BRILL

Social Media Data and the Dynamics of Thai Protests

Benjamin Nyblade
RAND Corporation

Angela O'Mahony
RAND Corporation

*Aim Sinpeng**
University of Sydney

Abstract

Traditional techniques used to study political engagement—interviews, ethnographic research, surveys—rely on collection of data at a single or a few points in time and/or from a small sample of political actors. They lead to a tendency in the literature to focus on "snapshots" of political engagement (as in the analysis of a single survey) or draw from a very limited set of sources (as in most small group ethnographic work and interviewing). Studying political engagement through analysis of social media data allows scholars to better understand the political engagement of millions of people by examining individuals' views on politics in their own voices. While social media analysis has important limitations, it provides the opportunity to see detailed "video" of political engagement over time that provides an important complement to traditional methods. We illustrate this point by drawing on social media data analysis of the protests and election in Thailand from October 2013 through February 2014.

Keywords

Studying Political Engagement: The Promises and Perils of Social Media Data

Understanding political engagement is a central topic in the social sciences. Why is it that at certain times large numbers of people are intensely engaged in politics, while others are not? What explains who votes, who engages in protests or who participates in social movements? Why do we see revolutions and mass uprisings? To answer these questions, scholars need reliable evidence and data concerning political engagement, which is surprisingly hard to come by.

Scholars have relied on two main primary research techniques to study political engagement, namely using quantitative surveys and qualitative narrative studies of particular examples of political engagement. Qualitative research of political engagement typically relies on direct observation of political engagement combined with interviews or ethnographic research in which scholars seek to get their research subjects to relate their experiences to them to incorporate into a narrative. Survey studies of political engagement involve asking a sample of subjects a series of standardised questions and, to the extent possible based on well-established statistical principles, scholars typically seek to generalise beyond their sample.

While many scholars engage in both qualitative and quantitative research concerning political engagement, there is a not-too-surprising disjoint between the two approaches. This is not to say we have not learned a great deal from both approaches—indeed, the vast majority of scholars recognise that work using both approaches has made important contributions to our understanding of political engagement. However, there are well-rehearsed critiques of the two approaches. Qualitative scholars typically point to the narrowness of survey research, which in many ways is inevitable given the cost and constraints surveyors face. Quantitative scholars often question the ability to generalise about political engagement from the narrative accounts and information gleaned from interviews and ethnographic research.

In this paper, we highlight how social media data analysis provides an opportunity to overcome some of the limitations that are common to both of the dominant approaches to studying political engagement. We suggest that the best qualitative and quantitative work on political engagement provides high

quality "snapshots" of political engagement and at best quite fuzzy "moving pictures". What has been lacking is large-scale micro-level data that shows the dynamics of political engagement over time. We believe that analysis of social media data allows for more detailed "moving pictures" of political engagement than has been possible through qualitative or survey research.

As more and more people leave digital records of their voices and views, scholars of political engagement need to recognise the opportunity. We can analyse the political engagement of millions of people online. We are hearing people in their own voices (not filtered through survey questions), we have the opportunity to observe their interactions with others politically (not just their interactions with a surveyor or researcher), and are not limited by the challenges of scaling participant observation or detailed ethnographic research. More importantly, in places where traditional forms of expression of political dissent, such as protests, are restricted, the Internet has opened up a plethora venues of engagement previously closed or inaccessible to ordinary citizens.

There are important limitations and challenges that analysing social media data entail, three of which strike us as particularly important to recognise up front. First, there are important practical and technical challenges scholars face in accessing and analysing social media data. Collecting and analysing social media data requires skills beyond what is taught in the standard research methodology courses taken in graduate school. Second, in contrast to the overly-optimistic views of some advocates of big data, the fundamental challenges of causal inference continue to exist in the big data world. The greater quantity, granularity and frequency of the data can aid in making stronger causal claims than has been previously possible, but the scale of the data does not obviate the need for careful theory development and research design in making causal claims. Finally, even though billions of people now regularly use social media, these people are certainly necessarily representative of all people, and it is not well-established how the dynamics of online political engagement relate to other forms of political engagement.

That being said, we believe that social media provides an important research opportunity for scholars to better understand political engagement, and in this paper we illustrate the opportunities and limitations of using social media data to analyse political engagement through discussion and presentation of social media data surrounding the protests and election in Thailand in late 2013 through early 2014.

The Increasing Importance of Online Political Engagement

Since the advent of the Internet, many commentators and scholars have hoped that widespread diffusion of Internet use could be a force for political change (Trippi, 2004; Benkler, 2006; Grossman, 1995; Gillmor, 2004). Scholars argue that there are several mechanisms by which the Internet can bring about political change (Shirky, 2008; Chadwick and Howard, 2008; Ferdinand, 2000). Use of the internet may possibly (a) increase political participation; (b) activate previously inactive citizens; (c) challenge politically vested interests of leaders; (d) narrow the gap between political elites and the mass public, and (e) devolve power from traditionally centralised institutions to the periphery. This article seeks to understand the dynamics of political participation online only in so far as understanding the role Twitter and Facebook play as sites of contention during an important political event: A series of mass protests and an election in Thailand in 2013–2014.

While the Arab Springs have spawned a growing scholarship on political participation and social media,[1] debate continues within the fields of political science and communication about exactly which role social media plays in relation to political participation (Segerberg and Bennett, 2012). Some scholars argue that people who use social media help drive political participation offline. Tufecki and Wilson's (2012) study found that protesters in Egypt's Tahrir Square not only learned of the protests through social media, but their social media use made it more likely for them to engage in political protest. Valenzuela et al.'s (2012) survey data on Facebook users in Chile also shows that there is a strong association between Facebook use and offline protest activity. Similarly, Howard and Hussein (2013: 5) argued that the Internet was so important to civil society actors in Arab states because they lacked other forms of political communication and Internet content could be hosted on servers that were out of reach from the state. In China, online activism, particularly political protest, thrived online partly because offline political opposition was strongly restricted (Yang, 2013).

Others scholars examine the potential democratising nature of social media and how they can play a role in facilitating political participation of those not usually engaged in politics (Gillmor, 2006). Social media channels, in particular, are viewed as having the potential to create new participatory culture due to its interactive and user-generated features (Leadbeater, 2008; Jenkins, 2006). In a

---

1 See, for example, Howard and Hussain (2011), Allagui and Kuebler (2011), and Eltantawy and West (2011).

large multi-country survey, Loader et al. (2014) find that there is a strong, positive relationship between social media use and political engagement among youths in Australia, the United States and the United Kingdom. In a continent where voting and political party membership has long been in a decline, social media tools provide some optimism that more networked European youths can enhance more traditional forms of political participation. Vitak et al. (2011) found Facebook activity to be a strong predictor of other forms of offline activities, such as petition signing, prior to the 2008 US presidential election.

Other scholars have highlighted important limitations and concerns about the wave of social media analysis. Perhaps most prominently, boyd and Crawford (2012) emphasise that even though online behaviour has become important for hundreds of millions, perhaps even billions, of people, "Facebook users" or "Twitter users" are not "all people", and scholars need to be careful in making claims to generality. As they also note, even when scholars are careful in their writing about the limitations of their work, they should be cognisant as to how popular media of research tends to distort or ignore scholarly recognition of the limitations of research.

This paper contributes to the debate on political engagement online in two key ways. First, much of the existing research on online political engagement focuses on the developed world, with the exception of a handful of Asian countries, such as China and Korea. However, social media use is more widespread in Asia (and beyond) than one might recognise from the existing literature, and we illustrate that through the presentation of data on social media engagement surrounding protests and an election in Thailand in 2013–2014. Thailand, is a middle-income developing country, whose Internet penetration mirrors an average in Asia and the Global South more broadly. The dynamics of political engagement online in a developing country may be quite different from that in the developed world due to the high degree of digital inequality but also because there are many weak democracies and authoritarian states in which governments seek to constrain online activism. Recent experiences in Thailand, can serve as an important case in which political engagement online is likely to have the most immediate impact: during a political crisis (Sinpeng, 2014).

Second, we marry a narrative account of digital activism and protests in Thailand with detailed micro-level data of social media use, a multi-method research approach we think is promising for future social scientists. While our use of the social media data in this article is purely exploratory, we hope to illustrate for other scholars who are less familiar with social media data and online political activism some of the opportunities and challenges involved in social media research.

Thailand's Cyber Coup

Background

The recent coup d'état in May 2014 in Thailand was the first ever coup announced on social media. For a country that had witnessed many military takeovers in its political history, the level of engagement with new media was unprecedented. Not only was the coup first announced on Twitter and Facebook accounts of the National Council for Peace and Order (NCPO)—the new coup government administrative body—but the Thai military has relied on social media and chat applications as the main methods of information dissemination.[2] While the nature of the political conflict that has marred this Southeast Asian nation has remained largely the same for the past decade,[3] the means have changed. Social media has emerged as one of the most important tools for mobilisation and governance.

The May 2014 coup was the 19th coup attempt, and the 12th successful coup, since Thailand transitioned to a constitutional monarchy in 1932. This most recent coup was triggered by a political impasse and ongoing unrest that began in the summer of 2013. The majority Pheu Thai Party (PTP) government of Yingluck Shinawatra, which came to power through the ballot in 2011, sought to give a blanket amnesty to all those involved in the political conflict in Thailand over the past decade. Anti-government forces saw this as a way for Yingluck to vindicate her brother, Thaksin Shinawatra, who was dislodged from power in the previous 2006 coup for alleged corruption, cronyism, and posing a threat to the monarchy. The amnesty bill helped to align various opposition groups under the rubric of a loosely organised protest movement, the People's Democratic Reform Committee (PDRC).

The PDRC was primarily led by the main opposition party in parliament, the Democrat Party. The Democrats, who have failed to win an election since 1992, decided to pursue street politics full-time and walked out of parliament to fight for power not through the ballot, but through the streets.[4] Building on the frustration of the various groups that had been disaffected by the Shinawatras' rule, the Democrats saw an opportunity to mobilise their supporters to not only oust PTP, but also to put an end to electoral democracy. As hundreds of thousands of anti-government protesters poured into the streets and seized key government offices, Yingluck felt that the only way out was

---

2 http://www.washingtonpost.com/blogs/monkey-cage/wp/2014/05/27/thailands-cybercoup/.

3 For a good discussion of the political conflict in Thailand, see Pongsudhirak (2012).

4 See Sinpeng (2013).

to get people to the polls, a decision she believed would lead to her party's victory. The PDRC retaliated with more protests and launched a "No Vote" campaign to get Thais to forgo their voting obligation (Thailand has mandatory voting). An escalation of violence leading up to election day prompted the government to declare a state of emergency. The efforts of the Democrats and the PDRC contributed to the annulment of the February 2014 election, the ousting of Prime Minister Yingluck, and sustained protests in Bangkok. After numerous failed negotiation efforts between pro- and anti-government elites, the military intervened to break the deadlock in Thai politics in May 2014.

The current conflict represents a continuation of a crisis that began in 2005, when opposition forces, known as the People's Alliance for Democracy (PAD), or more widely, "Yellow Shirts", mobilised against the Thaksin Shinawatra government. Thaksin was a hugely popular but controversial leader whom the Yellow Shirts saw as highly corrupt, tyrannical and manipulative. The PAD orchestrated a series of mass protests in Bangkok that eventually led to a military coup d'état in September 2006. Yet, for each national election held since, a Thaksin-aligned party has managed to secure an electoral victory, including the most recent government led by his sister, Yingluck. The Shinawatras were able to build a strong and loyal electoral support base, thanks in part to their populist appeals, including advocating universal healthcare, village funds, debt moratorium and various rural development programmes. A common feeling among many Thais is that never before has a politician met the demands of the rural poor like Thaksin, and with his sister continuing his pro-poor policies, the support base for the PTP seemed secure.

Those who oppose the Shinawatras' rule argue that the poor are being duped by populist "hand-outs". Much of the Thai middle class has opposed the PTP, seeing the middle-class position threatened by the growing power of the poor. Arguably, Thaksin effectively enfranchised the rural poor, who for the first time understand their own political power. Their power is based on not just the ability to vote, but on the knowledge that if they are united, they represent nearly 70% of the population and can thus choose the government they want. The middle class and the elites used to be decisive in selecting Thailand's government, but now they feel that they have forever lost this privilege. Moreover, during Thaksin's rule the tax-paying middle class felt that their hard-earned money was being wasted by a populist government on ineffective policies catering to the poor. The deep resentment that they as a class were paying for a government that completely ignored their interests sparked outrage at the Shinawatras.

Social Media and Popular Mobilisation

The PDRC is among the most media-savvy protest movements in contemporary times, and like the Thai military's media strategy, is largely a response to the increasing number of Thais who are "wired". The PDRC's media campaigns included a 24-hour satellite TV channel, Bluesky TV,[5] very popular Facebook pages, opposition media outlets, website,[6] magazines,[7] radio stations and highly active social networking supporters. The decision to "go viral" with the protest movement made the most sense, given that PDRC supporters were drawn largely from the urban middle class, particularly from Bangkok and southern Thailand. Coupled with growing Internet diffusion and a very high use of Facebook, organising the protest movement online was a natural step. Suthep Taugsuban, the PDRC leader, used to have a public page on Facebook when he served as a member of parliament (MP), and his page had only 30,000 "likes" in March of 2013. With Suthep assuming the leadership of the new protest movement, his page's popularity grew 6,000% to nearly 2,000,000 "likes" one year later. Suthep's Facebook page ranks as 26th the most popular Facebook page in Thailand, according to Zocialrank.

Thailand experienced one of the most rapid growth rates of ICT diffusion in developing Asia since the Internet became commercially available in 2002. More than one-third of the 67 million Thai people have access to the Internet. The Southeast Asian nation also tops the list of the highest rate of Facebook penetration in the world, with 24 million of the 27 million on the internet having a Facebook account, representing nearly 90% penetration.[8] Social media represents among the top reasons why Thais go online, according to the Thailand's Internet Profile 2013 survey. Bangkok is the world's leading Facebook city[9] and 2013's most Instragrammed place.[10] According to Bangkok-based digital agency, EdgeAsia, Thailand's monthly Facebook activities in 2013 include: 7.1 billion shares, 5.5 billion "likes", 1.3 billion comments, and 206 million status updates.[11] Although other Asian nations, such as China, Malaysia and South Korea, have an overall higher internet penetration rate, the intense adoption of social media in Thailand amongst those online is notable.

---

5   http://www.blueskychannel.tv/.

6   For example, http://www.reformthai.com.

7   https://www.facebook.com/lipsmag.

8   http://www.techinasia.com/facebook-thailand-grows-to-24-million-users-infographic/.

9   http://www.socialbakers.com/blog/647-top-10-biggest-facebook-cities.

10   http://www.cnn.com/2013/12/16/travel/most-instagrammed-places-2013/.

11   http://www.techinasia.com/facebook-thailand-grows-to-24-million-users-infographic/.

TABLE 1   *Thailand's most popular public Facebook pages (political category), 2014*

	Likes	Who	3-Mo gain	Rankings
Suthep Taugsuban	1,950,000	PDRC leader	200%	#26
Yingluck Shinawatra	2,500,000	Prime Minister	68%	#17
Panthongtae Shinawatra	2,100,000	Thaksin's son / PM's nephew	34%	#22
Abhisit Vejjajiva	2,000,000	Democrat Party leader / PDRC	22%	#23

SOURCE: ZOCIALRANK (BASED ON 19 MARCH 2014 FIGURES) AND AUTHORS' ANALYSIS

Unlike Mubarak's move to shut down the Internet in Egypt during the height of the Tahrir Square uprising, Yingluck believed that keeping the cyber communication line open would allow her government to counteract the PDRC's digital activism. The "cyber war" between the government, led by the Shinawatra clan (Yingluck and Thaksin's son, Panthongtae) and the opposition, under the leadership of the Democrat Party/PDRC (Suthep and Abhisit), began in the fall of 2013 with both sides seeking to mobilise their own supporters. The most pivotal moment was the lead up to the February 2014 election, following the House dissolution. Given its successful track record, the government believed it would win the election and pushed hard for another electoral victory. Meanwhile, the Democrat Party, having completely boycotted the election, launched a "No Vote" movement—an anti-election campaign against what they saw as a corrupt and unjust electoral system. Supporters of the No Vote campaign strongly believed that a low voter turnout would help debunk the myth of the PTP's popularity.

The PDRC formally began their No Vote campaign with a "Bangkok Shutdown Day" on 13 January, where thousands of supporters began occupying key government and commercial sites in the heart of Bangkok. Until the election day, massive campaigns online and offline, including daily protest rallies, were launched online prompting spikes in the activities of the top 4 political public pages (see Figure 1). Calculations based on the average interactions per post show that on "critical days", such as days leading up to an election or a major planned protest, we see spikes in Facebook activities. In the case of Suthep (green), on 23 February, his single post that day garnered nearly half a million interactions from Facebook users. It was this day when protests turned violent, producing casualties. Similarly, when Panthongtae (red) launched a counter campaign to get people to vote on 16 January, his Facebook activity

f Avg. Interactions Per Post | All Profiles

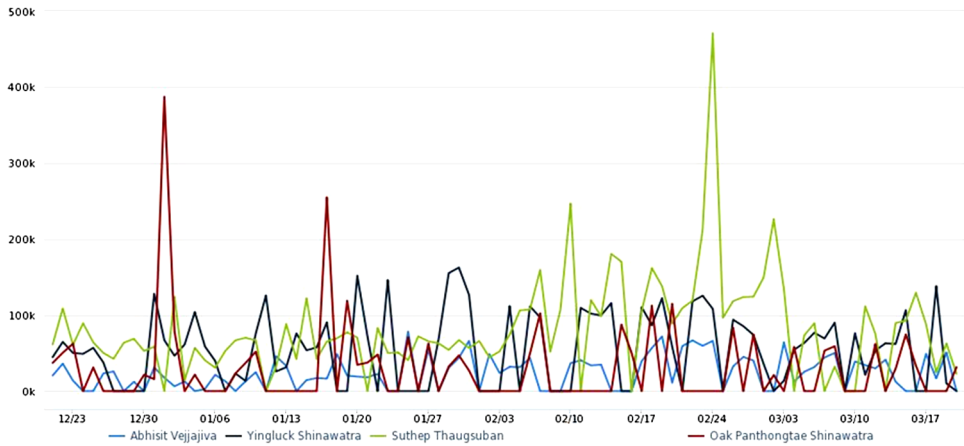12/21/2013 - 03/20/2014 (daily)

quintly

FIGURE 1   *Interaction rates of four Thai Facebook pages (20 December 2013–21 March 2014).*
Average interactions = sum of shares, likes and comments per post.
SOURCE: AUTHORS' CALCULATION USING DATA FROM QUINTLY

experienced a surge in interactions to 255,000 per post. The PDRC's No Vote campaign kept roughly an additional 10 million voters from the polls on election day (above the usual 10 million non-voters), and its protests led to scores of cancelled voting stations, including making elections impossible in districts in 9 out of 11 provinces in southern Thailand. There were further electoral irregularities, and the Election Commission of Thailand has been unable to resolve this issue, although it has recently invalidated the election, leaving Thai politics in a state of limbo.

Social media, Facebook, in particular, was monumental in the mobilisation of anti-government protests in Thailand. This round of protests, unlike the previous episodes, focused much effort on social media. Arguably this was because Thailand's cyberspace has reached a critical mass, with more than 1/3 of households online and active on social media. More importantly, because most Thais in the urban areas, the main source of supporters of the PDRC, are frequent Facebook users, the social media tool presents a relatively low cost means to overcoming collective action problems. Since the government showed no signs of active repression online, the cost of joining the protest movement online and ultimately on the streets became much lower as more and more people joined in.

Political Engagement via Twitter

While much of the political mobilisation surrounding the Thai protests and elections occurred on Facebook, there are over two million Twitter users in Thailand and, in this section, we use data on Twitter activity to illustrate the depth and level of data that such data can provide surrounding mass political engagement around the Thai protests and elections. The data were pulled from Twitter via Datasift, one of the handful of companies that has had access to the full Twitter historical archives and pipeline of current tweets. We collected data for October 2013 through February 2014.

We compiled a dataset of tweets based on a keyword search of the complete Titter archives. Our keywords centered around protests and the election. Specifically we choose four keywords in Thai that were representative of the types of activities and words used in social media surrounding the election and protests. Our four words were "election" (leuk tung, เลือกตั้ง), "demonstration" (choomnum, ชุมนุม), "protest" (pratuang, ประท้วง) and "mob" (mob, ม็อบ). As we were searching in Thai, we were comfortable using fairly general words, and an analysis of sample tweets from our data pull suggest an extremely low proportion of our observations were unrelated to the Thai protests or election. We chose not to focus on specific hashtags or specific terms associated with mobilisation of the protests (e.g., #BangkokShutdown), as our aim was to assess general political engagement surrounding the protests and election, rather than focus on specific mobilising or rhetorical strategies.[12]

> No matter which mob is out on the streets, they are fighting for their own self interests. Don't use "nationalism" as a front for what ultimately will be your own benefits

> The farmers' protest and the corn protests represent the true failures of populist policies

> Hey you PM (bitch) can you just get out? Photos of me going to protests are filling up all my available data space

---

12   Focusing on the use of particular hashtags as a mobilisation strategy is also an important avenue of research in this field (e.g., Bastos et al., 2013).

I'm in class (school) and we could hear the sounds of the mob. [Re-dacted teacher name] said if they [the government] disperse the mobs, we will have to get out and fight, hahahahahahahahaha

Anyone wanna protest at Victory Monument w 7 others?

Every day I am late for work because of all these mobs but today I also ran out of hot water! Had to use regular (cold) water and now can't feel my face

FIGURE 2   *Sample tweets concerning Thai protests, January 2014*

Examples of tweets that our search pulled are shown in Figure 2. In Figure 2, we provide anonymised translations of original tweets.[13] Many of our observations are also retweets—most commonly of news reports or journalists' coverage of the protests. As can be seen, while Twitter was used by protestors for mobilisation purposes or political commentary, others post about being inconvenienced by the protests.

Our dataset includes more than 120,000 tweets in October, 1.8 million tweets in November, 875,000 tweets in December, 720,000 tweets in January and 370,000 tweets in February. The distribution of tweets over time is illustrated in Figure 3. The density of tweets is strongly associated with the timing of major protest activities in Thailand. Most striking is the number of tweets in late November and early December, with nearly 50,000 tweets per hour mentioning one of our keywords during peak hours in the lead-up to major street protests. This is followed by a decline in online activity on and shortly following the King's birthday in early December, before a burst of activity again at the beginning of the following week. Another burst of activity occurs the day

---

13   Twitter use policies require that those who use their data adhere to certain rules, and the fact that all the Tweets analysed are nominally in the public record should not absolve researchers of potential privacy or other ethical concerns involved in reporting specific instances of social media use, as scholars should not assume users are familiar with the privacy settings and privacy policy of their social media services (see Lawler and Molluzzo, 2010). The particular tweets we present anonymised translations of were selected in order to provide an indication of the range of content and tones of tweets that were collected in our data pulls from within a single day within January 2015. We have taken particular care to de-identify these tweets, removing individually identifiably information. We have tested whether the Twitter users can be identified by reverse translation and searching twitter of our presented sample tweets and found that individuals could not be readily identified (cf. Young et al., 2014).
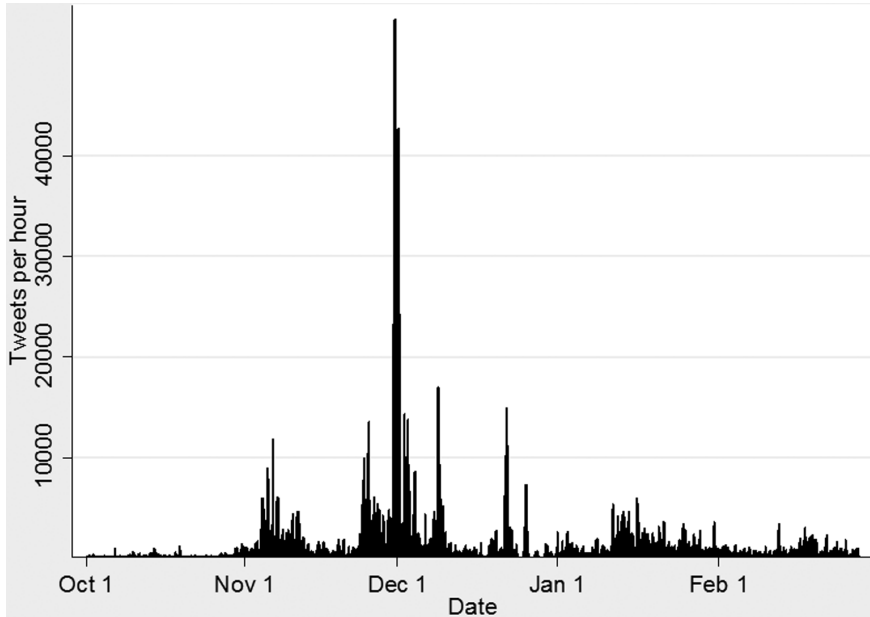
FIGURE 3   *Tweets mentioning Thai protests over time (1 November–1 March)*

prior to and day of the major protests on 24 December, and then the activity increases again with the lead-up to "Bangkok Shutdown Day" and the protests that followed through the election on 2 February.

Although Twitter is often focused upon as a venue for use by journalists and news outlets in Thailand, the Twitter activity we see in Figure 3 is not driven by a small number of accounts. In fact, the nearly 3.9 million tweets in our dataset are from 250,000 different accounts. A majority of accounts tweet about the elections or protests fewer than five times in our five month period, as can be seen in Table 2, which shows the distribution of the number of tweets across Twitter accounts in our dataset. The accounts that show the highest numbers of tweets concerning elections and protests during this period are the Twitter accounts of major Thai media outlets—the highest is 6,481 individual tweets, that is to say, an average of more than 40 tweets per day that include one of our keywords.

However, even excluding the accounts associated with media outlets, there is a wide range of levels of political engagement shown in the data. The bulk of people are politically engaged at fairly low levels and/or infrequently on Twitter—84% of the Twitter accounts that tweeted about the protests or elections had fewer than 20 tweets mentioning protests or the election in this five-month period. However, that means that more than 40,000 accounts posted at

TABLE 2   *Distribution of tweets across Twitter accounts*

Tweets mentioning keyword	Accounts	Proportion
1	70,848	27.9
2	32,908	13.0
3	20,308	8.0
4	14,760	5.8
5	11,392	4.5
6 to 10	33,464	13.0
11 to 20	29,073	11.4
21 to 50	26,412	10.4
51 to 100	9,447	3.7
101 to 200	3,480	1.4
201 to 6,481	1,777	0.7

least 21 tweets about the protests, and in fact more than 5,000 accounts posted at least 101 tweets.[14]

Social media data provides extremely detailed information about the timing (down to the fraction of a second) and the content of online social media communication, which can be extremely useful, but increasingly with the advent of geo-located internet capable devices (particularly smart phones), social media data includes precise coordinates for where the online communication originates. The number of Tweets and Facebook posts that include geo-location data have increased dramatically in recent years, although it is important to recognise that just as online political engagement may not necessarily be representative of offline political engagement, social media engagement by users who have enable geo-location may not be perfectly representative of all online users. That caveat aside, geo-located data may be increasingly important to analysts, and we discuss how it can be used in our case below.

Not surprisingly, given that our keywords were in Thai and the protest activity centred on Bangkok, the vast bulk of the tweets were written by Twitter

---

14   This is consistent with patterns of political engagement more broadly in other media as well: While some people are intensively engaged in political activities, the bulk of people are far less consistently involved. Social media data presents a great opportunity to examine and to better understand how the language, social networks and patterns of online engagement in politics differ between from those with different levels of political engagement, something not easily done with other data sources.

TABLE 3   *Distribution of geo-tagged tweet locations globally*

Country	Number	Proportion
Thailand	51,143	95.14
NA	1,844	3.43
Japan	130	0.24
USA	123	0.23
UK	61	0.11
Malaysia	56	0.1
Laos	53	0.1
Philippines	50	0.09
Germany	47	0.09
Canada	36	0.07
Myanmar	26	0.05
South Korea	26	0.05

Note: Locations with 1–24 Tweets include Argentina, Australia, Austria, Brazil, Cambodia, China, Czech Republic, Dominican Republic, Egypt, Finland, France, Greece, Hungary, India, Indonesia, Ireland, Italy, Jordan, Netherlands, New Zealand, Oman, Panama, Poland, Russia, Saudi Arabia, Serbia, Solomon Islands, South Africa, Sweden, Switzerland, Taiwan, United Arab Emirates.

users in Bangkok. Just under 54,000 of our tweets (1.5% of tweets) were original tweets (not counting retweets) that were posted from geo-tagging enabled devices with the geo-tagged information publicly available. We took the longitude and latitudes provided for those geo-tagged tweets and ran them through the *maps* library in R, and the distribution of locations is shown in Table 3. Some 95% of the geo-tagged locations were based in Thailand, and 3.5% of the remaining 5% were locations classified as Not Available (NA). The bulk of NA locations were locations in the Pacific Ocean that were plausible but not identified by the mapping software as associated within the land boundaries of a particular country (indeed, most are locations quite close to Thailand), although some are clearly errors in the geo-tagging process (e.g., precisely reported 0° latitude and 0° longitude).

The bulk of the remaining geo-tagged tweets are from countries with a reasonable population of Thai migrants or student populations, including other countries in Asia, such as Malaysia, Japan, Laos and the Philippines, as well as
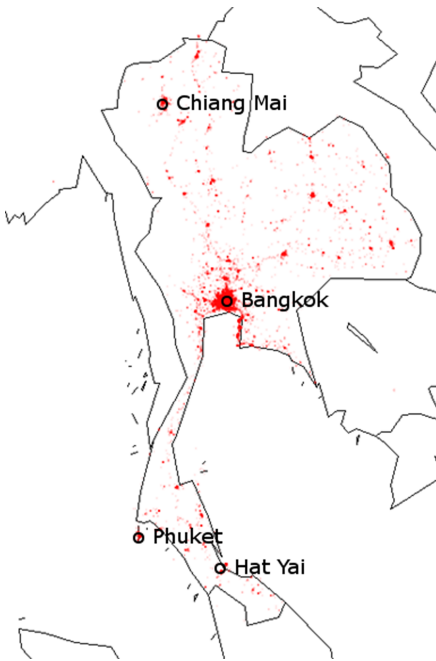
FIGURE 4   *Distribution of geo-tagged tweet
            locations in Thailand*

the US, Canada and the UK. Most of the tweets were by migrants or students who were residing abroad; however, in some cases, the tweets were from Thai nationals who were clearly travelling abroad for shorter durations and commenting about the protests back home. Given that our sample only includes those who were talking about the protests or elections in Thai, our sample doesn't include anyone discussing the protests in English and other languages, and thus does not fully capture international interest in what was occurring in Thailand. Even so, the diversity of locations that entered our sample is striking, with over 40 countries represented.

The geo-tagging allows us to differentiate locations within Thailand as well. Figure 4 shows the distribution of tweets across Thailand (again using the R *maps* library). Not surprisingly, the tweets are concentrated in and around Bangkok. Not only were the protests centred in Bangkok, but Internet penetration and social media use is lower outside of Bangkok. However, the distribution of engagement beyond Bangkok is quite wide, albeit concentrated in cities and major towns.

Overall, 82% of our geo-tagged tweets were from within Greater Bangkok, with an additional 15% of locations from within the extreme points of Thailand
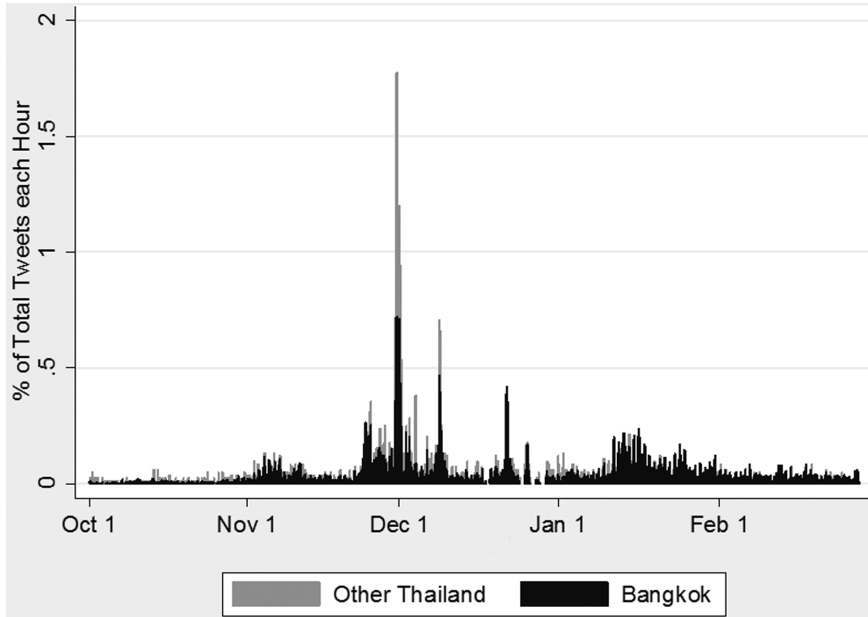
FIGURE 5   *Geo-tagged tweets over time by location*

(a latitude between 5.5° and 20.5° N and a longitude 97.3° and 105.6° E). Figure 5 shows how intensity of online political engagement varied over time based on whether the tweets were geo-located in Greater Bangkok or elsewhere in Thailand. Geo-tagged tweets include latitude and longitude to a relatively fine degree, which allows the potential for tailoring studies to fairly precise areas. Here, we identify Greater Bangkok broadly, and contrast the proportion of tweets in our samples by those located in Bangkok and elsewhere in Thailand that occurred at different periods of time.

In general, the patterns are broadly similar. The most notable difference is that the discussion of protests are more concentrated around the major early protest events in late November and early December outside of Bangkok, and are proportionately less common in the subsequent rounds of protests in January and February 2014. Well over half the total number of tweets from outside of Bangkok are concentrated in a 30-day period in November and December, whereas the levels of political engagement via Twitter in Bangkok are much more dispersed across our five-month period. This suggests a continually higher level of engagement in Bangkok, whereas political engagement beyond Bangkok was more concentrated around major events.

While we do not engage here in detailed analysis of the extent to which the sentiments expressed regarding the protests and election are different in

these groups, it is worth briefly noting that the choice of terminology regarding the protests tends to be different based on location. Tweets about the protests in Greater Bangkok include "demonstration" (choomnum, ชุมนุม) 54% of the time, "mob" (mob, ม็อบ) 41% of the time and "protest" (pratuang, ประท้วง) 7% of the time. However, tweets outside of Bangkok were more than twice as likely (16% of tweets) to use the term "protest" and less likely to refer to "demonstrations" than those in Bangkok (45% of tweets). The different word usage likely reflects the different ways that Thais across the country experienced the events, with the immediacy of the demonstrations more striking to those in Bangkok, whereas the more general "protest" term having greater prevalence elsewhere.

Discussion

The discussion and data presented above highlight the potential value of using "big" social media data to develop more nuanced "moving pictures" of political engagement and political mobilisation over time and space. While narrative accounts of political engagement are important, by pairing them with detailed dynamic micro-level data, at a minimum we have the opportunity to provide far more nuance and granularity to narrative accounts, and have the potential to marry such narrative accounts with more sophisticated statistical analysis as well. While we have not provided detailed statistical analysis of the particular data presented, we hope that by providing examples of social media data we can inspire scholars and students to explore the possibilities that the fine-grained, dynamic, micro-level data that social media use provides.

Studying political engagement using big data from social media sites such as Twitter and Facebook allows us to work with extremely detailed data to a degree that has simply not been possible previously. We have the opportunity to see political engagement by millions of people. We have the opportunity to assess people's own words on a scale previously not possible, and also consider their choices to like, share, retweet and comment on the words of others. However, using social media data to better understand political engagement faces serious challenges, both in terms of understanding political engagement via social media, and in terms of understanding the relationship between online and offline political engagement. We revisit the three challenges we previously highlighted concerning social media analysis. What is worth highlighting is that while the particulars of social media analysis may be new, the general concerns are common to more common social scientific research strategies—they are, after all, fundamentally concerns over external validity, internal validity and the logistics of research techniques.

Even in Thailand, where social media adoption and use is comparatively quite high, only one-third of the population is relatively active on social media, while the remaining two-thirds are not. Furthermore, even for the people active on social media, it is not always clear when online speech will or will not translate into political engagement or mobilisation elsewhere. Of course, understanding the relationships between different forms of political engagement is a general challenge in the literature that long predates the Internet. While typically the evidence points towards the common belief that higher levels of political and social engagement of any particular type are associated with higher levels of engagement of other types, this need not always be the case, and a great deal of work remains to better understand the nature of political engagement of online and how it relates to other forms of political engagement.

Furthermore, while careful analysis detailed micro-level data over time can improve the reliability of some forms of stronger causal inferences made using statistical techniques,[15] contrary to the claims of some advocates of big data, big data does do not obviate the importance of theory development and research design required to make compelling causal claims. Fundamentally, social media data and other sources of big data are merely data. The intrinsic challenges of evaluating counterfactuals, something that is involved in making and assessing any causal claim, will not simply disappear if hidden under a mountain of data.

While these two limitations (generalisability, causal inferences) are not unique to social media data, it is important to recognise that at a practical level there are some new challenges to gathering and working with the data. Social media data are not always readily available for interested academics,[16] and many scholars in areas that could be illuminated by working with social media data do not have the technical skills to manage and analyse the data. Even though the academic community studying political engagement has relied on and developed multiple quantitative and qualitative approaches in the past, with the rise of new data sources comes new technical challenges. Not only are

---

15   For example, both difference-in-difference and regression discontinuity statistical approaches to moving beyond simple correlational analysis require comparatively detailed data and fairly large sample sizes. In general, large-scale micro-level dynamic panels of data are generally seen as the best opportunity for using such techniques.

16   Some efforts are being made in this area. Twitter and Gnip jointly announced an initiative to provide access to Twitter for some academic projects; Bit.ly and Crimson Hexagon are committed to working with academics as well. A number of scholars have and are working with the Facebook Data Science Team. However, these efforts still remain much more limited than academic demand.

the statistical techniques used in social media analysis fairly different, the magnitude of the data being used presents challenges in ways that survey data do not.

For example, just the Twitter dataset summarised in Section 4 of this paper included nearly four million tweets, and given the amount of text included in each observation in the dataset, the file size of the full datasets exceeded 40 GB. The ability to make the most of social media data requires knowledge of techniques, such as textual analysis, network analysis, GIS mapping and potentially a wide range of other techniques, that differ greatly from traditional survey analysis, the most common quantitative technique for studying political engagement. While some of these techniques are becoming more commonly taught in political science and sociology graduate schools, to be on the cutting edge of work in this area requires interdisciplinary connections to scholarship in computer science and statistics that is comparatively rare.

While we don't wish to minimise these challenges or ignore the limitations of what can be learned through social media analysis, the technical challenges related to working with social media data are not dramatically different from those challenges facing prior generations of scholars who learned to adapt to the development of new technologies. Technological change provides both new opportunities and new challenges for researchers. We believe that new social media data provides a new opportunity to develop and analyse detailed moving pictures of political engagement based on the individual voices of millions of people using social media, which can provide an important complement to more traditional approaches to studying political engagement.

References

Allagui, I. and J. Kuebler (2011) "The Arab Spring and the Role of ICTs". *International Journal of Communication* 5: 1435–1442.

Bastos, M., R.L.G. Raimundo and R. Travitzki (2013) "Gatekeeping Twitter: Message diffusion in political hashtags". *Media, Culture & Society* 35(2): 260–270.

Benkler, Y. and H. Nissenbaum (2006) "Commons-based peer production and virtue". *Political Philosophy* 14(4): 394–419.

Boyd, D. and K. Crawford (2012) "Critical questions for big data: Provocations for a cultural, technological and scholarly phenomenon." *Information, Communication & Society* 15(5): 662–679.

Castells, M. (2001) *The Internet galaxy: Reflections on the Internet, business, and society*. Oxford: Oxford University Press.

Chadwick, A. (2007) "Digital network repertoires and organizational hybridity". *Political Communication* 24(3): 283–301.

Chadwick, A. and P. Howard (2008) *Routledge handbook of internet politics*. New York: Routledge.

Eltantawy, N. and J. West (2011) "Social Media in the Egyptian Revolution: Reconsidering Resource Mobilization Theory". *International Journal of Communication* 5: 1207–1224.

Ferdinand, P. (2000) "The internet, democracy and democratization". *Democratization* 7(1): 1–17.

Gillmor, D. (2006) *We the media: Grassroots journalism by the people, for the people.* Sebastapol: O'Reilly Media.

Grossman, L.K. (1995) *The electronic republic*. New York: Viking.

Hindman, M. (2009) *The myth of digital democracy.* Princeton: Princeton University Press.

Howard, P. and M. Hussain (2011) "The Role of Digital Media". *Journal of Democracy* 22(3): 35–48.

Jenkins, H. (2006) *Convergence culture: Where old and new media collide.* NY: NTU Press.

Lawler, J. and J. Molluzzo (2010) "A study of the perceptions of students on privacy and security on social networking sites (SNS) on the internet". *Journal of Information Systems Applied Research* 3(12): 3–18.

Leadbeater, C. (2008) *We-Think: Mass innovation, not mass production.* Profile Books.

Loader, B.D., A. Vromen and M.A. Xenoa (2014) "The networked young citizen: Social media, political participation and civic engagement". *Information, Communication & Society* 17(2): 143–150.

Norris, P. (2001) *Digital divide: Civic engagement, information poverty and the Internet worldwide.* Cambridge: Cambridge University Press.

Shirky, C. (2008) *Here comes everybody: The Power of organizing without organizations.* New York: Penguin.

Segerberg, A. and W.L. Bennett (2012) "The logic of connective action: Digital media and the personalization of contentious politics". *Information, Communication & Society* 15(5): 739–768.

Sinpeng, A. (2013) "Party-movement coalition in Thailand's political conflict (2005–2011)", in P. Liamputtong (ed.) *Contemporary Socio-Cultural and Political Perspectives in Thailand*. Dodrecht: Springer.

Pongsudhirak, T. (2012) "Thailand's uneasy passage". *Journal of Democracy* 23(2): 47–61.

Trippi, J. (2004) *The revolution will not be televised: Democracy, the internet, and the overthrow of everything.* NY: Harper Collins.

Tufekci, Z. and C. Wilson (2012) "Social media and the decision to participate in political protest: Observations from Tahrir Square". *Journal of Communication* 62(2): 363–379.

Valenzuela, S., A. Arriagada and A. Scherman (2012) "The social media basis of youth protest behavior: The case of Chie". *Journal of Communication* 62(2): 299–314.

Vitak, J., P. Zube, A. Smock, C.T. Carr, N. Ellison, and C. Lampe. (2011) "It's complicated: Facebook users' political participation in the 2008 election". *Cyberpsychology, Behavior, and Social Networking* 14(3): 107–114.

Yang, G. (2013) *The power of the internet in China: Citizen activism online*. NY: Columbia University Press.

Young, S.D., C. Rivers and B. Lewis (2014) "Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes". *Preventive Medicine* 63: 112–115.