*, Vol. 14 No. 6, pp. 872-895.

Snijders, T.A., Pattison, P.E., Robins, G.L. and Handcock, M.S. (2006), "New specifications for exponential random graph models", *Sociological Methodology*, Vol. 36 No. 1, pp. 99-153.

Taylor, V. and Whittier, N. (1992), "Collective identity in social movement communities: lesbian feminist mobilization", in Muller, C.M. and Morris, A.D. (Eds), *Frontiers in Social Movement Theory*, Yale University Press, New Haven, CT, pp. 104-129.

Thompson, J.B. (2005), "The new visibility", *Theory, Culture & Society*, Vol. 22 No. 6, pp. 31-51.

Thorson, K., Ekdale, B., Borah, P., Namkoong, K. and Shah, C. (2010), "YouTube and Proposition 8: a case study in video activism", *Information, Communication & Society*, Vol. 13 No. 3, pp. 325-349.

Thorson, K., Driscoll, K., Ekdale, B., Edgerly, S., Thompson, L.G., Schrock, A., Swartz, L., Vraga, E.K. and Wells, C. (2013), "YouTube, Twitter and the Occupy Movement: connecting content and circulation practices", *Information, Communication & Society*, Vol. 16 No. 3, pp. 421-451.

Treem, J.W. and Leonardi, P.M. (2013), "Social media use in organizations: exploring the affordances of visibility, editability, persistence, and association", *Annals of the International Communication Association*, Vol. 36 No. 1, pp. 143-189.

Tufekci, Z. (2013), " 'Not This One': social movements, the attention economy, and microcelebrity networked activism", *American Behavioral Scientist*, Vol. 57 No. 7, pp. 848-870.

White, H.D. and Griffith, B.C. (1981), "Author cocitation: a literature measure of intellectual structure", *Journal of the American Society for Information Science*, Vol. 32 No. 3, pp. 163-171.

Zamir, M.H. (2014), "Diffusion of protest information in Twitter during Shahbag movement of Bangladesh", *Proceedings of the Association for Information Science and Technology Annual Meeting, Vol. 51, Seattle, WA, October 31-November 5*, pp. 1-4.

**About the authors**

Rong Wang (PhD, University of Southern California, 2016) is a Postdoc Research Associate at the Department of Communication Studies, Northwestern University. Her research focuses on analyzing how to organize different networks for public goods. Rong Wang is the corresponding author and can be contacted at: rongwang@northwestern.edu

Wenlin Liu (PhD, University of Southern California, 2016) is an Assistant Professor of Organizational Communication at the University of Houston-Downtown. Her research interests include non-profit strategic communication, organizational networks and community building.

Shuyang Gao is a Doctoral Student at the Department of Computer Science, University of Southern California. His research focuses on information theory and machine learning.