

RESEARCH

Open Access

# The fifteen days: building claims and contexts



Eunkyung Song

Correspondence:  
[suhyoo.ek@gmail.com](mailto:suhyoo.ek@gmail.com)  
Sociology, Rutgers University, New  
Brunswick, USA

## Abstract

This paper examines the process of protest claim-making by reconstructing the semantic structure of online communication that took place prior to the first street event of a protest. Topic networks are identified on the basis of topic modeling outputs, deeming topics to be connected if they share the same terms. Semantic communities within such topic networks are specified by patterns of inter-topic relations with the aim to delineate the emergence of cohesive semantic communities as protest claims. This paper argues that protest claims are developed through the repetition of specific arrangement of topics, which generates relevance among topics discussed concurrently and continually over time. The iterated patterns shape semantic coherence through brokering terms that connect topics in consistent ways. Findings are presented by investigating 17 daily corpora of digital posts produced on a bulletin board from April 16 to May 2, 2008, which constitutes the pre-protest period of the Candlelight Protests in South Korea against a government policy on beef trade with the United States.

**Keywords:** Semantic communities, Topic networks, Topic modeling, Claim-making, Large-scale protests, Digital interactions

## Introduction

In the early evening of May 2, 2008, about 10,000 people gathered together in central Seoul, South Korea, each holding a candle in their hands to protest an upcoming policy change to allow imports of previously-prohibited beef products from the United States. The products contained Specified Risk Materials (SRMs) known to cause Bovine Spongiform Encephalopathy (BSE) in cattle. Addressing concerns over its human variant, Creutzfeldt-Jacob disease (vCJD), the protesters demanded the Korean government to rescind the policy change. Similar street events subsequently took place over the following three months, which were later collectively named the Candlelight Protests of 2008 (Candlelight Protests). This left indelible marks on the Lee administration into its third month in domestic politics. Globally, the Candlelight Protests, which were based on anonymous discussion boards and online communities, preceded similar protests such as the Occupy Movement in the United States, the Arab Spring in the Middle East, and the 15 M movement in Spain, the Gezi Park protests in Turkey, and the Umbrella movement in Hong Kong, which were largely organized on social media.

This paper examines the dynamics of claim-making online in the lead-up to a large-scale protest, using the case study of the Candlelight Protests. My goal is to specify the meanings of the claims of the Candlelight Protests by tracing their emergence. Locating this task at the intersection of social network analysis (SNA) and social movement studies (SMS), I argue that protest claims are shaped rather than spread in its invariable format from the outset of protest mobilization. To demonstrate, this study analyzes digital posts that were produced during their pre-protest period by employing topic modeling and community detection. Mapping out the dynamics of topic networks, this study finds that descriptive terms of the Candlelight Protests (e.g., “beef,” “produced in the US,” “mad cow disease,” etc.), which had initially existed as issues with no salience, eventually became brokering terms that contextualize seemingly-irrelevant issues and concrete protest logistics as the components of the claims.

The following section situates the current research inquiry at the nexus of SNA and SMS, particularly in the tradition of cultural analysis, which is crucial to understand collective action. The third section introduces data, followed by an analytic framework that underscores conditions that are considered in employing topic modeling and detecting semantic communities in this study. The fifth section illustrates the dynamics of topic networks through which protest claims are developed. The last section discusses the implication of the findings as well as suggestions for future research.

### **The meanings of protest claims**

The SNA scholarship in SMS has enhanced the structural analysis of collective action, centering on the roles of social movement organizations (SMOs) within it (Diani and McAdam 2003; McAdam 2003; Diani 2004). The inquiry of who participates in collective action has drawn attention to the interdependent dimension of one’s participation on another person’s participation (Snow et al. 1980; McAdam 1988; Marwell and Oliver 1993; Gould 1993, 2003; Passy 2003; Passy and Monsch 2014). At the individual level, one’s decision to participate increases the likelihood of a collective goal being achieved, which would subsequently offer either a positive or negative basis for another person’s participation (to name a few, Granovetter 1978; Oliver and Myers 2003). Beyond interpersonal connections, inter-organizational relations created through coalition-building or activists’ affiliations with multiple organizations have led the inquiry of how revolutions or popular protests build a broad base of solidarity that transcends boundaries between social groups (Fernandez and McAdam 1989; Mische and Pattison 2000; Rucht 2004; Mische 2008). Conceptualizing collective identity through the notion of structural equivalence, Gould (1995) points out that collective identities contribute to the formation of solidarity among individuals who ordinarily juggle multiple identities under situational contingencies. The same aspect regarding multiple identities is conceptualized in Mische’s work as to how mediating roles are played in different ways by activists affiliated with multiple organizations (Mische 2003, 2008).

This long and robust tradition of SNA in SMS, however, has faced obstacles when examining cases of large-scale protests that are mainly mobilized through online communication. A primary reason for this is that such recent protests possessed a tenuous organizational basis (Earl and Kimport 2011; Castells 2012; Gerbaudo 2012; Bennett and Segerberg 2013). Organizational affiliations have become surprisingly diminished in their

importance, if not outright irrelevant, as a factor in protest mobilization.<sup>1</sup> Instead, communicative ties, which are relatively instantaneous, have risen as a sufficiently manifest factor to ensure and indicate individual's decision on participation. SNA researchers, who are less burdened with SMS's theoretical focus on SMOs, have rapidly adopted communicative commonalities (e.g., common terms found in digital posts written by different authors) as a measure of ties. To this end, they have reconstructed a proxy protest network in their entirety with as many discernible ties as possible at a given point of time (Woolery et al. 2016; González-Bailón and Wang 2016).

In contrast, social movement researchers have rarely attempted to examine the largest possible number of protesters, due to both practical and theoretical limitations. On the practical side, such data have not been always available. It is often impossible to survey or interview all participants at the site of a protest to gather data on their social relations. This is likely why certain large-scale protests or revolutions (e.g., the French Revolution, the Civil Rights Movement in the United States) have been studied more widely than others, as their development processes have been documented more extensively in relative terms by various groups and government agencies. There have also been lingering concerns regarding the reduction of collective action to a standstill mechanism with a focus on dynamics that originate from the interplay of contingency and structural factors (Tarrow 1993; McAdam and Sewell 2001; McAdam et al. 2001). Visualizing a protest network as a single entirety may have rekindled such a concern.

On the theoretical front, the emphasis on SMOs and leading activists in SMS have privileged interactions initiated by activists, which is often not applicable to digitally-mobilized protests. This conventional approach has been rarely challenged, with a few exceptions found in the cultural strand in the field. Framing theory argues that protest mobilization cannot be accomplished without the sharing of "frames" that specifies an issue as social injustice that can only be resolved through collective action (Snow et al. 1986, 2014). Framing theory contends that, when recruiting new members, activists are supposed to articulate their claims and viewpoints in advance, and adjust the frames discursively in accordance with the audience's dispositions. This directionality limits new recruits to a passive position and consequently disregards mutually influential communication. Advancing this theory, later researchers have examined the interactive and relational dimensions of protest mobilization by studying various media such as stories (Polletta 2006), discourses (Steinberg 1998, 1999), and conversations (Mische 2003, 2008). These communicative media depict activists and non-activists as mutually influential and helpful in placing a protest in a broader context, where discursive counteractions by opponents or targets can be considered simultaneously, though this approach still prioritizes activists and SMOs. Thus, it is no surprise that social movement researchers have yet to reach a definitive conclusion expanding existing theories to incorporate multi-pronged communicative flows online where such a discursive center emerges through communication rather than existing in the first place.

Another possible theoretical reason for the divergence in the structural appreciation of collective action pertains to the scarcity of available information on digital media users,

---

<sup>1</sup>This point requires further discussions to specify its process. Empirically, experienced and existing SMOs in South Korea were not deeply involved in the earlier mobilization process of the Candlelight Protests. Nevertheless, the Candlelight Protests throughout their overall development cannot be explained without referring to a big coalition of SMOs nationwide.

which starkly contrasts with the abundance of their verbal traces. In the communicative environment of digital media, it is rare for enriched relational data to become available to researchers for ethical reasons and various digital media layouts. Moreover, the paucity of such data may need to be tolerated as it is, indeed, to properly reflect the aspect of the recent large-scale protests, where mere strangers with no previous social encounters coordinate actions together for a shared cause in reference to words circulated online.

Therefore, large volumes of digital posts imply a two-sided challenge for both SNA and SMS. First, SNA analysts have not fully explored the depth of verbal expressions, despite its own tradition in semantic network analysis (to name a few, Carley and Kaufer 1993; Carley 1994; Bearman and Stovel 2000). Semantic network analysts have pointed out that messages, i.e., the meanings of terms, are forged by the arrangement of the terms in relation to others in a given text beyond their lexical meanings. The negligence of the relationality of messages has promoted the presupposition that a common term used by different authors connotes the same meaning, thereby disregarding the arrangements in which it is embedded (DiMaggio et al. 2013; Fuhse 2015; McLean 2017). In this regard, digital posts and their constitutive terms become important, in the investigation of their constitutive when involving many people, not because they instantiate the shallow fact of whether those posts contain the same terms.

For social movement researchers, the abundance of digital posts presents another aspect of the puzzle of how, then, protest-related texts, i.e., posts, can be compartmentalized without noise. The nature of this puzzle has never fallen beyond the purview of SMS researchers in terms of utilizing text materials (Tilly et al. 1975; Johnston 1995; Wada 2012; Krinsky and Mische 2013). In particular, the cultural strand of SMS has developed its theories by incorporating various speech acts and different kinds of interaction settings. For instance, Mische (2003) argues that informal dialogues among activists can be a source of mediation among activists from different political groups. Nonetheless, there is some ambiguity regarding what speeches, conversations, and writings should be counted as protest-relevant in online communication, insofar as the degree of relevance is assumed to be defined by a researcher. An online community of amateur photographers can become a potent locus of an upcoming protest regarding abortion. A discussion board tends to present topical incoherence because of a variety of issues presented by different users. Nevertheless, it can unexpectedly become a hub through which protest communication thrives when its heterogeneous users come to concentrate on the protest by presenting their own viewpoints. If the communicative environment is permissive to various and heterogeneous issues, filtering out “relevant” digital posts from “irrelevant” ones can lead to an analytic bias, thereby overlooking the fact that messages are relational.

Based on these considerations, this study explores the semantic structures of digital posts, considering the following points. First, protest claims are defined here as challengers’ (or protesters’) demands against their targets (McAdam et al. 2001). According to social movement scholars, protest claims encompass a boundary drawn between challengers (connoted as “we”), opponents, and targets (referred to as “they”), substantive issues, specified goals to achieve, as well as concrete logistics such as tactics and strategies (Goodwin and Jasper 2004). Second, to trace the dynamics of claim-making process, temporal markers, i.e., by day, are adopted to subdivide digital posts into daily collections, instead of aggregating them all together. Lastly, posts are not filtered out by

their relevance to the Candlelight Protests to capture how users of the studied platform came to develop certain patterns in discussing the upcoming protests. In other words, this study utilizes all available digital posts within a given period of time on a digital platform, whose merits and disadvantages will be discussed below.

#### **Data: digital posts and digital platform**

This study analyzes 15,634 posts published from April 16 to May 2, 2008 on Agora's free discussion board (FDB). Serviced by the Korean portal site Daum,<sup>2</sup> Agora was a bulletin board that had multiple subdirectories allotted to broad domains of topics (e.g., politics, the economy, petitions, etc.) in addition to the FDB. It attracted high publicity to its posts by requiring posters to log in with their Daum account but allowing readers free access without the need to log in. Unlike other subdirectories with set topical boundaries, the FDB had no guidelines concerning the selection of topics. Combined with Daum's major news service, which offered news articles compiled from numerous news organizations, the FDB was sensitive to real-time updates of breaking news. These platform-specific characteristics constitute an ambivalent environment for Agora's FDB users who may want to mobilize specific agendas. Its fast turnover of issues could be useful in drawing initial attention in a timely manner, but at the same time, continuous floods of different issues have the potential to severely undercut efforts to develop coherent claims.

Considering these platform-based attributes, the daily dynamics of issues in the FDB is factored in by compartmentalizing the collected posts into daily corpora. The target period of time starts from April 18, the day when the core issue—a policy change to the inspection standards for importing beef products from the United States—was first announced and ends on May 2, the day of the first street gathering. The two days prior to the start point, April 16 and 17, were also included to observe ordinary activities in the FDB before the beef trade issue. Table 1 profiles all daily corpora by the quantities of posts, authors, and terms used. The numbers of posts and authors have remained stable up to April 27, and rapidly increased until May 2.<sup>3</sup>

#### **Topic modeling and community detection**

##### ***Topic modeling and topic networks***

Topic modeling is a generative and probabilistic model that treats a corpus of multiple documents as a set of topics by clustering its constitutive terms on the basis of their probabilistic co-occurrence (Blei and McAuliffe 2010; Blei 2012). Terms assigned to one topic can resurface in other topics with the assumption that a document can encompass multiple topics. These features render topic modeling a compelling method of identifying nuanced subtleties of terms by revealing how the same terms eventually connote different contexts by being embedded in different topics (DiMaggio et al. 2013). For this study, brokering terms in topic networks receive analytic attention regarding their relational roles that show how digital media users connect different

---

<sup>2</sup>The data were collected several years after 2008, and as a result, some posts may have been removed by their authors.

<sup>3</sup>This author appreciates a reviewer's comment on how the below analysis can also include the attributes of digital post authors. As an initial step towards developing a more comprehensive analysis of protest networks, this study exclusively examines topic networks.

**Table 1** Profile of the collection of digital posts

Date	Number of posts	Number of authors (number of unidentifiable authors <sup>a</sup> )	Number of terms <sup>b</sup>
April 16	295	188 (23)	18,251
April 17	433	251 (39)	24,872
April 18	435	246 (31)	25,697
April 19	418	239 (35)	22,934
April 20	399	229 (29)	23,265
April 21	491	245 (38)	20,630
April 22	580	327 (38)	27,533
April 23	403	236 (30)	23,540
April 24	460	260 (22)	22,066
April 25	606	322 (39)	30,600
April 26	474	250 (58)	23,440
April 27	437	257 (45)	22,820
April 28	818	469 (42)	33,086
April 29	924	523 (94)	45,574
April 30	1909	866 (122)	48,448
May 1	2475	1137 (152)	68,375
May 2	4209	1742 (169)	83,018
Total	15,634	7787 (1006) <sup>c</sup>	564,149

Note: <sup>a</sup>This indicates posts whose authors are left blank.; <sup>b</sup>These terms are included in a document term matrix.; <sup>c</sup>The number of unique authors for the entire collection is 4496

topics, whereby the protest claims become contextualized, and acquire their underlying connotations in relation to other topics.

Using topic modeling, the following decisions were made. First, unsupervised topic modeling is employed, despite the known concerns in terms of validity confirmation (DiMaggio 2015; McFarland and McFarland 2015). Developing a test corpus against which the rest of the documents can be trained would enhance the validity of research outputs, given that computing programs never supersede a researcher's sophisticated understanding of the studied documents. And yet, this study relies on baseline topic modeling because the lack of information about pre-protest communication, particularly those that takes place online.<sup>4</sup> Second, the number of topics that should be predetermined by the researcher was set to 10 and applied to all daily collections. This decision remains fundamentally arbitrary, and can make the model insensitive to daily variations in the studied collection, quantitatively and qualitatively. As Table 1 illustrates, the daily collections nearer May 2 obviously contain more digital posts, which may imply the appearance of more heterogeneous topics. The implication of this variation in the quantity of posts is somewhat ambiguous. For instance, there may be fewer discernible topics in the daily collections nearer the end point of the data, if reasonably coherent discussions over ideas on

<sup>4</sup>It is rare in SMS to study fledgling activism before the realization of desired outcomes. A notable exception is Blee (2012), who traces small groups at the initial stage of their activism by setting out in-person gatherings on the basis of broadly-defined agendas. In her analysis, she argues that many groups eventually fail to develop a ground for activism despite their general consensus, with which they are able to initiate meetings. Such failure, according to Blee, pertains to a challenge to embark on a relatively consistent flow of discussion, whose preceding conclusions serve the following communication, which she referred to as the effect of 'path dependency.'



an upcoming protest prevailed. In contrast, the same collections may contain more topics if digital media users are convinced that the upcoming protest would actually take place and become more enthusiastic by suggesting more ideas at a rapid pace.<sup>5</sup> Inferring the probable outcome for the studied platform is not a simple task, given its openness to topics as well as public access. To handle these uncertainties in accordance with the research question of this study, which searches for the emergence of a reasonably coherent protest claim, the author selected April 16, 2008, as a yardstick to present the ordinary setting of the studied digital platform, where a 10-topic model returned a decent performance of clustering topics.

Thirdly, the nature of the current study may exacerbate the known drawback of topic modeling regarding noise. Regardless of the number of topics preset by the researcher, topic modeling itself is never immune to imperfect clustering that makes some topics insensible. This limitation can be possibly aggravated in this study because the research design intends to trace the emergence of relevance out of noise. For instance, it can be the case where a discussed subject pertains less to the protest theme on the surface, but is captured as relevant in topic modeling outcomes, if the subject is discussed in the terms shared with the terms of the protest theme. To prevent the misinterpretation of noise as meaningful outputs, consistency in the appearance of similar topics will be discussed. Also, topic labels given by the author often contain two words to help the clarification of their contents.

Fourth, to run a topic model on digital posts written in the Korean language, two pre-processing principles were applied to combat the bag-of-words assumption in topic modeling, which completely disregards the grammatical roles of terms. On one hand, some preprocessing recommendations were disregarded, considering linguistic features in the Korean language such as its highly developed suffixes. Unlike the English language, in which a word has its own independent grammatical value and is therefore placed between whitespaces, the Korean language creates a word by combining two different grammatical units (e.g., nouns and postpositions) without a whitespace.<sup>6</sup> The drawback of this procedure is the redundancy that results from treating the same noun with different postpositions as different words altogether. In particular, it is very likely that 'core terms' for a topic appear multiple times because they are regarded as

<sup>5</sup>In SMS, there has been a lack of discussion over temporality in terms of how rapidly (or slowly) a protest is mobilized (McAdam and Sewell 2001). When this issue has been addressed, it mainly focused on a long-term structural effect in connection with the notion of protest cycles (Tarrow 1993, 2010; Oliver and Myers 2003) emphasizing that the effects of mobilizing factors that would last longer. What is intriguing in the study of online mobilization is that its communication process contains uncertainties in its temporal dimension. First, issues presented online with the intention of mobilizing public attention do not always achieve the desired attention at a desired pace. Second, it is also difficult to distinguish "long-term" effects from "short-term." For instance, activism that successfully obtains a large-scale following appears both to touch upon causes that have long been discussed and to trigger a certain mechanism that propels focused attention at an unexpectedly rapid pace. Further discussion on this topic will follow.

<sup>6</sup>When employing general text pre-processing protocols to texts in English, prepositions are removed. Therefore, the sentence "I gave a book to her" would contribute to a topic model without the preposition "to." In the Korean language, a suffix that plays the same grammatical function is attached to the noun, book without a whitespace between the two units, "*ch'aekül*" (*ch'aek* (book) + *ül*), specifying its role within a sentence (in this example as the object of the sentence). It is possible to remove suffixes by conducting a part-of-speech analysis. A challenge to the procedure considered in this study is that it is unknown how an optimal level of part-of-speech analysis can be found, preventing the increase of complexities in the data. These effects will be analyzed in follow-up studies. Names of people were not deleted from the corpora because Korean names do not contain a whitespace between the first and last name, thereby remaining as a specifiable single word. Words in non-Korean letters were also retained, reflecting the fact that digital posts often use many functional terms online (e.g., the word 'petition' in a hyperlink) and terms that do not have Korean translations.

different terms. The expected effect of redundancy is two-pronged. Coherence across terms within a topic can be enhanced, but such redundancy can also affect inter-topic relations because the redundant terms are regarded as different terms, which resultantly reduce the degree of inter-topic relations.

In addition, this study retains plural pronouns, e.g., “we” and “they,” and their derivatives, though this is rarely recommended because of the bag-of-words assumption. Although this reluctance is valid, this author retained first- and third-person plural pronouns, because these pronouns are major symbolic and strategic indicators of collectivity, not only in SMS in general (Polletta and Jasper 2001; Taylor and Whittier 1992), but also in digitally-mobilized activism (Gerbaudo 2012; Gerbaudo and Treré 2015). The rationale for this decision is as follows: If such pronouns for identifying opposing fronts prevail in the collection being examined, the pronouns can be captured as a relatively independent topic that may contain not only such pronouns but also terms that are more likely to be associated with such pronouns. Also, the semantic roles of these pronouns can be investigated by checking their associated terms. After text pre-processing, data-term matrixes in this study used 564,149 terms. Topic modeling analysis conducted using R utilized the lda gibbs sampler.

#### ***Semantic communities: clusters of topics***

Topic modeling outputs were translated as a two-mode network where one node set represents topics and the other terms, which was transformed as a one-mode topic network. The goal of this procedure is (1) to specify daily topic networks, (2) to investigate structural patterns across the daily topic networks by finding semantic communities and their dynamics by calculating modularity as a way to capture the relationship between intra- and inter-semantic communities, and (3) to examine the types of terms found at the brokering position.

To achieve these goals, the following decisions were made in accordance with the characteristics of the topic modeling in use. First, because topic networks can be sensitive to the number of terms, topic networks were created using different numbers of terms, respectively top 15 terms and 30 terms, though the discussion below mainly pertains to outcomes of the former. For instance, two topics that are not linked in a 15-term topic network can be tied, albeit lightly, in a 30-term network. Regardless of how many terms are used, this can occur because a shared term between two topics can have a different impact in each connected topic, whereby they are ranked differently. The critical point of this study is to obtain topic networks that (1) avoid reconstructing networks that are either too sparse or too densely connected, and (2) to secure a reasonable ground for interpretation. Although these procedures require improvement for future research, this author decided to (1) reduce redundancy originating from the current topic model by disregarding suffixes for the same terms to ensure that they are counted only once, which resultantly reduces the number of terms in topic networks smaller than the original topic modeling outcomes; and (2) rely on the number of ties by shared terms among topics without weighting.<sup>7</sup>

<sup>7</sup>A likely solution would be to utilize all classified terms, or to use a certain number of terms and weight ties by considering the probability of a term's assignment to a topic and the probability of the topic's presence in a given corpus. This author appreciates the reviewers who made suggestions regarding this point.



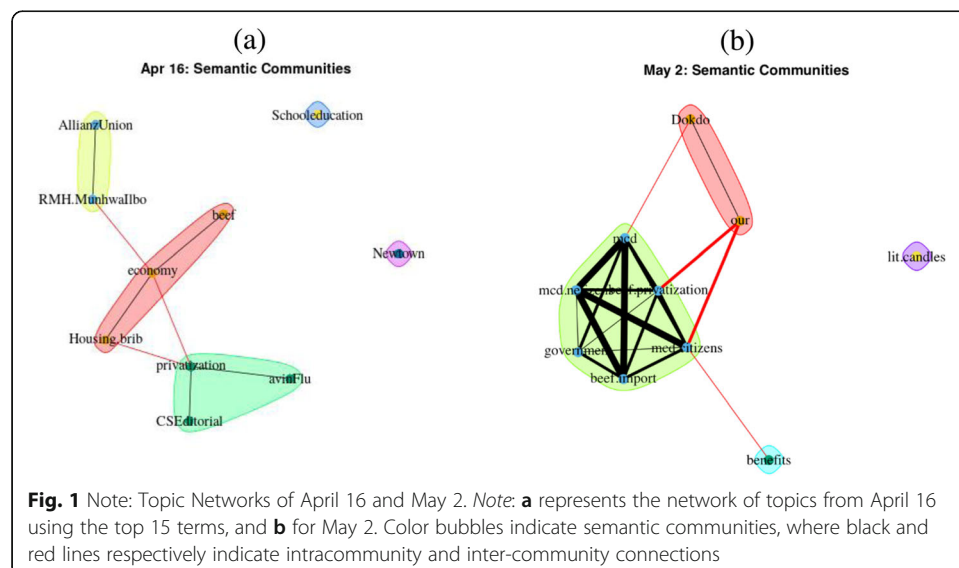
The findings are reported as follows. First, the two topic networks of April 16 and May 2 are briefly compared in Fig. 1, and contextualized with modularity scores profiled in Fig. 2. Second, using the modularity score of 0.2 as the standard, the daily topic networks are divided into two groups with either lower modularity (see Fig. 3) or higher modularity (see Fig. 4) to specify structural patterns.

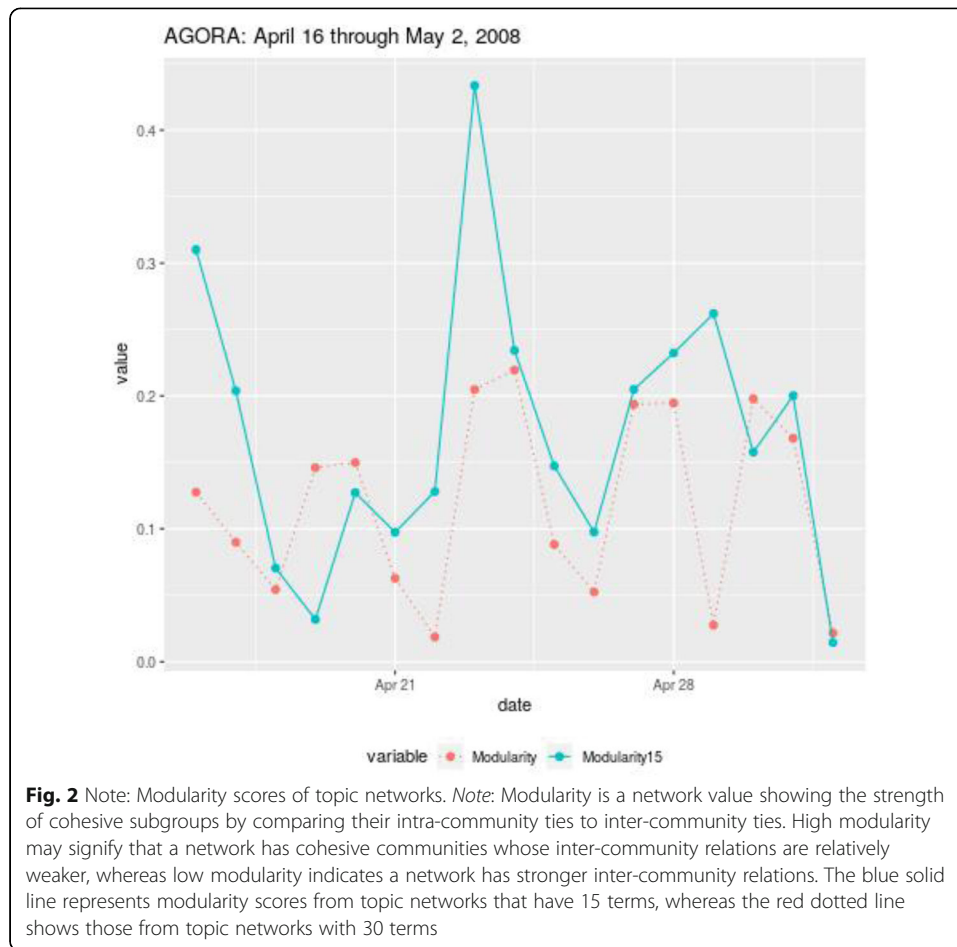
#### Topic networks and semantic communities: April 16 and May 2

Figure 1 profiles topic networks for April 16 (a) and May 2 (b). Color bubbles around the nodes indicate semantic communities whose member topics are linked to one another within the same community (black lines) than to nodes outside the community (red lines).

The openness of Agora's FDB yields a semantic network with a variety of topics. On April 16 (see Fig. 1(a) and Table 2), the scope of subjects ranges from education reform for secondary school (*SchoolEducation*); an outbreak of avian flu (*avianFlu*); a proposal for health insurance reform (*privatization*); a redevelopment plan for several districts in Seoul (*NewTown*); to the beef trade issue in concerns over mad cow disease (*beef*). These topics are not intrinsically relevant to one another. It is unlikely that there is any inherent connection between thoughts on introducing an honors class in secondary education (*SchoolEducation*) and ideas on how imported beef products should be inspected. Also, it is notable that some topics have terms that are more coherent, which helps interpretation, such as topic *SchoolEducation*, which contains terms that are intuitively related to education policy. Topic *beef* presents terms such as “beef,” “risk materials of mad cow disease in US beef products,” and “produced in the US,” alongside “wealthy” and “the rich,” insinuating that the beef trade issue was mentioned in connection to the idea of social inequality. In contrast, topic *Housing.brib* delivers terms related to housing and those close to a political scandal, which are captured together.

In terms of the structure of the April 16 topic network, it is important to consider the common terms that link topics: (1) abstract or general terms in reference to policy areas or politicians, such as “the economy,” and “politics”; (2) phrases including

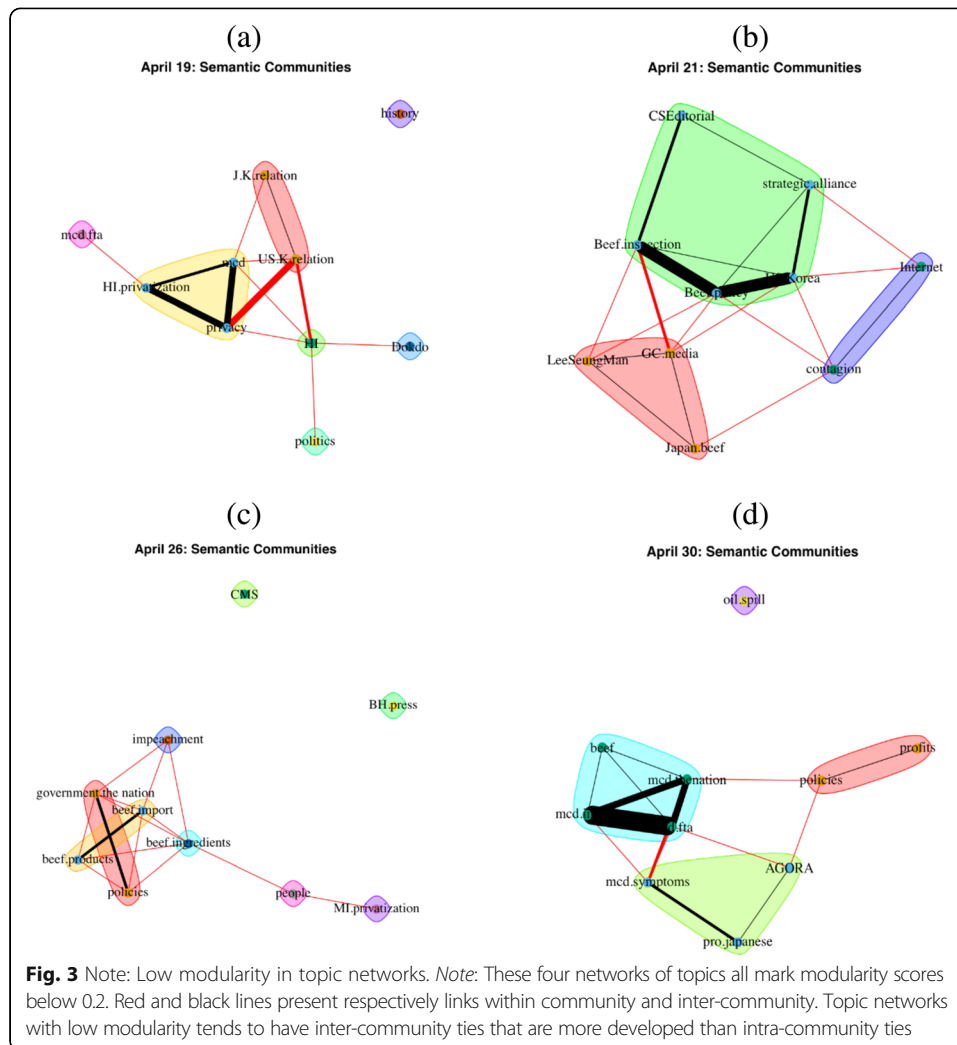




pronouns such as “they/their”; and (3) perspective-based and value-based terms such as “problems,” and “wish.”<sup>8</sup> More specifically, topics *beef* and *economy* share the term “the economy.” For topic *beef*, the term “economy” contextualizes beef import from the U.S. in relation to social inequality in relatively abstract ways alongside other terms such as “wealthy” and “the rich.” In contrast, topic *economy* concerns the politico-economic dimension of the Grand Canal Project, an initiative revealed as a key election promise by President Lee to build a large-scale waterway connecting major rivers in Korea. This type of semantic connection appears abstract and general in the sense that it portrays the two issues in the context of the economy. In addition, terms assigned to topic *beef*, such as “beef” and “risk materials of mad cow disease in the US beef products,” remain exclusively within their own topic.

To conclude, the various topics in the April 16 network imply that Agora users tended to discuss national politics in diverse areas of life. This also suggests that its semantic communities rarely share substantial contexts beyond a general subject area. Moreover, Agora users showed a tendency to appeal to the audience by using collective pronouns regardless of the topic at hand, although it is yet difficult to discern whether

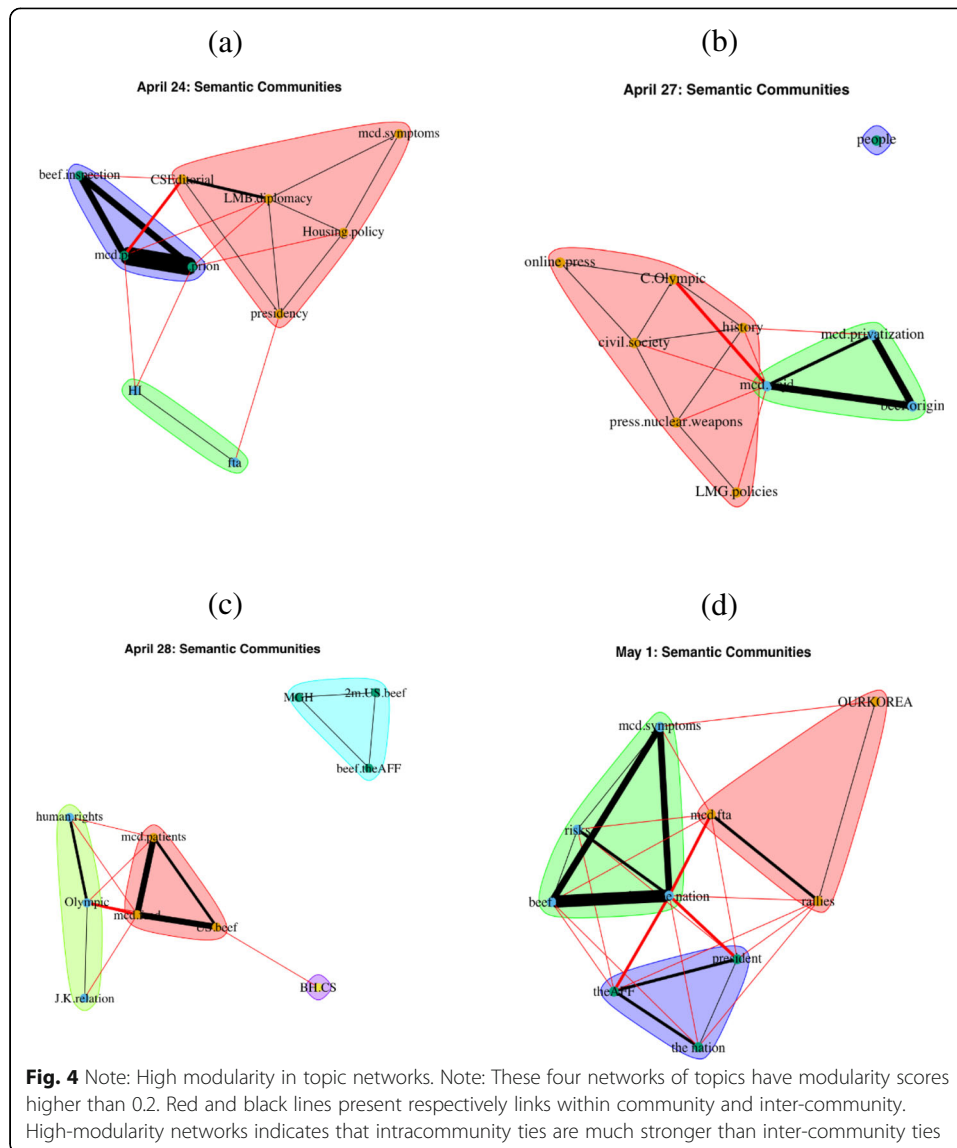
<sup>8</sup>When top 30 terms are used, the number of shared terms increases such as “President Roh,” “participatory government,” “from now on,” “to what extent,” “properly,” as well as other pronouns such as “for all of us,” “to you,” “of oneself,” and “we.”



these pronouns carry uniform connotations across the entire network. The beef trade issue was not entirely at the forefront of public attention, although it had been recognized and discussed. In this sense, it is unlikely that Agora users on April 16 foresaw a collective action on the beef trade issue, which was then simply one among myriad topics that momentarily commanded the interest of Agora users.

The May 2 semantic network (Fig. 1(b) and Table 3) demonstrates a different narrative. Above all, beef trade-related topics prevail. Seven out of ten topics address U.S. beef importation in a nuanced manner and contain terms such as “mad cow disease,” “beef,” “imported from the United States,” “Lee Myung-bak,” and “crazy cow.”<sup>9</sup> Six of them are classified as a single semantic community, which results in highly dense connections among them as thick black ties that demonstrate: (1) topic *beef.privatization* discusses imported beef products in relation to Lee’s political drive to privatize public sectors; (2) topic *mad.netizens* points out that netizens (a general term denoting Internet users) are aware of the possible danger of consuming U.S. beef products processed from cattle raised on animal-based feeds that could cause abnormal proteins such as

<sup>9</sup>The words “crazy” and “mad” in English are often used interchangeably in Korean.



prings; (3) topic *mcd* concerns how conservative politicians of the ruling party ignored public anxiety over the resumption of U.S. beef imports, alluding to a TV show, PD's Note, that focused on the danger of mad cow disease; (4) topic *beef.import* describes the poor handling of the beef trade issue by the Korean government in its negotiations with the U.S.; (5) topic *mcd.citizen* concerns the responses of political parties to the beef trade issue; and (6) topic *government* concerns the incompetence shown by Lee's presidency in an abstract way.<sup>10</sup> It is evident that Agora users on May 2 were predominantly occupied with the beef trade issue in diverse aspects, compared to April 16.

Four topics that are clustered into different communities are equally worthy of mention to demonstrate the complexity of this topic network. Topic *Dokdo* concerns the Japanese government's distortion of history through their secondary school history textbooks and territorial claims over Dokdo Island, while topic *benefit* criticizes exploitative economic

<sup>10</sup>Compared to other topics, topics *government* and *benefits* contain terms in a less detectable way.

**Table 2** Top 15 Terms of the Topic Modeling Result of the April 16 Collection

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
	School Education	Avian Flu	RMH-Munhwa	Privatization	Economy	Newtown	Allianz Union	CEEditorial	Beef	Housing, bribe
1	<i>uyōiban</i> (honors class)	<i>saramūn</i> (person)	<i>chōngch'ōngnae</i> (Chung Cheong-rae)	<i>han'gugūi</i> (Korean)	<i>ōlmana</i> (how much)	<i>hannaradang</i> (Grand National Party)	<i>allianch'ūsengmyōng</i> (Allianz Life)	<i>imyōngbak</i> (Lee Myung-bak)	<i>kyōngjerūl</i> (the economy)	<i>ap'at'ū</i> (apartment)
2	<i>kyoyugūi</i> (educational)	<i>chinboūi</i> (progressive)	<i>nomuhyōn</i> (Roh Moo-hyun)	<i>miyuge</i> (to/in the US)	<i>kyōngjega</i> (the economy)	<i>nyut'aun</i> (New Town)	<i>habū'igwōn</i> (right to negotiate)	<i>hyōksindoshi</i> (Hyuk-shin Dosi (Innovation City))	<i>sogagi</i> (beef)	<i>mujogōn</i> (unconditionally)
3	<i>hakkyoesō</i> (at school)	<i>pūrusella</i> (Brucellosis)	<i>munhwalibo</i> (Munwha Ilbo)	<i>chōgōdō</i> (at least)	<i>hanbando</i> (Korean Peninsula)	<i>osehun</i> (Oh Se-hoon)	<i>hoesanūn</i> (a company)	<i>hongjaehui</i> (Hong Jae-hee)	<i>ap'ūro</i> (in the future)	<i>ihanjōng</i> (Lee Hang-jung)
4	<i>chayurhwa</i> (school privatization)	<i>soriga</i> (sound)	<i>kūdūrii</i> (their)	<i>imyōngbak</i> (Lee Myung-bak)	<i>mandūrō</i> (being made)	<i>mun'guk'yōn</i> (Mun Gook-hyun)	<i>chahabwōn</i> (union member)	<i>kaunde</i> (in the middle of)	<i>kungmindūl</i> (the people)	<i>chōnhwarūl</i> (a phone call)
5	<i>taet'ōngnyōng</i> (President)	<i>sorirūl</i> (sound)	<i>chōdūrii</i> (of those people)	<i>yōlshimhi</i> (diligently)	<i>tansunhi</i> (simply)	<i>mun'guk'yōnūi</i> (Mun Gook-hyun's)	<i>kanūnghan</i> (possible)	<i>sajangūi</i> (the owner's)	<i>marhanūn</i> (speaking)	<i>chōnhwaga</i> (a call)
6	<i>konggyoyugūi</i> (of public education)	<i>haengsajang</i> (event site)	<i>ūiwōnim</i> (dear representative of the Korean National Assembly)	<i>chejudo</i> (Jeju Island)	<i>munjenūn</i> (problem)	<i>kuk'oeūiwōn</i> (Representative of the Korean National Assembly)	<i>hoesaga</i> (a company)	<i>chosōnilbo</i> (Chosun Ilbo)	<i>moūgo</i> (without knowing)	<i>saenggak'amnida</i> (think)
7	<i>haksae'ngdūl</i> (students)	<i>ch'amūro</i> (truly)	<i>kūdūrii</i> (they)	<i>oegugin</i> (foreigner)	<i>palk'yōtta</i> (revealed)	<i>min'judangūi</i> (the Democratic Party's)	<i>ilbangōkūro</i> (unilaterally)	<i>pangsanghun</i> (Bang Sang-hoon)	<i>chōngsangjōgin</i> (normal)	<i>pomnida</i> (regard)
8	<i>haksae'ngdūrii</i> (students)	<i>ōpsōtta</i> (was missing)	<i>kundaerūl</i> (the military)	<i>min'yōnghwa</i> (privatization)	<i>manmyōngi</i> (ten-thousands of people)	<i>chōngch'ijōk</i> (political)	<i>nodongjohabe</i> (to a labor union)	<i>chosōnsasōrūn</i> (Chosun Ilbo's editorial)	<i>moūnūn</i> (not knowing)	<i>yushimin</i> (Rhyu Si-min)
9	<i>aidūl</i> (school children)	<i>kamsahamnida</i> (thank you)	<i>moduga</i> (everyone)	<i>chōngch'irūl</i> (politics)	<i>chunggane</i> (amid)	<i>araero</i> (below)	<i>nodongbōpkwa</i> (with the labor law)	<i>chujanghago</i> (claiming)	<i>migwangubyōngwihōminja</i> (Risk materials of Mad Cow Disease in US beef products)	<i>KTF</i> (KT Freetel Co., Ltd)
10	<i>charhanūn</i> (competent)	<i>shijak'aetta</i> (started)	<i>haengt'ae</i> (at such deeds)	<i>urūi</i> (our)	<i>chōnggwōnūl</i> (the regime)	<i>maegungno</i> (traitor)	<i>nodongjohap</i> (labor union)	<i>chōngch'irūl</i> (politics)	<i>Tommanūn</i> (wealthy)	<i>ajikro</i> (still)
11	<i>haksae'ngdūrii</i> (students)	<i>pūrusellae</i> (Brucellosis)	<i>changjōmūl</i> (merits)	<i>urinūn</i> (we)	<i>hongjūnp'yōnūn</i> (Hong Jun-pyo)	<i>mushik'anja</i> (fool)	<i>paramnida</i> (wish to)	<i>chiyōgūi</i> (local)	<i>chugigo</i> (killing)	<i>munjega</i> (problem)

**Table 2** Top 15 Terms of the Topic Modeling Result of the April 16 Collection (Continued)

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
School Education	Avian Flu	RMH-Munhwa	Privatization	Economy	Newtown	Allianz Union	CEEditorial	Beef	Housing, brib
12 <i>chinjŏngha</i> (sincere)	<i>aika</i> (avian influenza)	<i>hongjunp'yŏi</i> (Hong Jun-pyo's)	<i>kudŭrŭi</i> (their)	<i>miwŏhanŭn</i> (hate)	<i>ilbonapchaebi</i> (pro-Japanese collaborator)	<i>chijŏnjang</i> (local branch chief)	<i>sudogwŏn</i> (Seoul metropolitan area)	<i>kŏn'gangŭl</i> (health)	<i>ŭmurŭl</i> (obligation)
13 <i>sŭsŭro</i> (independently)	<i>hwakshirihi</i> (surely)	<i>chŏn'gyŏnghanŭn</i> (respecting)	<i>saramdŭrŭl</i> (people)	<i>ap'esŏ</i> (in front of)	<i>miguksadaejŭjja</i> (Pro-US figures)	<i>nodongbuŭi</i> (Ministry of Labor)	<i>chŏnggwŏnŭi</i> (the regime)	<i>pujaga</i> (the rich)	<i>kŭrŏgo</i> (in addition)
14 <i>chungyohan</i> (important)	<i>p'riyohan</i> (needed)	<i>sumanŭn</i> (numerous)	<i>unharŭl</i> (Canal)	<i>ttarŭmyŏn</i> (according to)	<i>marhaetta</i> (said)	<i>sŏnggwagŭpche</i> (performance-based pay system)	<i>konggiŏp</i> (public company)	<i>shipsŭmnida</i> (want to)	<i>chŏege</i> (to me)
15 <i>pŏbinse</i> (corporation tax)	<i>chigwŏni</i> (staffer)	<i>paramnida</i> (wish to)	<i>munjenŭn</i> (problem)	<i>kaemando</i> (lowly)	<i>tangsŏnja</i> (the elected)	<i>nodongjaŭi</i> (worker's)	<i>poinda</i> (seen)	<i>chedaero</i> (thoroughly)	<i>yŏnhamnyusŭ</i> (Yonhap News)



**Table 3** Top 15 Terms of the Topic Modeling Result of the May 2 Collection

Topic 1 Benefits	Topic 2 Beef privatization	Topic 3 mcd.netizen	Topic 4 mcd	Topic 5 Our	Topic 6 Lit.candles	Topic 7 beefimport	Topic 8 Dokdo	Topic 9 Government	Topic 10 mcd.citizens
1 <i>idŭkkwa</i> (profit on earth)	<i>sogagi</i> (beef)	<i>kwangubyōng</i> (mad cow disease)	<i>sogagi</i> (beef)	<i>uriŭi</i> (our)	<i>neibō</i> (Naver)	<i>sogagi</i> (beef)	<i>yōksagogyasō</i> (history textbook)	<i>shikkaji</i> (by when)	<i>uriga</i> (we)
2 <i>todaech'e</i> (what on earth)	<i>chinjunggwŏn</i> (Chin Jung-kwon)	<i>sogagi</i> (beef)	<i>miguksan</i> (imported from the USA)	<i>uriga</i> (we)	<i>myōngbakchiok</i> (Hell with Myung-bak)	<i>chochi'riŭl</i> (actions)	<i>tonghaep'yogi</i> (labeling the East Sea)	<i>sonhakkyu</i> (Sohn Hak-kyu)	<i>solchik'i</i> (frankly)
3 <i>ōmch'ōngnan</i> (immense)	<i>sahoja</i> (moderator)	<i>miguksan</i> (imported from the USA)	<i>kwangubyōng</i> (mad cow disease)	<i>chashinŭi</i> (one's)	<i>t'anhaekch'ōn'guk</i> (Heaven after Impeachment)	<i>srn</i> (specified risk material)	<i>kwagōsa</i> (past history)	<i>saeroun</i> (new)	<i>kwangubyōng</i> (mad cow disease)
4 <i>kaemand</i> (deplorable)	<i>mindongsōk</i> (Min Dong-seok)	<i>p'uraion</i> (prion)	<i>fta</i> (Free Trade Agreement)	<i>saenggak'amnida</i> (think)	<i>sait'ŭ</i> (website)	<i>kaewŏllyōng</i> (months old)	<i>yōngŏra</i> (in English)	<i>sahoerŭl</i> (moderating)	<i>miguksa</i> (US beef)
5 <i>chadŭri</i> (people [derogatory])	<i>ansŭmmikka</i> (isn't it)	<i>tongmulsōng</i> (animal based)	<i>embissi</i> (Munhwa Broadcasting Corporation)	<i>kungmindŭri</i> (the people)	<i>ch'ōnggyecheon</i> (Cheong-gye cheon)	<i>kwangubyōng</i> (mad cow disease)	<i>ilbonŭi</i> (Japanese)	<i>sogagi</i> (beef)	<i>sogagi</i> (beef)
6 <i>kyōngjeŏk</i> (economic)	<i>modudŭl</i> (all)	<i>sanyoŭl</i> (feed)	<i>marhaetta</i> (spoke)	<i>urinŭn</i> (we)	<i>paramnida</i> (wish)	<i>subŭl</i> (importation)	<i>kongnuwŏn</i> (civil officials)	<i>onŭllarŭi</i> (current)	<i>Nongnim haeyang susan wiwŏnhoe</i> (Committee on Agriculture, Forestry, and Fisheries)
7 <i>angnarhan</i> (helinous)	<i>mōkūmyōn</i> (by eating)	<i>mad</i> (mad)	<i>hannaradang</i> (Grand National Party)	<i>urinara</i> (our country)	<i>petition</i> (petition)	<i>chep'umŭl</i> (products)	<i>kimjōnggu</i> (Kim Jeong-gu)	<i>kanŭngsōngi</i> (possibility)	<i>korean</i> (Korean)
8 <i>inhaesō</i> (caused by)	<i>mič'inso</i> (crazy cow)	<i>qjd</i> (Creutzfeldt Jacob disease)	<i>imyōngbak</i> (Lee Myung-bak)	<i>nararŭl</i> (country)	<i>agora</i> (Agora)	<i>miguge</i> (to the USA)	<i>imyōngbak</i> (Lee Myung-bak)	<i>hannaradang</i> (Grand National Party)	<i>report</i> (report)
9 <i>chunggane</i> (amid)	<i>kungmindŭl</i> (the people)	<i>isangŭi</i> (more than)	<i>palk'yōtta</i> (announced)	<i>kyōngjerŭl</i> (the economy)	<i>chōngch'aek pandae shiwi'yōndae</i> (OUR Korea)	<i>miguksan</i> (imported from the USA)	<i>homp'eiji</i> (homepage)	<i>kūngjōngjōgin</i> (positive)	<i>abuse</i> (abuse)

**Table 3** Top 15 Terms of the Topic Modeling Result of the May 2 Collection (Continued)

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
Benefits	Beef privatization	mcd.netizen	mcd	Our	Lit.candles	beefimport	Dokdo	Government	mcd.citizens
10 <i>nôhûdûri</i> (you folks)	<i>taeunha</i> (Grand Canal Project)	<i>tanbaekchil</i> (protein)	<i>disease</i> (disease)	<i>yôrbun</i> (you [plural])	<i>int'ônet</i> (internet)	<i>migukchôngbunûn</i> (US government)	<i>sôngsuk'an</i> (mature)	<i>ch'arari</i> (rather than)	<i>posted</i> (posted)
11 <i>palsaengdoenûn</i> (emergent)	<i>kat'ayo</i> (seems like)	<i>net'jûnûn</i> (netizen)	<i>pdsuch'ôp</i> (PD's Note)	<i>kiûrûn</i> (they)	<i>chip'oe</i> (at/to protest)	<i>kiyejôk</i> (mechanical)	<i>hwangusôk</i> (Hwang Woo-seok)	<i>nomuhyôn</i> (Roh Moo-hyun)	<i>april</i> (April)
12 <i>mandûlgo</i> (making)	<i>miguksan</i> (made in USA)	<i>vcjd</i> (variant cid)	<i>han'gugini</i> (Korean)	<i>sûsûro</i> (by oneself)	<i>chehandoen</i> (limited)	<i>saengsandoen</i> (produced)	<i>naraûi</i> (country's)	<i>ibun</i> (this person)	<i>mad</i> (mad)
13 <i>yôjadûlgwa</i> (with women)	<i>ponikka</i> (looking at)	<i>roir'ô</i> (Reuter)	<i>ansangsu</i> (Ahn Sang-su)	<i>saenggagi</i> (thought)	<i>hwa'it'ing</i> (Cheer up)	<i>nongnimbunûn</i> (Ministry of Agriculture and Forestry)	<i>chuihan'guktaesagwan</i> (Embassy of the Republic of Korea in Japan)	<i>sig</i> (sig)	<i>nidûri</i> (you folks)
14 <i>aktok'an</i> (vicious)	<i>minyônghwa</i> (privatization)	<i>soege</i> (to the cow)	<i>kaunde</i> (midst)	<i>naraûi</i> (country's)	<i>soragwangjang</i> (Sora Square)	<i>kangnyôkan</i> (strong)	<i>toktonûn</i> (Dokdo Island)	<i>migugi</i> (the United States)	<i>urinara</i> (our country)
15 <i>ikûl</i> (interest)	<i>urinara</i> (our country)	<i>môgôdo</i> (even when eaten)	<i>chosônllbo</i> (Chsun Ilbo)	<i>kiûrûi</i> (their)	<i>meine</i> (on the main page)	<i>yôn'guso</i> (research institute)	<i>hompe'ijiesô</i> (on homepage)	<i>ch'agidaet'ongnyông</i> (next president)	<i>chilmune</i> (to a question)

activities in an abstract manner. Topic *our* is a list of terms that divides society into two classes: one group (we, our, us, and citizens) against the other (they, their, them, and the President) with a trace of the beef trade issue with the term “produced/made in the United States. Topic *lit.candles*, which appears as an isolate,<sup>11</sup> concerns online communities whose members proposed street rallies on the beef trade issue, encouraging participation in the candlelight protests in downtown Seoul and signing Andante’s petition.<sup>12</sup>

It is noteworthy that the two-pronged relationship between the cohesive community largely focusing on the beef trade issue (the green community in Fig. 1(b)) and the rest of the communities urges interpretive work. It is counterintuitive that the community containing topics *Dokdo* and *our* are connected to the beef community on the beef trade issue, whereas topic *lit.candles* appears as an isolate. Topic *Dokdo* shares the term “Lee Myung-bak” with topic *mcd*, whereas topic *our* has in common (1) the terms “our country” and “the people” with topic *beef.privatization*, and (2) the terms “our country” and “we” with topic *mcd.citizens*. In the sense that topic *Dokdo* addresses international relations with Japan, the connotation of the shared term (“Lee Myung-bak”) is slightly different from its contexts in topic *mcd*. In contrast, topic *lit.candles* whose constitutive terms portray specific logistics of the upcoming street rallies and list the names of leading online communities remains isolate without sharing the most shared terms “mad cow disease” or “beef.” Nevertheless, it is hard to judge topic *lit.candles* has no contribution to claim-making. It implies that the structure of topic networks requires careful interpretation regarding the semantic connections between topics.

### Semantic communities in dynamics

Despite the stark contrast between the two juxtaposed topic networks at the start and end points of the studied period, the above discussion lacks dynamic accounts. Figure 2 charts variations in the modularity scores for the two versions of topic modeling, which differ only in the number of extracted terms (15 terms in solid blue line vs. 30 terms in red dotted line). Modularity is often used to detect cohesive subgroups or communities that are strongly connected nodes in comparison with their connections to other nodes by measuring the proportion of the strength of intra-group connections to that of inter-group connections as a single point score for a given network as a whole. Higher modularity indicates the presence of strong cohesive subgroups whose inter-community connections are less salient, whereas low modularity implies fewer cohesive communities or stronger connections among distinguishable communities.

Figure 2 illustrates that modularity scores fluctuate rather than move in a consistent way over time. This implies that Agora users did not discuss the protest claims of the Candlelight Protests in a linear way, under which semantic communities grow consistently. Although Fig. 2 portrays the dynamics for this fluctuation, it does not reveal the definitive cause of the fluctuation, for which multiple scenarios can be hypothesized.

<sup>11</sup>When the May 2 topic network is drawn with top 30 terms, topic *lit.candles* is connected to the cohesive community by sharing the terms such as “netizens.”

<sup>12</sup>Andante, an Agora user whose actual identity as a high school student, was revealed to the public later, wrote an online petition on the petition section of Agora on April 6, 2008, asking to push the National Assembly to impeach President Lee for his “already unacceptably irresponsible” governance for the two months after inauguration. Although the petition did not gain popularity at the moment of its writing, it eventually received more than a million signatories soon after first two to three candlelight protests in early May.

To begin, quantitative changes in modularity scores require careful inter-community and intra-community investigations, because the changes can result from the two separately or in combination. It is equally important to appreciate modularity in more interpretative ways. For instance, a topic network that contains a growing semantic community regarding the beef trade issue over time can also contain isolated topics, whose semantic relevance often defies explicit surface evaluation, as in the May 2 network. Also, it is possible that protest claims appear in multiple connected semantic communities rather than appear together in a single community, as the same issues are addressed in slightly different sets of terms. In other words, the topic networks and their semantic communities require interpretative procedures as well. In the following two sections, I discuss topic networks with low and high modularity as presented in Figs. 3 and 4, respectively, with a focus on resilience in the patterns of connections.

#### ***Low modularity: different terms connect different topics***

In Fig. 3, the common trait across all four topic networks, April 19, 21, 26, and 30,<sup>13</sup> is that they have either more red lines (linking different topical communities) or thicker red lines despite the presence of thicker black lines. A noticeable phenomenon is that different terms tend to connect different topics, although some terms iteratively appear over time. This feature is evident in the April 26 network (Fig. 3(c)). There are eight communities, only two of which contain two topics: one community includes topics *beef.products* and *beef.import*, and the other consists of topics *government.the.nation* and *policies*. The first community is linked by the term of “mad cow disease,” whereas the second is linked by “we.” The term “mad cow disease” also appears in topic *beef.ingredients*, which is also connected to topic *government.the.nation* and topic *policies*. Topic *impeachment*, which mentions that the online community ANTI-MB demands the impeachment of Lee from presidential office, is connected to topic *government.the.nation* by the term “Lee Myung-bak.” These communities tend to be tied to one another through different terms.

In the topic networks for April 19 and 21, it is observed that the terms “mad cow disease,” “beef,” “Lee Myung-bak,” derivatives of first-person plural pronouns, and “produced/made in the United States,” connect different topics. The April 21 topic network has three communities. The biggest community has five topics: *CSEditorial* (commenting on the conservative newspaper Chosun Ilbo’s editorials outlining its perspective on Korea’s diplomatic approaches towards the U.S. and North Korea), *strategic.alliance* (shifts in Korea’s military alliances), *beef.inspection* (inspection issues surrounding the importation of U.S. beef not through bilateral trade but under regulation by international organizations), *beef.policy* (in relation to health insurance), and *U.S.Korea* (concerning the opening of the Korean market to the United States through a Free Trade Agreement). Among these, topics *CSEditorial* and *beef.policy* have no shared terms, though three topics, *beef.inspection*, *beef.policy*, and *U.S.Korea* are connected by the terms “beef” and “mad cow disease.” Topic *contagion* is connected to topics in the green community by sharing the term “how much” with topics *beef.policy* and *U.S.Korea*, whose meaning remains relatively rhetorical in each topic.

<sup>13</sup>See Figure 1(b) for the May 2 topic network.

In this regard, the April 30 network deserves discussion at length for the appearance of a relatively cohesive community (in the blue community in Fig. 3(d)). The four topics, *beef*, *mcd.the.nation*, *mcd.fta*, and *mcd.import*, are connected by the three terms of “beef,” “mad cow disease,” and “produced/made in the US.” These three terms build consistency and coherence concerning the beef trade issue by contextualizing beef importation in relation to the risk of mad cow disease from U.S. beef products and the ongoing negotiations over the Korea-U.S. Free Trade Agreement. The green community contains topic *mcd.symptoms* that delineate the graveness of the vCJD threat in reference to a TV show that aired on April 29, alongside two other topics, *Agora* and *pro.japanese*, which also appear to be irrelevant to the beef trade issue. Topic *Agora* carries terms such as “you,” “Grand Canal Project,” “petition,” “Naver,” and “Agora,” which seemingly appear to be unrelated to the beef trade issue at face value. However, the links interestingly adumbrate the field of public discussion on the beef trade issue in the digital platforms of Naver and Agora, where the former’s search engine was under suspicion of technical manipulation to hide increasing public interest in the beef trade issue. In contrast, topic *pro.japanese* contains plural pronouns such as “their,” alongside terms such as “Japanese,” “now,” and “achievements and drawbacks.”<sup>14</sup> In other words, topic *pro.japanese* neither carries nor shares terms such as “U.S. beef” or “mad cow disease” at all. Its connection to other topic communities is that the topic itself is set up in the dueling and boundary-making languages of “we” (indicating Korea or Koreans) and “they” (indicating Japan or Japanese). In this sense, this part of the connection should not be overemphasized in order to avoid misinterpretation of the pronouns that obviously serve different social and political contexts.

In sum, topic networks with low modularity scores can be summarized as follows. They tend to have many communities because their member topics share different terms with different topics. Simultaneously, some topics such as beef trade-related topics focusing on various aspects, such as mad cow disease, its human variant CJD, KOR-US economic and military aspects, domestic policies proposed by President Lee since his presidential election campaign, tend to appear consistently and grow into cohesive topic groups over time as demonstrated in Fig. 3.

#### **High modularity: cohesive topical communities as claim**

Figure 4 presents four topic networks with modularity scores higher than 0.2. A common pattern here is that topics are more coherently connected by sharing multiple terms, which appear continuously. In the April 24 network, the beef trade issue is presented in connection to how the newly-proposed inspection process would fail to control mad cow disease. In it, terms such as “beef,” “mad cow disease,” and “produced in the US” all connect the three topics *mcd.prion*, *mcd.policy*, and *beef.inspection*. In the April 27 network, the beef trade issue is closely related to the context of mad cow disease and vCJD. Topics *mcd.privatization* (regarding mad cow disease alongside health insurance and the Grand Canal project), *mcd.vcjd* (regarding mad cow disease in cattle and vCJD in humans), and *beef.origin* (regarding the origins of beef products), which are clustered as a community, share the terms “Lee Myung-bak,” “beef,” and “mad cow

<sup>14</sup>The same topic *pro.japanese* also includes terms such as “pro-Japanese traitors,” “history,” “pro-Japanese,” “apologies,” “the Korean Peninsula,” etc. in the outcome of its 30-term topic modeling.

disease.” Topic *mcd.vcjd* particularly contains more terms that make it as a brokering center with another community in red in Fig. 4(b). The term “we” in the topic is shared with topics *C. Olympic* (ethnic oppression in China before the 2008 Summer Olympics) and *civil.society* (concerning various social activism in abstract terms), and the term “problem” is shared with topics *press.nuclear.weapons* (regarding reports on the North Korean nuclear weapons program) and *LMG.policies* (critical remarks on policies in abstract terms). Compared to the green community, the red community contains relatively independent topics, whose connecting terms are less substantial.

In the April 28 network, the four terms—“beef,” “mad cow disease,” “we,” “our country,” and “produced in the US”—builds a tightly-connected community whose members are topics *US.beef*, *mcd.feed*, and *mcd.patients*. The May 1 topic network has three strongly-connected topical communities: one mainly focuses on the risk of importing U.S. beef at a quality that poses the risk of a lethal disease, vCJD; one on relevant state-level offices such as topic *theAFF* (carrying terms “Grand Canal Project,” “mad-cow cattle,” “president,” “the Ministry of Agriculture, Forestry, and Fisheries”), *president* (“president,” “our,” “economy,” “people,” “Naver”), and *thenation* (“country,” “citizens,” “you,” etc.); the third topic community on topics *mcd.fta*, *rallies* (“political,” “our,” “election regulations,” “your,” “social,” etc.), and *OURKOREA* (“protest,” “Our Korea” (an online community that organized initial protests), “vCJD,” “Chosun Ilbo” (popular Korean conservative newspaper). To conclude, these four high-modularity networks present the protest claims of the Candlelight Protests through topics that share the same major terms consistently. Topics that are not clustered into the protest claims, however, tend to share abstract terms such as “problems” or “we,” whose indicative meanings can vary drastically.

In addition to the specified brokering terms in the high-modularity topic networks, a noteworthy point regardless of modularity is that, from April 30 to May 2, all semantic networks contain more explicit terms showing the growing involvement of digital interactions across online communities and web portal sites in the ongoing debate on the beef trade issue. Topic *Agora* in the April 30 semantic network suggests that some Agora users began to proactively persuade people to sign an online petition to demand Lee’s impeachment, a demand which is also briefly mentioned in topic *mcd.import*. On May 1, topic *OURKOREA* refers to legislation on collective action and protests that prohibit protests from being held in public spaces at night,<sup>15</sup> while notifying that potential rally participants would need to bring candles. It is worth noting that expressions denoting group boundaries such as “we” versus “they,” “the nation,” “citizens,” and “netizens” appear continually, and tend to appear as part of the brokering terms over time.

In a similar vein, topic *rallies* in the May 1 network contains terms such as “rallies” and “protests,” and on the day of the first street rally that took place on May 2, topic *lit.candles* carried the following terms: “Naver,” “Hell with Myungbak,” “Heaven after impeachment,” “Cheonggyecheon,”<sup>16</sup> “petition,” “Agora,” “Our Korea,”<sup>17</sup> “internet.” In particular, the slogans “Hell with Myungbak” and “Heaven after impeachment” were

<sup>15</sup>The Constitutional Court of Korea has since ruled this law partially unconstitutional as of 2010.

<sup>16</sup>A stream running through downtown Seoul. Candlelight protesters started their events in a public square by this river, called Sora Square.

<sup>17</sup>Online community launched to stage street rallies against beef imports.



often used to encourage public participation in the impeachment petition posted to Agora. Also, the online communities Our Korea and ANTI-MB were the most active in staging the street event on May 2, although the political differences between the two soon resulted in an online dispute. On the day, people also gathered with lit candles at Sora Square, a public space constructed alongside Cheonggyecheon Stream, in a gathering titled “A Cultural Event with Lit Candles,” which was so named in order to avoid the police intervention that would follow an overtly political rally. This also signals that “netizens” in particular became prominent in digital interactions over the beef trade issue with the abstract concept of citizens or the nation.

Some of the semantic and relational traits described above can be summarized as follows. First, the beef trade issue did not appear on Agora as a clean-cut claim in its definitive form from the outset. Rather, it was constituted through repeated arrangement and rearrangement in relation to other terms and topics. One of the topics that consistently appear throughout the all daily collections is *privatization*. In the April 16 network, it is one of the major topics used in discussions in reference to beef importation. Unlike the U.S. healthcare system, Korea has developed a universal health insurance (UHI) system while allowing private insurance companies.

In 2008, President Lee alluded to a policy to allow for-profit hospitals that can refuse patients with universal health insurance. The rationale for this policy mainly focused on the expected economic benefits that for-profit hospitals would create in the future. Arguments against for-profit hospitals and private health insurance had been a mainstay on Agora before the beef negotiations took place, and Agora users linked beef trade-related news to their concerns over the privatization of the healthcare industry. This rhetorical connection became stronger and more consistent to the extent that Lee’s policies were predicted to result in a sequence of unpleasant events, namely that lower-income classes of the Korean public would become more likely to purchase cheaper U.S. beef products, thereby causing the spread of vCJD since the new inspection standard would allow SRM to enter the Korean food supply system. Such contexts raised by topics related to privatization are overshadowed due to the growing prevalence of mad cow disease-related topics, although they return in the May 2 network. In other words, the claim-making process occurs concurrently with the process of context-building that inscribes multiple layers into the claim of resistance against U.S. beef, which offers specific styles and tones in discussing the beef trade issue, regardless of the extent of outward visibility.

Likewise, people whose primary interests are not the beef trade issue continued to link their own agenda to the beef issue by addressing the two topics together and using terms such as “Lee Myung-bak” or plural pronouns. In this regard, the publication of *Who’s Who: Traitors during the Japanese Colonial Period*, which was published by the Presidential Committee on the Settlement of Public Past Affairs, is rarely mentioned as a potential influence on the Candlelight Protests. The beef trade issue is neither topically nor temporally connected in an intuitive sense to the settlement of social conflicts surrounding the traces of Japanese colonialism. Nevertheless, it appeared repeatedly during the 2008 Protests, as well as in the topic networks used in this study. Agora users later published an anthology entitled *Agora: The Encyclopedia of Republic of Korea*, in the third month of the Candlelight Protests by compiling selections of digital posts announced with their own timeline and points. According to the book, online

communities that did not take up the beef trade issue as their primary agenda nonetheless both participated enthusiastically in the Candlelight Protests and obtained publicity.

Second, there are many new issues that appear in topic networks, though the degree of their repeated appearance over time varies, as well as how those topics connected to other beef-related topics. An apt example is the aforementioned April 27 topic network, while other topic networks with lower modularity scores also present similar features. The continual appearance of new issues across all topic networks results from the characteristics of the studied platform, which was never devoted to a specific topic or issue. It should be noted that the effect of new topics can be more salient in high-modularity networks as their protest claims form a closely-connected community, compared to low-modularity networks where similar cohesive communities are rare. In this regard, it is notable that topic *beef* in the April 16 network was one of the topics that may have been buried quickly, unless it was repeatedly discussed in relation to other topics.

## Discussion and conclusion

The main finding of this study is that protest claims are shaped through communication rather than appear in its fully-developed form. While being formed, protest claims are contextualized in relation to other issues. Relationality in the topic networks does not originate from the arbitrariness of connected topics. The topics that obtain semantic relevance tend to be connected in more consistent ways over time through a set of brokering terms that directly denote the beef trade issue. Also, consistency does not imply a gradual and linear increase in the degree of coherence of protest claims. The reconstructed topic networks demonstrate that their semantic structures undergo changes through which some layers of protest claims become reinforced and incorporate new layers.

This study yields a few implications in the development of semantic analysis, whose importance will grow for both SNA and SMS in the era of digital media. First, it is evident that semantic network analysis of online posts requires rigorous text pre-processing. A critical point is seeking balance between generic suggestions that aim to reduce noise in text materials and specific interests pursued by a given study. As discussed above at length, a fundamental challenge specific to this study has been that the topics and terms to be removed remained unknown or needed to be selected, given that this study intended to trace how relevance emerges through ongoing communication. This conflicting situation invites further discussions to enhance the validity of similar research. Noise, when used as a technical term, must be carefully understood in the theoretical sense. At the earliest stage of text pre-processing, particularly for Korean-language texts, tagging the part-of-speech of terms or recoding terms into N-gram terms can be helpful, but it can introduce unwanted complexities into data. It is equally important to reflect the characteristics of digital posts and online communication that involves many people with various viewpoints and styles of speech, whereby the task of reducing noise entails a challenge in and of itself. With regard to this matter, the accumulation of future empirical research will be helpful.

By extension, it is important to develop more rigorous classificatory systems that enrich quantitative and qualitative analysis simultaneously. This study has employed

multiple classificatory schemes by clustering co-occurring terms into topics, which were subsequently grouped into semantic communities, whose structural characteristics, i.e., modularity, were ultimately used to divide topic networks into two groups. As discussed above, inter-topic relationships and inter-community relationships captured in quantitative notions do not complete analysis regarding what topics are connected and what terms tend to be widely used beyond topical boundaries. To investigate such subtleties in dynamics, it is necessary to conduct interpretive analysis, aptly informed by semantic network analysis. Improvements can be made in multiple dimensions. To begin, temporarily can be considered in a more elaborate way. In addition, it also should be noted that this study has not yet benefited from other attributes of digital posts such as the attributes of authors, as their motivations could have varied both because of their personal preferences that are largely unknown and because of growing focus on the Candlelight Protests over time.

Avoiding the use of protest-related terms as a primary reference to reconstruct communication networks on large-scale protests, this study has sought to identify how terms and issues constitute protest claims. To trace the process of claim-making, the analysis relied on posts available on a digital platform. In part, this decision consulted the environment of online communication, where multiple types of platforms prevailed in 2008 in South Korea. But more importantly, this study presupposed that digital platforms may have distinctive impacts on the ways in which the similar issues are discussed. Drawing on this point, it would be intriguing to study what patterns of topic networks were under construction in other digital platforms during the same pre-protest period. This follow-up research would contribute to more general questions such as how ideas diffuse when their spread is likely to carry subtle semantic connotations from their originating place, despite some commonalities.

#### Abbreviations

BSE: Bovine Spongiform Encephalopathy; FDB: Free Discussion Board; SMO: Social Movement Organization; SMS: Social Movement Studies; SNA: Social Network Analysis; SRM: Specified Risk Material; vCJD: variant Creutzfeldt-Jacob disease

#### Acknowledgements

I appreciate the time and advice from the reviewers of this manuscript. I would also like to thank Paul McLean, Ann Mische, Judith Gerson, Richard Williams, John Krinsky, and Misty Dawn Ring-Ramirez for their insightful suggestions. Earlier drafts were presented at multiple conferences, including the Annual Conference of American Sociological Association of 2018, and the Annual Conference of International Network Social Network Analysis of 2016.

#### Authors' contributions

The author read and approved the final manuscript.

#### Author's information

Eunkyung Song received her PhD degree in sociology from Rutgers University. Her research interests include social network analysis, social movements, cultural sociology, and computational social science.

#### Availability of data and materials

This manuscript was developed from the author's Ph.D. dissertation. Part of the data used in this paper can be provided upon request.

#### Competing interests

The author declares that she has no competing interests.

Received: 16 November 2018 Accepted: 30 July 2019

Published online: 23 August 2019

#### References

- Bearman P, Stovel K (2000) Becoming a Nazi: a model for narrative networks. *Poetics* 27:69–90
- Bennett WL, Segerberg A (2013) *The logic of connective action: digital media and the personalization of contentious politics*. Cambridge University Press, New York
- Blee KM (2012) *Democracy in the making: how activist groups form*. Oxford University Press, New York

- Blei DM (2012) Probabilistic topic models. *Commun ACM* 55:77–84. <https://doi.org/10.1145/2133806.2133826>
- Blei DM, McAuliffe JD (2010) Supervised topic models, pp 1–22
- Carley K (1994) Extracting culture through textual analysis. *Poetics* 22:291–312
- Carley KM, Kaufer DS (1993) Semantic connectivity: an approach for analyzing symbols in semantic networks. *Commun Theory* 3:183–213
- Castells M (2012) *Networks of outrage and Hope: social movements in the internet age*. Polity Press, Malden
- Diani M (2004) Networks and participation. In: Snow DA, Soule SA, Kriesi H (eds) *The Blackwell companion to social movements*. Blackwell Publishing, Malden, pp 340–359
- Diani M, McAdam D (eds) (2003) *Social movements and networks: relational approaches to collective action*. Oxford University Press, New York
- DiMaggio P (2015) Adapting computational text analysis to social science (and vice versa). *Big Data Soc* 2 (July–December):1–5
- DiMaggio P, Nag M, Blei D (2013) Exploiting affinities between topic modeling and the sociological perspective on culture: application to newspaper coverage of U.S. government arts funding. *Poetics* 41:570–606. <https://doi.org/10.1016/j.poetic.2013.08.004>
- Earl J, Kimport K (2011) *Digitally enabled social change: activism in the internet age*. The MIT Press, Cambridge and London
- Fernandez RM, McAdam D (1989) Multiorganizational fields and recruitment to social movements. *Int Soc Mov Res* 2:315–343
- Fuhse JA (2015) Networks from communication. *Eur J Soc Theory* 18:39–59. <https://doi.org/10.1177/1368431014549683>
- Gerbaudo P (2012) *Tweets and the streets: social media and contemporary activism*. Pluto Press, London
- Gerbaudo P, Treré E (2015) In search of the ‘we’ of social media activism: introduction to the special issue on social media and protest identities. *Inf Commun Soc* 18:865–871. <https://doi.org/10.1080/1369118X.2015.1043319>
- González-Bailón S, Wang N (2016) Networked discontent: the anatomy of protest campaigns in social media. *Soc Networks* 44:95–104. <https://doi.org/10.1016/j.socnet.2015.07.003>
- Goodwin J, Jasper JM (eds) (2004) *Rethinking social movements: structure, meaning, and emotion*. Rowman & Littlefield Publishers, Inc, Lanham
- Gould RV (1993) Collective action and network structure. *Am Sociol Rev* 58:182–196. <https://doi.org/10.2307/2095965>
- Gould RV (1995) *Insurgent identities: class, community, and protest in Paris from 1848 to the commune*. The University of Chicago Press, Chicago and London
- Gould RV (2003) Why do networks matter. In: *Social movements and networks: relational approaches to collective action*. Oxford University Press, New York and Oxford, pp 233–257
- Granovetter M (1978) Threshold models of collective behavior. *Am J Sociol* 83:1420–1443
- Johnston H (1995) A methodology for frame analysis: from discourse to cognitive schemata. In: Johnston H, Klandermans B (eds) *Social movements and culture*. University of Minnesota Press, Minneapolis, pp 217–246
- Krinsky J, Mische A (2013) Formations and formalisms: Charles Tilly and the paradox of the actor. *Annu Rev Sociol* 39:1–26
- Marwell G, Oliver P (1993) *The critical mass in collective action: a micro-social theory*. Cambridge University Press, New York and Cambridge
- McAdam D (1988) Micromobilization contexts and recruitment to activism. *Int Soc Mov Res* 1:125–154
- McAdam D (2003) Beyond structural analysis: toward a more dynamic understanding of social movements. In: Diani M, McAdam D (eds) *Social movements and networks: relational approaches to collective action*. Oxford University Press, New York, pp 281–298
- McAdam D, Sewell WH (2001) It’s about time: temporality in the study of social movements and revolutions. *Silence and Voice in the Study of Contentious Politics*, In, pp 89–125
- McAdam D, Tarrow S, Tilly C (2001) *Dynamics of contention*. Cambridge University Press, New York
- McFarland D, McFarland HR (2015) Big data and the danger of being precisely inaccurate. *Big Data Soc* 2(July–December):1–4
- McLean P (2017) *Culture in networks*. Polity, London and Malden
- Mische A (2003) Cross-talk in movements: reconceiving the culture-network link. In: Diani M, McAdam D (eds) *Social movements and networks: relational approaches to collective action*. Oxford University Press, New York, pp 258–280
- Mische A (2008) *Partisan publics: communication and contention across Brazilian youth activist networks*. Princeton University Press, Princeton
- Mische A, Pattison P (2000) Composing a civic arena: publics, projects, and social settings. *Poetics* 27:163–194
- Oliver M, Myers D (2003) Networks, diffusion, and cycles of collective action. In: Diani M, McAdam D (eds) *Social movements and networks: relational approaches to collective action*. Oxford University Press, Oxford and New York, pp 173–203
- Passy F (2003) Social networks matter. But how? In: Diani M, McAdam D (eds) *Social movements and networks: relational approaches to collective action*. Oxford University Press, New York, pp 21–48
- Passy F, Monsch G-A (2014) Do social networks really matter in contentious politics? *Soc Mov Stud* 13:22–47
- Polletta F (2006) *It was like a fever: storytelling in protest and politics*. University of Chicago Press, Chicago
- Polletta F, Jasper JM (2001) Collective identity and social movements. *Annu Rev Sociol* 27:283–305
- Rucht D (2004) Movement allies, adversaries, and third parties. In: Snow DA, Soule SA, Kriesi H (eds) *The Blackwell companion to social movements*. Blackwell Publishing, Malden and Victoria, pp 197–216
- Snow DA, Benford RD, Mcammon HJ et al (2014) The emergence, development, and future of the framing perspective: 25+ years since “frame alignment”. *Mobilization* 19:23–45
- Snow DA, Burke Rochford E Jr, Worden SK, Benford RD (1986) Frame alignment processes, micro-mobilization and movement participation. *Am Sociol Rev* 51:464–481
- Snow DA, Zurcher L, Ekland-Olson S (1980) Social networks and social movements: a microstructural approach to differential recruitment. *Am Sociol Rev* 45:787–801
- Steinberg MW (1998) Tilting the frame: considerations on collective action framing from a discursive turn. *Theory Soc* 27:845–872
- Steinberg MW (1999) The talk and Back talk of collective action: a dialogic analysis of repertoires of discourse among nineteenth-century English cotton spinners. *Am J Sociol* 105:736–780. <https://doi.org/10.1086/210359>
- Tarrow S (1993) Cycles of collective action: between moments of madness and the repertoire of contention. *Soc Sci Hist* 17:281. <https://doi.org/10.2307/1171283>

- Tarrow S (2010) Dynamics of diffusion: mechanisms, institutions and scale shift. In: Givan RK, Soule SA, Roberts KM (eds) *The diffusion of social movements*. Cambridge University Press, New York, pp 204–219
- Taylor V, Whittier NE (1992) Collective identity in social movement communities: lesbian feminist mobilization. In: Moriris AD, Mueller CM (eds) *Frontiers in social movement theory*. Yale University Press, New Haven and London, pp 104–132
- Tilly C, Tilly L, Tilly R (1975) *The rebellious century, 1830–1930*. Harvard University Press, Cambridge
- Wada T (2012) Modularity and transferability of repertoires of contention. *Soc Probl* 59:544–571
- Woolery, Liz, Ryan Budish, and Levin Bankston (March 2016) *The Transparency Reporting Toolkit: Best Practices for Reporting on U.S. Government Requests for User Information*. Berkman Cent for Internet & Society

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---