Article

The International Journal of Press/Politics
2015, Vol. 20(2) 166–184

DOI: 10.1177/1940161214566709
ijpp.sagepub.com

Hero on Twitter, Traitor on News: How Social Media and Legacy News Frame Snowden

Jie Qin[1,2]

Abstract

Is Edward Snowden a hero or traitor? In what ways do frames on social media and legacy news differ in covering the incident of Edward Snowden? Utilizing the approach of semantic network analysis, the study found social media users associated Snowden's case with other whistle-blowers, bipartisan issues, and personal privacy issues, while professional journalists associated the Snowden incident with issues of national security and international relations. Frames on social media portray Edward Snowden as a hero while the frames on legacy news make him a traitor. The study further identified media frames on social media and legacy news differ in two ways: word selection and word salience. In addition, the study discussed the challenges and opportunities of framing analysis in the context of social media.

Keywords

framing theory, semantic network, social media, legacy news, public opinion

While the debates on whether Julian Assange and Bradley Manning are heroes or villains still linger, another whistle-blower arrived. Edward Snowden, the former employer of National Security Agency (NSA) in the United States, hid in Hong Kong, leaking and accusing the NSA of massive surveillance. As usual, a global debate titled "Edward Snowden: a hero or traitor" was on. President Obama called it "modest encroachments on privacy" (Dorning and Strohm 2013), while Ron Paul, the Representative for Texas

[1]City University of Hong Kong, Hong Kong
[2]SCAD Hong Kong, Hong Kong

Corresponding Author:
Jie Qin, Professor of Communications, Department of Liberal Arts, SCAD Hong Kong, 292 Tai Po Road, Sham Shui Po, Kowloon, Hong Kong.
Email: jqin@scad.edu

and former presidential candidate, said, "We should be thankful for individuals like Edward Snowden and Glenn Greenwald who see injustice being carried out by their own government and speak out, despite the risk" (Walsh 2013).

Which camp will prevail, Snowden's attackers or defenders? According to Kirn (2013), the winner is the latter, "at least on Twitter," as Twitter is different from newspapers or magazines, "where muddiness flourishes and space is at less of a premium." His assumption does not only concern who is the winner, more importantly, it triggers an important, yet rarely studied research question: In what ways do frames on social media and legacy news differ? To address this question, this article is organized as four parts. The first part reviews the classic conceptualization, mechanism, and operationalization of framing analysis, and discusses the challenges toward the three issues in the context of social media. The second part introduces hashtags as new framing devices and semantic network analysis as an alternative to traditional framing analysis in the context of social media. The third part begins with a brief overview of the Snowden incident, and then offers a semantic network analysis and compares in what ways frames on social media and legacy news differ in covering the Snowden incident. The fourth part concludes and discusses new opportunities for framing analysis in the context of social media.

Rethinking the Conceptualization, Mechanism, and Operationalization of Framing Analysis

Framing Analysis in the Context of Legacy News

Exhaustive discussions have been made on the origins of framing theory. Pan and Kosicki (1993) suggested two traditions: sociological tradition and psychological tradition of framing theory. The sociological tradition started with Goffman (1974), and the psychological tradition has a longer history, which can be traced back to scholars like Piaget (1929) in early twentieth century. Goffman defined frame as a "strip" or "any arbitrary slice or cut from the stream of ongoing actives" (Goffman 1974: 10). According to Piaget, frames are interchangeable with schemas. A schema outlines an object or a construct, such as cloud, life, and so on, and will "be taken as the general" (Piaget 1929: 204).

Regarding the conceptualization of frames in communication research, Scheufele (1999) made a distinction between two types of frames: media frames and individual frames. A media frame is "a set of interpretive packages that give meaning to an issue" (Gamson and Modigliani 1989: 3), while individual frames "are internal structures of the mind that help individuals to order and give meaning to the dizzying parade of events" (Kinder and Sanders 1990: 74). For example, Gamson and Modigliani's (1989) study on media coverage of nuclear issues is a classic example of media frame analysis. The authors examined five types of framing devices in daily newspapers and network evening news broadcasts, including metaphors, exemplars, catchphrases, depiction, and visual images to identify different interpretive packages from 1940s to 1980s. Take the interpretive package in 1950s for example. During that time, the progress package was identified as the dominant interpretive package. An editorial in *New York Times* was quoted as an example of the progress package: "We face the prospect either of destruction on a scale which dwarfs anything thus far reported, or of a golden era of social change which would satisfy the most romantic utopian" (Gamson and Modigliani 1989: 12).

Regarding the mechanism of framing process, it involves two elements: selection and salience (Bakker and Hellsten 2013; Scheufele 1999). As Entman (1993) elaborated,

> To frame is to select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation for the item described. (Entman 1993: 52)

For example, Gamson and Modigliani (1989) found that the dominant frames that media agencies employed to cover nuclear issues kept changing from 1940s to 1980s. Media agencies had selected some aspects of nuclear issues and made them more salient. Meanwhile, different media agencies had selected different frames regarding this controversial issue, which implied "a range of position" of the media agencies, "rather than any single one, allowing for a degree of controversy among those who share a common frame" (Gamson and Modigliani 1989: 3).

Regarding the operationalization of framing analysis, let us go back to Gamson and Modigliani's (1989) study again. The first step was to collect a considerable amount of texts on a given issue. And then, the authors developed several "meta-frames" and "sub-frames" as the coding schemes based on researchers' subjective or even arbitrary selections. At the end, the authors recruited coders to go over all the texts and put each article into corresponding sub-frame. When this top–down approach is applied in the context of social media, it becomes problematic, which will be elaborated in the next section.

Challenges for Framing Analysis in the Context of Social Media

Regarding the conceptualization of frames, a major challenge is whether the frames in social media are media frames or individual frames in nature, given the fact that social media are a mixture of institutional accounts and individual accounts? Take Twitter for example. Wu et al. (2011) found that Twitter users can be classed into at least five types: celebrities, media agencies, other organizational accounts, bloggers, and ordinary individual accounts. So the social media frames in nature are not simply the interpretive packages that media agencies employ, but also are far more complex than individual schemas. However, existing dualistic conceptualization, media frames and individual frames, fail to capture the nature of social media frames. To address this issue, I employed a network perspective to conceptualize social media frames. In this article, social media frames are conceptualized as the networks that are composed of various framing devices, such as hashtags, keywords, and so on. Compared with existing dualistic conceptualization, my conceptualization is innovative in two aspects.

First, it involves a bottom–up perspective, in contrast with the top–down perspective in existing dualistic conceptualization. Second, it sheds light on the connections between different framing devices, which has been overlooked in extant studies.

Regarding the mechanism of framing process, the challenge is how to measure the process of selection and salience, given the massive amounts of contents on social media, which are beyond the processing capacity of traditional content analysis? In this article, I proposed two new measures based on the network perspective. Selection refers to the framing devices that are included in a network, and salience is measured by the degrees of centrality of each framing device in the network. More details on the two measures will be elaborated in analyzing the incident of Edward Snowden in the following part.

Regarding the operationalization of framing analysis, the challenges are twofold. First, the framing devices identified in legacy news are either absent or transformed on social media. Tankard (2001: 101) offered a list of framing devices, which has been widely used in extant framing research. The list contains eleven common framing devices on legacy news: (1) headlines and kickers; (2) subheads; (3) photographs; (4) photo captions; (5) leads; (6) selection of sources or affiliations; (7) selection of quotes; (8) pull quotes; (9) logos; (10) statistics, charts, and graphs; and (11) concluding statements or paragraphs of articles. However, on social media, take Twitter for example, headlines, leads, and news sources are absent while tweets, retweets, replies, hashtags, attached hyperlinks, photos, and videos are what we have. To analyze social media frames, we need to identify appropriate framing devices. Hashtags, which have drawn considerable media attention (e.g., *ABC News* 2010) and preliminary academic attention (e.g., Hemphill et al. 2013; Parmelee 2013) will be introduced as the framing device on social media in the next section.

The second challenge is the scholar-centered approach. Unlike news articles which bear rich contextual cues for framing analysis, tweets are usually too short to carry sufficient contextual cues. For example, in Sukosd and Fu's (2013) study, the authors aimed to explore how netizens discussed seven major environmental conflicts in China on Weibo. The authors borrowed five media frames from Semetko and Valkenburg's (2000) study, along with seven self-selected frames as the benchmarks. Coders were allowed to choose the one and only one proximal frame for each tweet. The authors were at risk of over-interpreting a considerable amount of tweets. For example, the decision of putting "Concerned!" into the "what's next frame" is questionable, since it could not be wrong to put it into any other frames. As Tankard strongly criticized,

> This approach makes frame identification a rather subjective process. Does one reader saying a story is using a conflict frame make that really the case? Indeed, coming up with the names for frames itself involves a kind of framing. (Tankard 2001: 98)

New Method for Framing Analysis on Social Media

Hashtags as Framing Devices on Social Media

Hashtags are words or phrases with the hash symbol "#." For example, "#wikileaks" is a hashtag to highlight this tweet is about WikiLeaks. I offered a detailed explanation of the most visible hashtags in the semantic network of Snowden incident on Twitter in Table 3. In practice, Twitter users are encouraged to add hashtags in their tweets to increase the visibility in Twitter Search (Twitter 2014). From an academic perspective, hashtags "are both text and metatext, information and tag, pragmatic and metapragmatic speech" (Rambukkana 2013: 1). Specific hashtags such as "#qldfloods" (Bruns et al. 2012) and "#Egypt" (Meraz and Papacharissi 2013; Papacharissi and de Fatima Oliveira 2012) have been utilized to identify event-related discussions on social media. However, studies regarding hashtags as framing devices are just emerging.

Hemphill et al.'s (2013) research is among this emerging body of studies regarding Twitter hashtags as framing devices. The authors performed an algorithmic approach to detect how politicians use hashtags to frame what issues and identified forty topics based on 10,546 hashtags. For example, users added hashtags like "#ACA" and "#Obamacare" in their tweets when they talked about health care issues. Hashtags like "#JOBS" and "#4jobs" appeared in the tweets discussing employment issues. An interesting finding is that Republicans and Democrats had different preference toward hashtags. Republicans prefer hashtags concerning macro-level issues while Democrats prefer micro-level issues, as the authors quoted the Republican National Committee's saying, "while Democrats tend to talk about people, Republicans tend to talk about policy" (Hemphill et al. 2013: 15).

Hemphill et al.'s study touched an important character of hashtags, that is, multiple hashtags may emerge regarding a complex issue. However, they just stopped at identifying partisan preferences on hashtags and did not go further into the connections between various hashtags on the same issues. In fact, when Koenig (2004) developed the idea of routinizing frame analysis with computer-aided tools, he mentioned two key steps. First, we need to identify the keywords in the texts. Second, we need to build up a word net based on the co-occurrence of these keywords to map the connections among them. The word net is actually a semantic network that will be discussed next.

Semantic Network Analysis

Semantic network analysis is a technique to map the associations among concepts. In the network, the nodes are words or phrases, and the edges are the co-occurrences or various associations among the nodes. For example, Doerfel and Barnett (1999) utilized this technique to study the structure of International Communication Association (ICA). They extracted words from titles of the papers presented to ICA to draw the semantic networks. In addition, they built up an affiliation network based on ICA memberships. They found the two networks were significantly correlated, which suggested scholars in the same division spoke the same language. Another up-to-date project of Yuan et al. (2013) used semantic network analysis to capture the privacy issue on Chinese social media. They found that the full semantic network was constituted of eleven concept clusters, which yielded diverse interpretations of privacy issues in China.

As a bottom–up approach, the foremost merit of semantic network analysis is it lets frames emerge by themselves. This is what greatly differentiates semantic network analysis from the top–down scholar-centered approach. In addition, it also has many advantages. First, the processing capacity leaps from kilobyte level to terabyte level. According to Daly (2009), the smallest ePub book is 1.6 kilobytes, the largest is 233 megabytes, and the total size of 35,854 books is 20 gigabytes. The largest dataset by far is located in NSA Data Center, which is reported to have 5 zettabytes of data (Herridge 2013). This is far beyond the processing capacity of traditional content analysis, but a normal size for semantic network analysis. Second, semantic network analysis expands the research scope from single frame to the associations among multiple frames. In traditional framing analysis, the definition of event boundary is quite arbitrary, and associations among frames are absent due to limited processing capacity. What a semantic network reveals is a natural situation of public opinion, where events connect with events by associations. Third, it brings down the costs and improves efficiency. Content analysis is labor-intensive; recruiting coders and coding processes take time and money. On the contrary, computed-aided semantic network analysis is a technology-intensive task, requiring computers and software, which I believe is affordable (or even zero-cost) for researchers.

Research Questions and Hypotheses

This section begins with a very brief review of the incident of Edward Snowden, a 30-year-old American citizen, who revealed the classified documents on PRISM, a surveillance program of NSA in the United States. It is reported that Snowden flew to Hong Kong in May 2013, and contacted media outlets including the *Guardian*, the *Washington Post*, and Hong Kong local press *South China Morning Post*. On June 7th, Snowden's story went public. Before Snowden left Hong Kong for Russia on June 23rd, Snowden leaked more and more documents that the U.S. government had the least interest of seeing in print. The world got to know that the U.S. government not only collected phone calls and online records of its citizens, but also intercepted foreign embassies and hacked the backbone networks in Hong Kong and mainland China. Public opinions on the incident were divided. In the United States, the White House felt "extremely disappointed" (Associated Press 2013b), but a petition on www.Whitehouse.gov had attracted more than 100,000 signatures to support Snowden, who, as the petition put it, was a "national hero" (Associated Press 2013a).

How do legacy news frame whistle-blowers? According to Wahl-Jorgensen and Hunt (2012: 399–407), legacy news in the United Kingdom "mostly cover whistle-blowers in neutral or positive ways," since "within the UK national newspaper cultures, blowing the whistle on corruption and malpractice is constructed as a brave act in the public interest," consistent with British's supportive public opinion on whistle-blowers. In the United States, Gallup polls found 53 percent Americans disapproved of the government surveillance programs (Newport 2013), while legacy news seemed to be taking an opposite direction. According to Grey (2013), "the American media has lined up squarely behind the Obama administration, the NSA and the military in defense of the massive spying operations exposed by former NSA contractor Edward Snowden."

Why do the legacy news in the United States take the position against public opinion and in chorus with the Governments? In fact, as Zhang (2013) recalled Daniel Ellsberg's releasing of the Pentagon Papers to *New York Times* in 1971, legacy news praised Ellsberg as a hero ending the Vietnam War. However, as Hewitt and Lucas (2009) observed, the ways that legacy news cover intelligence issues have been shifting. After the September 11 attacks, "national security" or "the War on Terror" became the dominant frame in media coverage of intelligence issues (Hewitt and Lucas 2009: 107). So, it is rational to assume that legacy news in the United States will cover Snowden incident with the frame of "national security" or "the War on Terror," which portrays Snowden a traitor.

Little research has been done on how social media frame whistle-blowers and the difference between social media and legacy news in covering the same issue, although it has been widely accepted that legacy news and social media are two different institutions in terms of actors, logics, routines, structure, and so on (Bennett and Segerberg 2012; Dijck and Poell 2013; Dutton 2009). To fill the gaps, the present study aims to address the following research question:

> **Research Question 1:** In what ways do frames on social media and legacy news differ in covering the incident of Edward Snowden?

Zhang (2013) has made a good summary on four frames on the incident of Edward Snowden: the frame of employee loyalty, the frame of freedom of speech, the frame of international relations, and the frame of citizen privacy. In the frame of employee loyalty, Edward Snowden is a disloyal employee. As his employer Booz Allen Hamilton (2013) put in its statement: "this action represents a grave violation of the code of conduct and core values of our firm." In the frame of freedom of speech, Snowden has the freedom to express himself. In the frame of international relations, Snowden turns out to be a traitor who brings disgrace on his home country. In the frame of citizen privacy, Snowden is a hero who sacrifices himself to protect the people all over the world.

As discussed in previous section, the central mechanism of framing process includes two elements: selection and salience. If we regard the above four frames as four networks, the framing devices such as keywords that are involved in each network and the most salient keyword in each network are very likely to be different. Here, I propose two hypotheses:

> **Hypothesis 1 (H1):** Regarding the selection process, frames on social media and legacy news differ in selecting framing devices.
> **Hypothesis 2 (H2):** Regarding the salience process, frames on social media and legacy news differ in making certain framing devices more salient than other framing devices.

Table 1. Summary of Data Collection.

To Get the Semantic Network	Tool	Sample	Approach
On Twitter	http://hashtagify.me/	1% sample offered by Twitter Streaming API	Snowball sampling
On Legacy News	http://www.sensebot.net/	Samples collected by SenseBot	Snowball sampling

Research Method

Data Collection

How can the researcher build up a semantic network based on the texts that are collected? The general procedure includes three steps: text segmentation, identifying word co-occurrence, and retrieving the semantic network. Text segmentation is the process of cutting passages and sentences into smaller units such as phases and words. Co-occurrence refers to that two words appear in one sentence. A network is composed of two elements: nodes and links. The nodes in a semantic network are the words yielded from the process of text segmentation. The links indicate the connected two words co-occur in one sentence.

In this study, multiple tools and approaches were employed as reported in Table 1. The semantic network of Snowden incident on Twitter was retrieved from http://hashtagify.me/ in June 2013. It is a Twitter hashtag search engine. The database of this search engine is based on the 1 percent sample from Twitter Streaming Application Programming Interface (API) (CyBranding 2013). It offers the semantic network for each hashtag as illustrated in Figure 1. In other words, the tool has completed the first two steps in the general procedure as mentioned before. The third step—to retrieve the semantic network—needs to be done manually. In the study, snowball sampling was used to retrieve the semantic network.

The semantic network of the incident of Edward Snowden on legacy news was retrieved from http://www.sensebot.net/ in June 2013. It is a semantic search engine. The users can choose one from two databases to do the search: SenseBot or Google, and refine the search in news only by ticking "Search news only," as well as refine the search in English/French/German/Spanish. For example, I searched "Snowden" using the database of SenseBot, ticking "Search news only" and selecting "English" as the language of the query. Then, I got many words that co-occur with "Snowden" in English news reports. Next, I used snowball sampling to map the whole semantic network. I used the words co-occurring with "Snowden" as seeds and searched each of them. Each seed recalled more words. Then I searched the new words one by one to find more words. I evaluated the face validity of the results after each wave. After three waves, the new words were apparently unrelated with Snowden incident, so I stopped. In addition, I used both of the databases of SenseBot and Google to cross-validate the results.

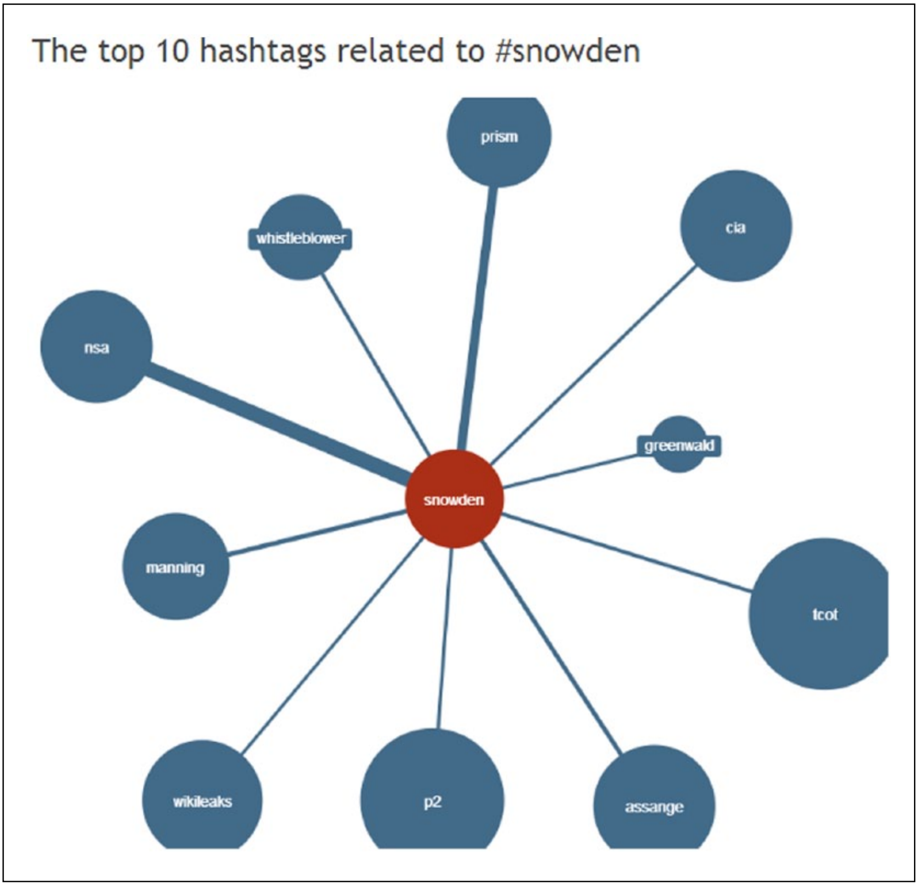

Figure 1. A snapshot of the semantic network of the hashtag "Snowden" generated by http://hashtagify.me/.

Measures

Word selection. H1 argues that frames on social media and legacy news differ in selecting framing devices. In the semantic network of Snowden incident on Twitter, hashtags were employed as framing devices. In the semantic network of Snowden incident on legacy news, keywords were employed as framing devices. Word selection measures the extent to which the hashtags that were involved in the semantic network on Twitter overlap with the keywords that were involved in the semantic network on legacy news.

Word salience. H2 argues that frames on social media and legacy news differ in making certain framing devices more salient than other framing devices. In the semantic network, the degree of centrality of a word suggests the importance of the word in the

Table 2. Descriptive Analysis of Two Semantic Networks.

Semantic network	Number of unique words	Number of ties among words	Network centralization	Top ten words with largest values of out-degree centrality	Top ten words with largest values of in-degree centrality
On Twitter	81	170	2.14	#snowden	#tcot
				#greenwald	#p2
				#nsa	#teaparty
				#p2	#obama
				#assange	#tlot
				#prism	#gop
				#tcot	#ocra
				#wikileaks	#sgp
				#manning	#nsa
				#irs	#wikileaks
On Legacy News	462	699	1.11	NSA	Government
				Hong Kong	Internet
				Intelligence Committee	Facebook House
				Terrorists	President
				Supporters	National Security
				High School	Congress
				Obama	NSA
				Internet	Terrorists
				Cheney	Surveillance
				Washington	

network. In other words, the more central the position is, the more important the word is in the semantic network. In the study, two measures were used to measure word salience: in-degree centrality and out-degree centrality of the framing device.

Results

As reported in Table 2, in the semantic network of Snowden incident on Twitter, there are eighty-one unique hashtags with 170 edges. Based on the top ten words with the largest out-degrees and in-degrees, we can see that when social media users mentioned Edward Snowden on Twitter, they were also talking about other whistle-blowers such as Glenn Greenwald, Julian Assange, and Bradley Manning, as well as the NSA and its PRISM project. Several unexpected words also appear like "#p2," "tcot," and "#irs." Meanings of more hashtags are listed in Table 3. The semantic network of Snowden incident on legacy news contains 462 unique words and 699 edges as reported in Table 2. As mentioned before, I used both of the databases, SenseBot and Google, to cross-validate the results. It turns out the results yielded from the two databases are quite different in terms of recall and precision. The results from Google

Table 3. Meanings of Selected Hashtags in the Semantic Network of Snowden Incident on Twitter.

Hashtag	Meaning
#tcot	Top conservatives on Twitter—a coalition of conservatives on the Internet.
#lnyhbt	Let not your heart be troubled—created by the supporters of Sean Hannity, Fox News Personality, and national Conservative radio talk host.
#gop	Grand Old Party—the U.S. Republican Party.
#sgp	Smart Girl Politics—A Conservative Women's Movement.
#teaparty	Tax protests held nation-wide against the spending for TARP, stimulus, and big-budget government.
#ocra	Organized Conservative Resistance Alliance.
#p2	Progressives On Twitter.
#irs	Internal Revenue Service.
#nsa	NSA.
#prsim	NSA's PRISM program.
#snowden	Edward Snowden.
#assange	Julian Paul Assange, the founder of WikiLeaks.
#wikileaks	It is a website which publishes secret information, news leaks, and classified media from anonymous sources.
#manning	Bradley Edward Manning.

Note. TARP = Troubled Asset Relief Program; NSA = National Security Agency.

contained more noise like "album," since there is a rock band called "Snowden." SenseBot recalled more words and had better precision. So I used the semantic network based on the database of SenseBot to do following comparison with that on Twitter. Google Fusion Tables was used to visualize the networks (see Figure 2).

Hypothesis 1 argues that frames on social media and legacy news differ in selecting framing devices. In the semantic network of Snowden incident on Twitter, hashtags were employed as framing devices. In the semantic network of Snowden incident on legacy news, keywords were employed as framing devices. Word selection measures the extent to which the hashtags that were involved in the semantic network on Twitter overlap with the keywords that were involved in the semantic network on legacy news. It is supported. There are only fifteen words that were found both in the two semantic networks. The fifteen words include Benghazi, CIA, Facebook, Google, GOP, Internet, IRS, NSA, Obama, security, Snowden, spy, surveillance, tax, and whistle-blower. It turns out that frames on social media and legacy news select different framing devices in covering the Snowden incident.

Hypothesis 2 argues that frames on social media and legacy news differ in making certain framing devices more salient than other framing devices. It is supported. As reported in Table 2, the top ten words in the two frames that have the largest in-degrees or out-degrees are quite different. The most important words on Twitter are individual

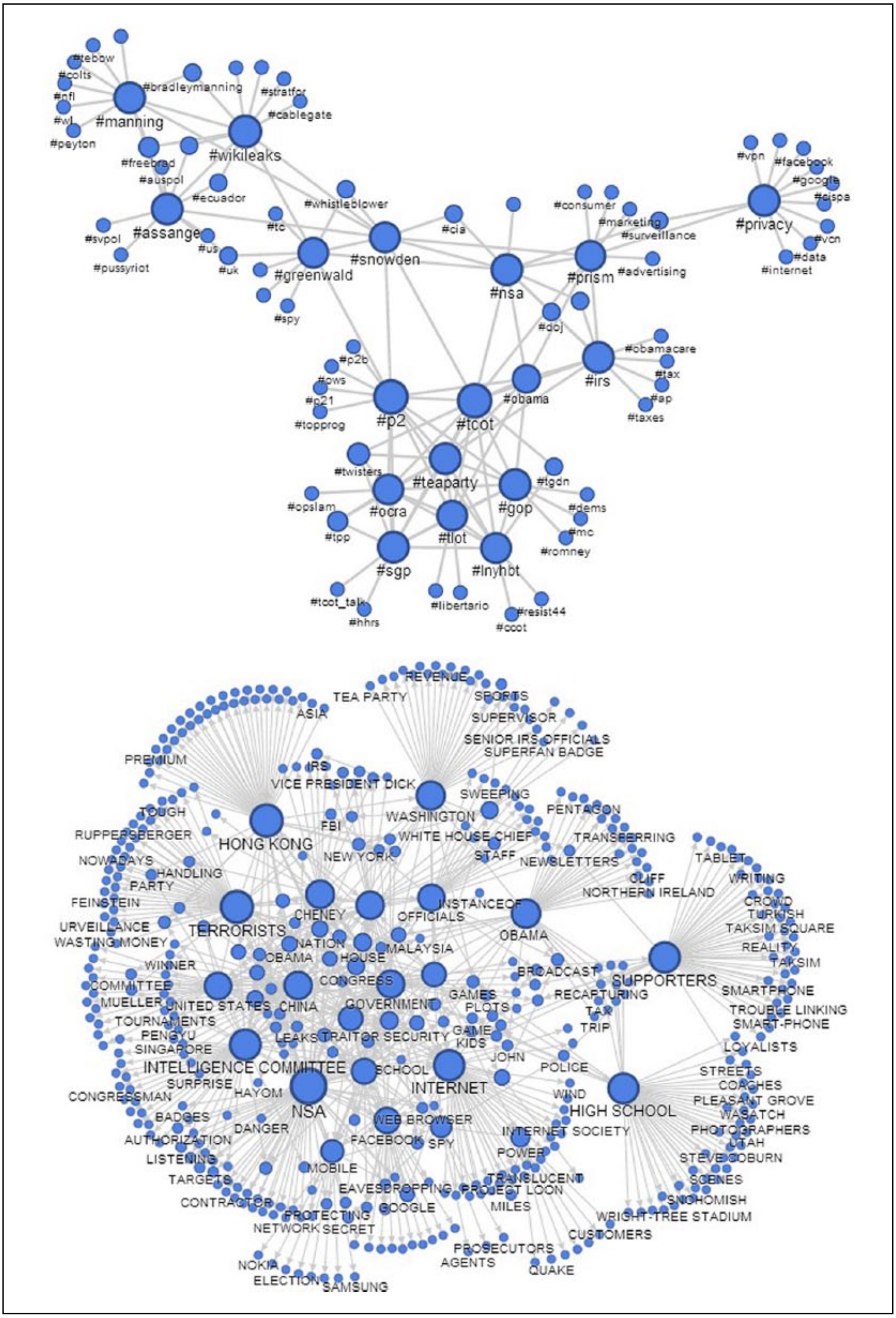

Figure 2. The full semantic networks of Snowden incident on social media (above) and legacy news (below).

names (e.g., #snowden, #greenwald, #manning, etc.) and partisan entities (e.g., #tcot, #p2, #teaparty, etc.). The most important words on legacy news are geographical locations (e.g., Hong Kong, Washington, etc.) and governmental entities (e.g., government, NSA, house, etc.).

Conclusion and Discussion

The study started with Kirn's (2013) interesting and insightful assumption that Edward Snowden would be a hero on Twitter but a traitor on newspapers or magazines, since social media and legacy news are distinct institutions with different logics. The study has proved that social media and legacy news have built different frames on Snowden incident. Frames on social media and legacy news differ in two ways: word selection and word salience. Word selection refers to the different framing devices that are selected in the two networks. Word salience refers to that frames on social media and legacy news differ in making certain framing devices more salient than other framing devices.

Three clusters of hashtags can be identified by visualization in Figure 2. The first cluster is the whistle-blower frame, including hashtags such as "#assange," "#wikileaks," "#greenwald," "#manning," and so on. The second cluster is the bipartisan frame. This cluster is led by two flagship hashtags: "#p2" and "#tcot." The former is the symbol of "Progressives on Twitter," and the latter represents the "Top Conservatives on Twitter." The two flagships associated with issues such as Tea Party movement and supporting organizations such as "#sgp" (Smart Girl Politics). The third cluster is the privacy frame. It includes major hashtags such as "#nsa," "#prism," and "#privacy."

To get a clearer view, twenty most important words were kept in each network (see Figure 3). We can still see the three frames of whistle-blower, bipartisan, and privacy on Twitter. On legacy news, two specific frames can be identified: The national security frame, which includes words like "national security" and "terrorists," and the international relations frame, which includes words like "Washington," "China," and "Hong Kong." It is noteworthy that "High School" is also a keyword on legacy news. In fact, Edward Snowden himself did criticize the U.S. media coverage that "the mainstream media now seems far more interested in what I said when I was 17 or what my girlfriend looks like rather than, say, the largest program of suspicionless surveillance in human history" (Mirkinson 2013). This gives more credits to the validity of this approach.

As a conclusion, social media users associated Snowden's case with other whistleblowers, bipartisan issues, and personal privacy issues. The three frames are independent but loosely connected. On legacy news, which appeared a more unified discourse, professional journalists connected the Snowden incident with issues of national security and international relations. In addition, all the three frames on Twitter are in favor of Snowden, while the frames in news reports make him a traitor.

Next, I would like to discuss the challenges and opportunities of framing analysis in the context of social media. As discussed in previous section, the conceptualization of frames, the mechanism of framing process, and the operationalization of framing analysis need to be reconsidered in the context of social media. Regarding the conceptualization of frames, the challenge is that whether the frames in social media are media frames or individual frames in nature, given the fact that social media are a mixture of institutional accounts and individual accounts? Regarding the mechanism of framing process, the challenge is how to measure the process of selection and salience, given the massive amounts of contents in social media, which are beyond the processing capacity of traditional content analysis? Regarding the operationalization of framing analysis, the challenges are twofold. First, the framing devices identified in legacy news are either absent or transformed in social media. The second challenge is the scholar-centered approach that involves subjective frame identification, given that fact that tweets are usually too short to carry sufficient contextual cues.

As responses to the above challenges, the contributions of the current study are threefold. First, regarding the conceptualization, the study used a network perspective to conceptualize social media frames. Social media frames are conceptualized as the networks that are composed of various framing devices, such as hashtags, keywords, and so on. Compared with existing dualistic conceptualization, the new conceptualization is innovative in two aspects. On one hand, it involves a bottom–up perspective, in contrast with the top–down perspective in existing dualistic conceptualization of media/individual frame. On the other hand, it sheds light on the connections between different framing devices, which has been overlooked in extant studies. Second, regarding the mechanism of framing process, the study proposed two new measures based on the network perspective. Word selection refers to the framing devices that are included in a network, and word salience is measured by the degrees of centrality of each framing device in the network. Third, regarding the operationalization, hashtags were introduced in the study as a new framing device on social media. The study also proposed a bottom–up approach as the alternative to the scholar-centered approach to minimize the risk of over-interpretation.

Meanwhile, new opportunities for framing analysis in the context of social media exist. First, future study is suggested to look into social media routines. Routines of legacy news refer to the journalist-centered daily practices that have been routinized and institutionalized as what Gans (1979) called "organizational routines." In contrast, I define social media routines as the user-centered daily practices that have been routinized and institutionalized by technological features of social media and the collective influence or the "mutual shaping" (Dijck and Poell 2013: 8) among social media users. Based on the comparison of the two semantic networks, I came up with the following three fundamental social media routines, which call for further empirical tests. First, social media are user-centered rather than journalist-centered. Thus, social media produce personalized contents unlike news products, which are constrained by professional journalistic standards and code of ethics. Personalized contents do not necessarily seek reliable sources, news neutralism, and so on. Second, social media tend to interpret the event in micro-level frames, while legacy news prefer macro-level frames. Third, social media create a loosely connected semantic network with several frames, while legacy news make a unified entity with strongly connected interpretive packages.

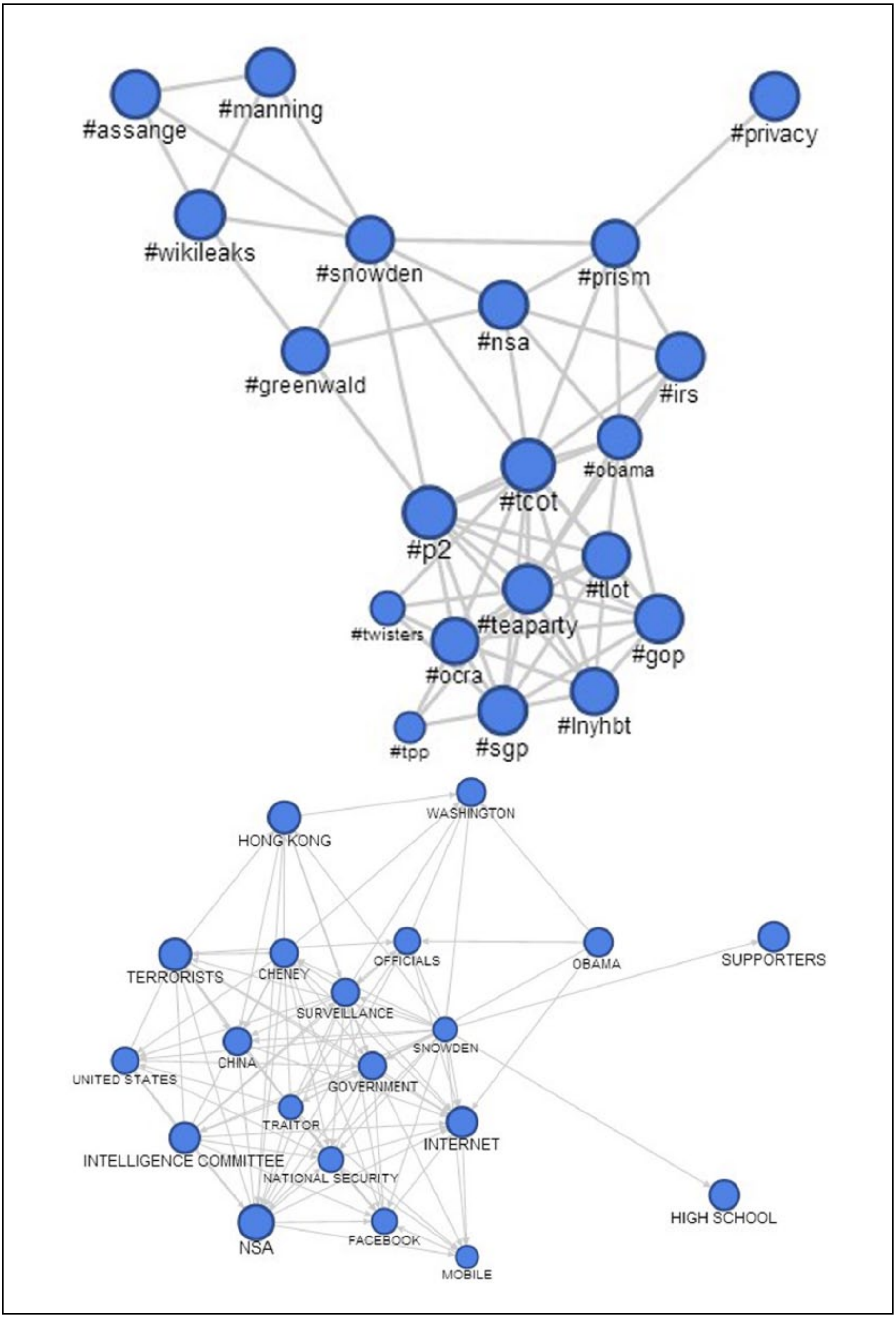

Figure 3. The simplified semantic networks of Snowden incident on social media (above) and legacy news (below).

The second direction for future exploration is social media manipulation. In China, the gray economy of manipulating public opinion on social media is in full swing (Wu et al. 2013). In fact, the Chinese government now plays a vital role in breeding such manipulation. Take the "50 Cent Party" in China as an example. As the name implies, each member of the party will get 50 cents for each comment that is in favor of the government. A line in the latest WikiLeaks movie *The Fifth Estate* goes, "Give a man a mask, and he will tell you the truth." However, the anonymity of the Internet is a double-edged sword. On one hand, it facilitates whistle-blowing. On the other hand, it makes the attempts to identify such manipulations unfeasible, as it is almost impossible to tell a regular comment from a paid comment. However, I doubt the government has the ability to bribe millions of Internet users. But we should be aware that there are entitles like the "50 Cent Party" interfering the online opinion climate and whose impact is probably grossly underestimated.

The third direction is semantic network analysis as a new approach of framing analysis. The foremost merit of semantic network analysis is it lets frames emerge by themselves. If the task of content analysis is to confirm the "known knowns" (Rumsfeld 2002), the greatest strength of semantic network analysis is to piece together the "unknown unknowns" (Rumsfeld 2002), and the anomalies, which are exactly the prime agenda-building activities of intelligence agencies. When doing content analysis, we come up with several frames and then fit the texts in. However, we are at risk of overlooking some elementary mechanisms, over-manipulation, and even misinterpretation. Compared with content analysis, of which the research objects are largely the institutionalized and relatively small amount of media contents, semantic network analysis seems more suitable for dealing with large-scale personalized contents on the Internet. The sizes of the digital texts are usually far beyond the processing capacity of content analysis, but an easy task for semantic analysis with the assistance of computers. In addition, it offers an unobtrusive alternative other than interviews or experiments to detect individual frames.

However, I have no intention to encourage uncritical acceptance of the new approach, given the existence of some problems. Two very likely challenges are that whether the dataset of SenseBot is big enough to cover all the news reports, and whether the one percent tweets is a representative sample of Twitter? In fact, this is a problem that all the Big Data researchers have to face. Big Data enthusiasts may argue that it is not necessary to do sampling as we are equipped to analyze the whole population. However, the fact is the big-data-owners will not share the "whole population" with researchers but a sample. Take Twitter for instance. The 1 percent sample offered by Twitter Streaming API is the only publicly available dataset. The 100 percent public tweets offered by Twitter Firehose are not available for most of researchers. Morstatter et al. (2013) compared the two datasets and found the 1 percent sample covers top hashtags for a large sample size well, which supports the present study, but most results "depend strongly on the coverage and the type of analysis that the researcher wishes to perform." Future studies are encouraged to perform validity tests of this approach and explore more utilities.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was supported by the Department of Media and Communication, City University of Hong Kong.

Note

1. Full lists of words in the two semantic networks are available upon request.

References

ABC News. 2010. "The Conversation: Understanding 'Hashtag Politics.'" http://abcnews.go.com/WN/tcot-drives-tea-party-republican-message-social-media/story?id=11682222.

Associated Press. 2013a. "Snowden 'National Hero' Petition On White House Website Passes 100,000 Signatures."

Associated Press. 2013b. "U.S. 'Extremely Disappointed' after Russia Defies American Wishes, Grants Asylum to NSA Leaker Edward Snowden: White House." http://news.nationalpost.com/2013/08/01/edward-snowden-has-left-moscow-airport-and-entered-russia-lawyer/.

Bennett, W. Lance, and Alexandra Segerberg. 2012. "The Logic of Connective Action." *Information, Communication & Society* 15 (5): 739–68. doi:10.1080/1369118X.2012.670661.

Booz Allen Hamilton. 2013. "Booz Allen Statement on Reports of Leaked Information." *boozallen.com*. http://www.boozallen.com/media-center/press-releases/48399320/statement-reports-leaked-information-060913.

Bruns, Axel, Jean Burgess, Kate Crawford, and Frances Shaw. 2012. *# Qldfloods And@ QPSMedia: Crisis Communication on Twitter in the 2011 South East Queensland Floods*. http://eprints.qut.edu.au/48241.

CyBranding. 2013. "Branding Savvy Social Media Analytics." *cybranding.com*. http://www.cybranding.com/home/about.

Daly, Liza. 2009. "How Big Is the Average ePub Book?" http://blog.threepress.org/2009/11/16/how-big-is-the-average-epub-book/.

Dijck, José Van, and Thomas Poell. 2013. "Understanding Social Media Logic." *Media and Communication* 1 (1): 2–14. http://www.librelloph.com/ojs/index.php/mediaandcommunication/article/view/37.

Doerfel, Marya L., and George A. Barnett. 1999. "A Semantic Network Analysis of the International Communication Association." *Human Communication Research* 25 (4): 589–603. doi:10.1111/j.1468-2958.1999.tb00463.x.

Dorning, Mike, and Chris Strohm. 2013. "Obama Defends Data Spying as Modest Privacy Encroachment." *Bloomberg*. http://www.bloomberg.com/news/2013-06-07/obama-defends-data-spying-as-modest-privacy-encroachment.html.

Dutton, William H. 2009. "The Fifth Estate Emerging through the Network of Networks." *Prometheus* 27 (1): 1–15. doi:10.1080/08109020802657453.

Entman, Robert M. 1993. "Framing: Toward Clarification of a Fractured Paradigm." *Journal of Communication* 43 (4): 51–58. doi:10.1111/j.1460-2466.1993.tb01304.x.

Gamson, William A, and Andre Modigliani. 1989. "Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach." *American Journal of Sociology* 95 (1): 1–37.

Gans, Herbert J. 1979. *Deciding What's News: A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. New York, NY: Pantheon Books.

Goffman, Erving. 1974. *Frame Analysis*. New York, NY: Harper & Row.

Grey, Barry. 2013. "The US Media and the Case of Edward Snowden." http://www.wsws.org/en/articles/2013/07/02/snow-j02.html?view=mobilearticle.

Hemphill, Libby, Aron Culotta, and Matthew Heston. 2013. "Framing in Social Media: How the US Congress Uses Twitter Hashtags to Frame Political Issues." http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2317335.

Herridge, Catherine. 2013. "NSA Data Center Front and Center in Debate over Liberty, Security and Privacy." *FoxNews.com*. http://www.foxnews.com/tech/2013/04/12/nsa-data-center-front-and-center-in-debate-over-liberty-security-and-privacy/.

Hewitt, Steve, and Scott Lucas. 2009. "All the Secrets That Are Fit to Print? The Media and US Intelligence Agencies Before and After 9/11." In *Spinning Intelligence: Why Intelligence Needs the Media, Why the Media Needs Intelligence*, ed. R. Dover and M. S. Goodman, 103–16. New York, NY: Columbia University Press.

Kinder, Donald R., and Lynn M. Sanders. 1990. "Mimicking Political Debate with Survey Questions: The Case of White Opinion on Affirmative Action for Blacks." *Social Cognition* 8 (1): 73–103.

Kirn, Walter. 2013. "Why Edward Snowden Wins on Twitter." *The New Republic*. http://www.newrepublic.com/article/113457/why-edward-snowden-wins-twitter.

Koenig, Thomas. 2004. "Routinizing Frame Analysis through the Use of CAQDAS." Paper presented at the *RC33 Conference*, Amsterdam.

Meraz, Sharon, and Zizi Papacharissi. 2013. "Networked Gatekeeping and Networked Framing on #Egypt." *The International Journal of Press/Politics* 18 (2): 138–66. http://hij.sagepub.com/cgi/doi/10.1177/1940161212474472.

Mirkinson, Jack. 2013. "Edward Snowden Criticizes Media's Coverage of Him." *The Huffington Post*. http://www.huffingtonpost.com/2013/06/17/edward-snowden-media-coverage_n_3454892.html.

Morstatter, Fred, Jürgen Pfeffer, Huan Liu, and Kathleen M. Carley. 2013. "Is the Sample Good Enough? Comparing Data from Twitter's Streaming Api with Twitter's Firehose." *CoRR*. http://arxiv.org/abs/1306.5204.

Newport, Frank. 2013. *Americans Disapprove of Government Surveillance Programs*. http://www.gallup.com/poll/163043/americans-disapprove-government-surveillance-programs.aspx.

Pan, Zhongdang, and Gerald M. Kosicki. 1993. "Framing Analysis: An Approach to News Discourse." *Political Communication* 10(1): 55–75. doi:10.1080/10584609.1993.9962963.

Papacharissi, Zizi, and Maria de Fatima Oliveira. 2012. "Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt." *Journal of Communication* 62 (2): 266–82. doi:10.1111/j.1460-2466.2012.01630.x.

Parmelee, John H. 2013. "The Agenda-Building Function of Political Tweets." *New Media & Society* (May): 1–17. http://nms.sagepub.com/cgi/doi/10.1177/1461444813487955.

Piaget, Jean. 1929. *The Child's Conception of the World*. London: Routledge and Kegan Paul.

Rambukkana, Nathan. 2013. "CFP-Hashtag Publics: The Power and Politics of Networked Discourse Communities." http://complexsingularities.net/wp-content/uploads/2013/06/Hashtag-Publics_CFP_new-deadline.pdf.

Rumsfeld, Donald H. 2002. “DoD News Briefing—Secretary Rumsfeld and Gen. Myers.” *defense.gov*. http://www.defense.gov/Transcripts/Transcript.aspx?TranscriptID=2636.

Scheufele, D A. 1999. “Framing as a Theory of Media Effects.” *Journal of Communication* 49(1): 103–22. doi:10.1111/j.1460-2466.1999.tb02784.x.

Semetko, H. A., and P. M. Valkenburg. 2000. “Framing European Politics: A Content Analysis of Press and Television News.” *Journal of Communication* 50 (2): 93–109. doi:10.1111/j.1460-2466.2000.tb02843.x.

Sukosd, Miklos, and King-wa Fu. 2013. “How Chinese Netizens Discuss Environmental Conflicts? Framing and Functions on Sina Weibo.” http://com.cityu.edu.hk/COMDOC/Seminar/paper/2013/seminarPaper-2013-09-16-1.pdf.

Tankard, J. W. 2001. “The Empirical Approach to the Study of Media Framing.” In *Framing Public Life: Perspectives on Media and Our Understanding of the Social World*, ed. S. D. Reese, O. H. Gandy, and A. E. Grant, 95–106. Mahwah, NJ: Lawrence Erlbaum.

Twitter. 2014. “Using Hashtags on Twitter.” https://support.twitter.com/entries/49309#.

Wahl-Jorgensen, Karin, and Joanne Hunt. 2012. “Journalism, Accountability and the Possibilities for Structural Critique: A Case Study of Coverage of Whistleblowing.” *Journalism* 13 (4): 399–416.

Walsh, Kenneth T. 2013. “Ron Paul: Edward Snowden a Hero.” *U.S. News*. http://www.usnews.com/news/blogs/ken-walshs-washington/2013/06/11/ron-paul-edward-snowden-a-hero.

Wu, Mei, Peter Jakubowicz, and Chengyu Cao. 2013. *Internet Mercenaries and Viral Marketing: The Case of Chinese Social Media*. Hershey, PA: IGI Global.

Wu, Shaomei, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts. 2011. “Who Says What to Whom on Twitter.” *Proceedings of the 20th international conference on World wide web - WWW* ’11: 705. doi:10.1145/1963405.1963504.

Yuan, Elaine J., Miao Feng, and James A. Danowski. 2013. “‘Privacy’ in Semantic Networks on Chinese Social Media: The Case of Sina Weibo.” *Journal of Communication* 63:1011–31.

Zhang, Xinzhi. 2013. “Four Interpretive Frames to Make Sense of Snowden Incident.” *Media Digest* 07:4–5. http://rthk.hk/mediadigest/20130715_76_123002.html. [In Chinese].

Author Biography

Jie Qin got her Ph.D. from the Department of Media and Communication, City University of Hong Kong, and now is the Professor of Communications at the Hong Kong campus of Savannah College of Art and Design. Her research interests are in areas of social media, nonprofit organizations’ use of ICTs, and social network analysis.