

**Weibo, social protests and government legitimacy in China: A data-driven analysis
of Weibo messages on “mass incidents”**

Jingrong Tong and Landong Zuo

This is the final version submitted before publication.

To Cite:

(2014) ‘Weibo communication and government legitimacy in China: a computer-assisted analysis of Weibo’, in Special Issue: The Internet, Social Networks and Civic Engagement in Chinese Societies, *Information, Communication & Society*, 17(1), 66-85

This article examines whether and to what extent online communication challenges the legitimacy of the authoritarian Chinese government through an analysis of data from Sina's Weibo (China's version of Twitter) about the 2011 Wukan and Haimen mass incidents, collected from December 2011-June 2012. The activities of Weibo users and the discourse of tweets they have posted are examined and the extent to which online communication on Weibo challenges government legitimacy is assessed.

China's social protests in the digital environment

The last few years have witnessed the proliferation of social protests across the country, as an expression of the public's demands that cannot be fulfilled in other ways (Gilboy and Read 2008). Farmers, workers, and home owners are among the most prominent protest groups. Contextual factors such as "inequality and corruption", "growing rights consciousness, organisational skill", and "comparisons with other social movements", are responsible for the growing social unrest (Gilboy and Read 2008; Lum 2006).

Social unrest is also facilitated by the fast growth of the internet. According to CNNIC, the number of internet users increased to 532 million by April 2012 and around 40 percent of the Chinese population is now connected to the Internet¹. Apart from entertainment activities, internet users participate in online discussions

on social and political issues, ranging from official corruption to the kidnapping of school children. The development and use of mobile devices and of apps also advances online participation. The popularity of social media sites, such as Weibo, has connected like-minded people more closely together.

Compared to traditional media that are banned from reporting topics like social protests, the online medium enjoys relatively more autonomous speech. Recent cases, such as the Weng'an incident² in 2008, have drawn the Chinese leadership's attention to the online medium. The internet has been used to organise protests, to disseminate information about protests, and to interact with the press (Fewsmith 2008). Most recently, Weibo has joined other social network sites, becoming an effective field for publicising and discussing these events (AFP 2011).

What threatens government legitimacy?

Social protests in China are thought of as being short-lived, curtailed by the repressive actions of local governments, and by a weak connection between mass unrest and organised political opposition, as well as being "largely incoherent and disorganised", and "fragmented" in themselves (Gilboy and Read 2008; Cai 2008). How and to what extent online discussion of social protests can impair government legitimacy lies in whether such online discussion generates factors that threaten government legitimacy or not.

Since the 1980s economic reform, China has seen rapid growth and the development of the market, while the Party leadership retains “political supremacy” in the process (Heberer and Schubert 2006). The leadership, however, is not able to guarantee legitimacy for its rule in the new century for three reasons. First, the legitimacy crisis results from the Party’s dependence on the dominant Communist ideology that is dying in the face of economic reform and the dismantling of the old socialist system. Other ideologies, such as liberalism, are rising, which enables a critique of the regime’s legitimacy (Holbig 2006; Guo 2003; Gilley and Holbig 2009). Second, the slim possibility of having political reforms puts political legitimacy in danger (Schubert 2008). Third, social problems, such as corruption, social inequality and injustice, and environmental problems, have further reduced the level of popular consent to the political legitimacy of the regime (Wong 2004; Potter 1994; Chen, Zhong et al. 1997; He 2000).

The central government increasingly relies on both stirring popular emotion, through nationalism for example, and on domestic economic achievements to regain and maintain its legitimacy (Downs and Saunders 1998/99; Shue 2004). A unique “political structural reform” has been launched as part of an attempt to secure more legitimacy, introducing an idea of “socialist democracy with Chinese characteristics”, though with the one-party leadership untouched (Heberer and Schubert 2006;

Dickson 2004). Village elections, for example, are being promoted in rural China, helping to solve problems at local level and gaining new legitimacy for the Party (Schubert 2009; Xu 1997). Even the central-local government relationship, where powers are decentralised, is utilised for the sake of consolidating government legitimacy. The current political arrangement that separates the central government from local governments' activities, for instance, keeps the central government from being blamed for repressing popular resistance (Cai 2008). Inner-party purity is also maintained through punishing local governments and officials for their wrongdoing (He 2000).

For some scholars, the efforts of the central government have been rewarded. For example, Cai (Cai 2008) argues that the multi-layered power structure of the government has saved the regime from collapse by leaving local governments to be held responsible for the failure in governance. However the main body of literature has developed a critique of these efforts, and hardly considers them as successful. These initiatives are thought to be limited by various factors including "the tensions between the central and local governments, institutional weaknesses, inconsistent policies, and the inability or unwillingness to undertake fundamental political reforms" (Lum 2006: 8). Scholars (e.g. Heberer and Schubert 2006; Dickson 2004) argue that the lack of a dominant ideology, the absence of substantial political

reforms leading to democracy and the persistence of social problems in reality, such as corruption and growing inequality and injustice, undermine the CCP's ability to retain its legitimacy, as these are the main threats to it. This provides some implications for the present study: that is, that it would cast more doubts on government legitimacy if online discussions of social protests touch on these issues rather than others.

Government legitimacy and online public participation

Scholars, who have examined whether online public participation has reinforced or worsened the government's problematic political legitimacy, have come to different conclusions. In general, scholarly views on this matter can be classified into three types. One view sees public participation on the web as a challenge and even a threat to the rule of the CCP (e.g. Zheng 2008; Yang 2009; Esarey and Xiao 2011). Another regards it dysfunctional or even as supporting the rule of the CCP, as a result of the CCP's control over the internet (e.g. MacKinnon 2008; Jiang 2006). The third perspective sees the function of online expressions varying according to whether the print media or online discussions have initiated the process and set the agenda for public attention (e.g. Hassid 2012).

Despite these disagreements, scholars agree that both the physical network and web content are under state control (Harwit and Clark 2001; Dong 2012). Some

scholars (e.g. MacKinnon 2008) even regard censorship and regulation as having made the control of the internet successful. The internet is seen as reinforcing the existing social structure and conditions, as it does not directly encourage the kind of political activism that might bring democracy to China. The primary aim of Internet censorship in China has been found to suppress social unrest and eliminate the possibility that online discussions will lead to collective action (Orcutt 2012). Posts about collective action or social protest are among the top banned topics that are required to be removed from the internet (Hunt 2012).

In the most recent period, social media sites have been playing an increasingly important role in China's social and political life, through means such as eroding government censorship (Wines and LaFraniere 2011), and providing a platform for internet users to call into question government policies, decisions and actions (MacKinnon 2012). Despite its importance, the content on social media sites such as Weibo, remains under-researched as a whole, let alone research specifically into Weibo posts about social protests. Few previous studies that have examined web content to understand the nature of online communication have mainly looked at blogs and forums (e.g. MacKinnon 2008; Yang 2003; Wu 2008; Zhou 2009; Hassid 2012), leaving the content of social media sites a virgin area for research. This gap in the existing literature limits our ability to understand the nature of online

communication and its challenge to government legitimacy. This study aims to fill the gap by analysing Sina's Weibo data about the Wukan and Haimen mass incidents, which was collected from December 2011-June 2012.

Examining social media

Weibo is a Chinese microblogging social media site where users disseminate and read short messages (up to 140 Chinese words) to/from the public or among a particular circle of contacts. Sina's Weibo was the first Weibo, appearing in 2009, and was immediately embraced by Chinese people. It had 195 million users by mid-2011. Various internet portals now have their own Weibo websites, among which Sina's and Tencent's Weibo are the most influential.

Compared to its English language counterpart Twitter, Weibo has been less examined, partly due to the language barrier presented by Chinese to Natural Language Processing (Chinese words are not separated by whitespace). Scholars (such as Pak and Paroubek 2010; Davidov, Tsur et al. 2010; Go, Bhayani et al. 2009; Diakopoulos and Shamma 2010; Mendoza, Poblete et al. 2010; Jansen, Zhang et al. 2009; Hughes and Palen 2009) have studied the content on Twitter, such as patterns, sentiment, and opinions, as well as the activities of Twitter users, and proposing methods for studying Twitter. Given the similar technical architecture of Twitter and Weibo, these previous researches provide a basis for this present study to develop its

research design, with the aim of answering the following four research questions:

- 1) What are the overall and over-time patterns of the activities of Weibo users, in terms of tweeting the two incidents?
- 2) What themes have emerged in the tweets about the two incidents and how did these change over time?
- 3) How does the online discourse of the two mass incidents develop over time?
- 4) Are there any differences between the two cases in terms of overall patterns, themes and development?

We have taken an exploratory research approach: we had no preconceptions about what the data might tell us and kept our research questions open-ended. The answers to these research questions have been found through the analysis of the data, from data query, data exploration, to eventually establishing contextual meanings in the discourse of the data.

Methodology

The two mass incidents

Mass incidents are events involving large numbers of people that can directly challenge government legitimacy. The Wukan incident refers to a series of demonstrations and petitions which villagers in Wukan, in Guangdong province, made from September- December 2011. They were unhappy principally because of

seizures of village land by corrupt local officials for sale to developers without proper compensation to the villagers. In December 2011, Xue Jinbo, a village representative, died in custody two days after his arrest. His death led to a severe confrontation between Wukan villagers and local police. The Guangdong province later showed a supportive attitude toward the Wukan villagers and assigned a working team to investigate local land and election issues on 20 December 2011. In February and March 2012, Wukan villagers elected their own village government.

The Haimen incident, also in Guangdong province, started with an appeal letter published on a web forum. It revealed that despite the dramatic pollution the first coal-fired electricity plant had caused there, Haimen's local government had decided to open a second plant of this kind. This letter aroused residents' fury over the local government's decision (Shangguan 2011). A message mobilising local residents to "protest in front of the Town government and take collective action" was reposted by many Weibo users (Zhang 2011). Thousands of local residents, many of whom were young people, including even teenagers, held protests from 20-23 December 2011.

Data Collection and sampling

Sophisticated tools (for data collecting and sampling, data storing and querying, as well as analysis and visualisation) have been developed and utilised in the

different stages of data analysis during this research. The data collection and sampling of Sina's Weibo³ tweets focused on the period between December 2011 and June 2012, in which the most dramatic and comparable trends were present in both incidents. The data collection was completed via Web request to the Weibo search API. This routine task was executed on a daily basis throughout the collection period with relevant keywords, which were first suggested by Weibo trends and then extended manually. The keyword list was manually amended, working around the search censorship of Weibo APIs. Experiments have been done to test the quality and reliability of data collection and prepare for the dataset that will be used for further analysis. The experiments have considered the most influential factors in the dataset, such as the popular trends, the active persons and the frequent word clusters. Noise has been minimised by deleting irrelevant topics and tweets. The possible impacts of noise to the overall findings are therefore seen as insignificant. The analysis also shows the high quality of the dataset as the development of the discourse very much fits that of the events. The final fully-examined dataset includes 278,980 tweets.

Data analysis and visualisation

The data analysis consists of two parts. One is the data-driven content analysis, which examines the content and makes reference from the content to the context (Tian and Stewart 2005; Popping 2000). This research method has been used in

analysing a variety of media content including internet, broadcasting, and newspaper content (e.g. Hassid 2012; Kirilenko and Stepchenkova 2012; Tian and Stewart 2005).

Second is the data-driven semantic analysis of tweets and of the activities of twitter users, in the way that the studies on Twitter content cited above have done. The data-driven analysis was deployed in these studies not only because of its efficiency over human manual coding, but also because of the higher reliability and functional flexibility of data processing.

The data storing and querying was driven by the analytical cases, i.e. the key information demanded from the data collection. MySQL was used to discover the tweet/user and tweet/retweet relationships, while Solar⁴ was used to perform the Natural Language Processing and faceted search function, i.e. fact numbers across timeline or categories, and SPSS was employed for clustering analysis. The analysis and visualisation tools used in the process enable the specific data insights over the data stored by illustrating the influential factors, e.g. frequency analysis and scheme clustering.

There is a three-step operation in the data analysis. First, the quantitative data insight revealed the basic facts of Weibo activities, such as the top 100⁵ most active users ranked by the amount of posts they made and grouped by profile types, the ratio of original tweets against retweets, and the top 100 most popular retweets, in

order to answer Q1. Second, the content analysis of Weibo tweets answered Q2 and Q3. The data query crossed all tweets to develop the frequency list of words. And then a group of words with the top 100 frequencies were statistically clustered by the correlating possibilities within the same tweet. The approaches of hierarchical clustering (creating dendrograms using Ward's method) and networking analysis were used. The clustering and networking analysis, which has been used in previous studies (such as Murphy 2001; Tian and Stewart 2005; Balasubramanyan, Lin et al. 2010), offered an understanding of themes emerging and semantic relations among these words. The event developments include several phrases. Clustering analysis for each phase was conducted to find more subtle themes emerging in different stages to identify the changes over time in the discourse on Weibo. Thirdly, research findings for the two incidents were compared, to measure the similarities and differences, in order to answer Q4.

Findings: two modes of online communication

An overall image

Overall, communication on Weibo about the two incidents has been very active. During the sampling period, 159,902 internet users have posted a total number of 278,980 tweets about the two incidents with a mean of 1.52 tweets per user. In addition, the number of retweets (186,545) was about twice the number of original

tweets (92,435). The Wukan case (Wukan for short in this section) saw much more active user participation than the Haimen case (Haimen for short in this section), in respect both to the numbers of tweets and the number of participants. 136,907 users posted 243,309 tweets in Wukan, while 35,671 tweets were sent by 22,995 users in Haimen.

Four types⁶ of internet users - local residents or someone who has personal connections to the place, elites⁷, news organisations (both new media companies such as Sina and traditional media such as newspapers) and ordinary people whose identity cannot be identified - are found to be most active users (see Figure 1 and Figure 2). Notably most traditional news organisations participating are from the print media rather than television. Among the top 100 most active users, local residents and ordinary people were responsible for the majority of the posts: around 72% in Wukan and 81% in Haimen. They were more likely to retweet posts made by other users rather than tweet original posts made by themselves.

By contrast, elites and news organisations were more likely to send their original messages. In both cases, except for only 22 messages in Wukan, most tweets posted by news organisations were original. Elites behaved differently in the two cases. Elites tweeted and retweeted a nearly equal number of tweets in Wukan, but only posted the messages they themselves generated in Haimen. That is to say, more

messages produced by other users were regarded as credible by elites and news organisations in Wukan than in Haimen.

The distribution of tweets in the two incidents displays similarities as well as differences in their trends, viewed on a weekly basis (see Figure 3). On one hand, Weibo communication in both incidents reached a peak at December 22, 2011. This date was crucial, as it marked the final day of consecutive protests in Haimen and of confrontation between the villagers and the authorities in Wukan. On the other hand the attention given to Haimen was rose sharply to a peak and then fell abruptly, dropping to trough within a short period of time, while the attention to Wukan lasted for a longer period of time, with ups and downs within the overall pattern. The longer the online attention lasted, the stronger the impact it potentially had on the real world.

Attention to both cases on Weibo showed variations in accordance with the development of actual events over the sampling time. The peaks more or less match the turning points of the events. As noted already, in Wukan the highest peak occurred on December 22, 2011, when the government came to negotiate with the villagers. Other peaks corresponded to the start of the December protest, the election of the village committee and the village elections (respectively on December

7, 2011, February 01, 11, and 16, 2012). In Haimen, the highest peak matched the occurrence of the protests from 20-23 December, 2011.

Among the top 100 most active users, Wukan has fewer profiles that no longer exist (around 6.4%) than Haimen (around 11.4%). It is difficult to trace why these accounts no longer exist. If the main reason why these profiles have disappeared is due to censorship by the Chinese government, then, one could argue that Haimen has experienced a higher level of censorship than Wukan. But, indeed, in the data collection stage, we found censorship changed over time, sometimes lighter, but sometimes tighter.

In general, the Weibo discussion about the incidents presents two modes of online communication. Online communication about Wukan was both led by local population elites and featured themes with a national perspective, enjoying long-term attention from online users and two-way interaction between elites and the local residents of Wukan. Communication on Weibo about Haimen, however, was led by the local population, transient, locally-focused and characterised by one-way communication from local residents to outsiders.

Wukan on Weibo: populace and elite-led two-way initiation of communication

Communication on Weibo in Wukan was elite-led for two reasons. One reason is the active participation of elites and news organisations in the discussion. The

second reason is the association of elite names with the themes of the discourse, which implies the role of elites in giving the Weibo discourse national meaning. The elite-led nature of the discussion suggests that the online communication was initiated by both local residents and those from outside.

Elites and news organisations both appeared to be active in the Wukan discussion. Nearly a third (31.3%) of the top 100 most active users and two out of the top 10 most active users (see Table 1) were elites and news organisations. Despite the fact that ordinary people were most active type of users, elites were the category of users whose messages were most frequently retweeted by other users (see Figure 4). Users preferred to retweet messages created by elites (accounting for 60%), followed by those by local residents (14%) and news organisations (13%). The more intensive attention from elites and news organisations in Wukan gives the incident a more national dimension.

Wukan has a discourse with focuses on keywords such as “democracy”, “enlightenment”, “reform”, “rights”, “hope” and “China”. Its seven clusters (see Table Table 4) emphasise the aftermath (demands for democracy and reform) and meaning of the incident rather than the incident itself. No themes suggested by the clusters were related to the demonstrations by Wukan villagers and their clashes with local government or the police. Except for one theme on the actual election, the majority

of the clusters are about democracy, reform, and the historical significance of this event, suggesting that the solutions to the problems raised by Wukan and the guarantee of the protection of rights are democracy, reform and the achievement of local autonomy. The themes in Wukan include the demands of the protesters that are about “democracy”, “reform”, “election”, “rights protection” and “justice”. In addition, the Wukan discourse on Weibo is plural and has historical and contextual meanings. For example, on one theme that examines the historical meaning of Wukan, the discourse relates Wukan to what happened in Xiaogang Village in the 1980s, a symbol of China’s economic reform.

These themes are also associated with the names of members of the elite, such as “Xiong Wei”, “Xue Manzi”, “Xiao Shu”, “Han Zhiguo”, which appear in the clusters of “democracy”, “reform”, “elites”, and “meaning”, and are associated with keywords such as “democracy”, “election”, “enlightenment”, “reform”, “change”, and “Xiaogang Village”. The name of the Guangdong provincial governor, Wang Yang, also appeared in the solution cluster. This association between elites and the discourse of democracy is prominently displayed in different phases of the incident.

The distribution of tweets over time in Wukan (see Figure 3), matching the development of the incident itself, is divided into four phases: protest (before December 20, 2011), negotiation (from December 20, 2011 to January 21, 2012),

election (from February 1 to March 3, 2012), and post-election (from March 4, 2012 afterwards). The themes in the protest phase are unclear, merely suggesting two focuses on the protest itself and on Weibo functionality, such as retweeting (see

Table 6). No keywords are relevant to democracy or election or elites at all in the top 100 keywords. When it comes to the negotiation phase, themes of “literacy”, “election”, “democracy”, “demand” and “justice” emerged and keywords about elites such as “Yu Jianrong”, “Xiong Wei”, “Xue Manzi”, “Xiao Shu”, “lawyer”, started appearing in the top 100 keywords. The theme of “literacy” was prominent throughout the last three phases. In the theme, members of the elite, such as “Xiaoshu” and “Xue Manzi”, mobilised Web users to make donations to create a “book fund” for Wukan villagers in order to increase their reading abilities and literacy, which was seen as crucial for achieving democracy and enlightenment. This discourse of democracy becomes stronger in the third phase- of election. The theme of “historical meaning” and new keywords of elites “Wu Jiaxiang” and “Han Zhiguo” appeared. The keywords of “Wang Yang”, the Guangdong governor, and “Liu Jianfeng”, a journalist, appeared in the final post-election phase. The themes in the second, third, and fourth phases remain similar, about democracy. And the theme of “literacy” is also connected with the themes of “democracy”, “election”, and “historical meaning”.

Haimen: populace-led one-way initiation of communication

By contrast, Haimen is more like a local rather than a national event. The lack of enthusiasm of elites and news organizations for the topic in contrast to the level of

activity local residents in Haimen suggests the Weibo communication about this event is led by the local population and is characterized by the initiation of communication from local residents to the outside world in one direction only.

In Haimen, local residents were the type of users who were most active, as well as those producing the messages which were mostly retweeted by other users. Only around 3% of the top 100 most active users in Haimen were members of the elites and news organisations. The messages that were most likely to be retweeted are produced by local residents (29, 29%), followed by ordinary people (19, 19%) and then elites (18, 18%). All the top 10 most active users were local residents and ordinary people (see

Table 2).

The themes in Haimen are from a local perspective, emphasising keywords such as “protest”, “pollution”, “oppose”, “town government”, “environment” and “solution”. As illustrated in Table 3⁸, the seven clusters in Haimen reflect an emphasis on the mass incident itself: a mass incident happened in Haimen (cluster 1), its cause: a concern for possible pollution from the second power plant (cluster 2) and on-site scenes (clusters 3 and 5): such as, people blocking the Shenshan highway, dialogues being sought, tear gas being used, police and armed police arriving on the scene, and “mothers” (old female Haimennese, according to relevant news reports) appearing as a symbol of support. In addition, Weibo messages also appealed for a solution in cluster 6. The demands of protesters in Haimen were limited to “solving” the problem and “opposing” the establishment of the power plant. No elite names were mentioned in Haimen, except the name of a newspaper journalist who was on the scene, Li Yong, which appeared in the cluster of “on-site scene”.

The distribution of tweets in Haimen (see Figure 3) suggests three phases in its development: rising (before December 23, 2011), declining (from December 23, 2011 to January 6, 2012), and fading-out (after January 7, 2012). A detailed clustering analysis for each phase suggests the discourse of the first two phases is about the incident itself and on-site scenes, while the discourse of the third phase is driven by

mainstream media coverage of government opinion about this incident (see **Error! Reference source not found.**). In the third phase, most tweets were retweeting a link to a news report rather than posting self-generated content by Weibo users. The keywords in cluster 1, 2 and 3 are highly consistent with the news report⁹ covered in the Guangzhou Daily on January 15, 2012. The discourse in the first two phases also mentioned “Wukan”, which shows a sort of link between these two cases in the discourse of social unrest on Weibo. Few elite names were mentioned here. Only the names of “Wang Keqing” and “Xu Xing” were mentioned, and this was in the context of the incident itself.

The role of Weibo

Notably, the analysis implies the importance of Weibo in publicising the two cases. One theme common to the two incidents concerns the functionality of Weibo itself, which often appears with key words such as “attention”, “retweet”, “share” and “support”, as suggested in the clusters. Both cluster 1s that include keywords of the events themselves as well as “http” and “cn” imply that the discourse about what happened is often associated with web links.

Discussion and conclusion

Different to the arguments raised by the previous studies discussed in the literature review, the present study regards to what extent Weibo communication

can shake government legitimacy depends on the actual situation, rather than being homogeneously determined by the technology itself.

It is clear from the analysis above that Weibo functioned as a platform for protestors to express their demands and for other Internet users to communicate their views. Local residents actively disseminated what was happening in the two locales. They spelled out their demands and voiced their concerns, breaking down the control over the terms of the debate by local authorities. This function has facilitated the publicising and dissemination of information about politically sensitive events, such as these mass incidents, and won the attention of wider society. Such attention has given “oxygen” to social protests as well as helped to connect different social protests together, and increasing their capacity to imitate others’ tactics and successes.

State control of the Internet is apparently ineffective in preventing the circulation of news and discussions about social protests among the public, at least in these two cases. Absolute control over information is seemingly impossible in the digital environment. Internet users enjoy certain autonomy in communicating information. People from all parts of the country can be informed about the things that exist beyond their immediate environment. This autonomy certainly does not mean that information control by the authorities in China has completely

disappeared. The level of information control largely varies from case to case, as has been found in the two cases examined here.

The two incidents involved protests against the wrongdoing of local governments or officials, occurring nearly simultaneously in the same province. Online communication about the two cases, however, reveals significant differences. The length and intensity of online attention, the types of participants or the themes in discourse all signified the Haimen case as a transient populace-led local event with local demands and concerns, but revealed the Wukan case to be a long-lasting event, led by the local population as well as members of elites, producing demands for democracy and a critique of the current political system in China. Enjoying a longer span of online attention, the Wukan case gave rise to a discourse beyond purely local demands and reflected more or less the discourse in the wider society, while the Haimen incident has a discourse from a local perspective, reflecting merely local demands.

The online communication in the two cases did not follow the same pattern. In the Wukan incident, the communication mode was a two-way one between protesters (local residents) and outsiders. From the second stage of its development, online communication about the Wukan events became led by elites from outside the village. However, in the Haimen incident, the communication remained mostly

one-way from protesters (local residents) to outsiders. The communication in this case was mainly led by the local populace as they wanted to attract the attention of the wider society, while the notice paid to the events by members of the elite was short-lived and weak.

What made the situations different was how internet users employed Weibo, for what reasons and with what interests. In theory, the technology of Weibo appears to be neutral: it is available for anyone who has access to the platform and the willingness and skills to use it. In practice, however, this is not the case. Differences resulted from the level of attention of online users, their interpretation of the events and the discourse of the events forged on Weibo. More specifically, differences in local residents' new media use and organisation skills in the two localities, as well as in the attention level of elites and news organisations, accounted for the variance in the two cases.

The strong organisation of the Wukan activists has been attributed to young people in their twenties and thirties who returned from their urban jobs for the protest (Zhang and Lu 2011; Zhang 2011). They used new media technologies, such as QQ, DV, Tudou (Chinese version of Youtube), forums, BBS and Weibo, skilfully and fluently for organising and communicating the protests (Zhang and Lu 2011). Over the three month period of the protests, these young people developed good skills in

managing the news media and strategically publicising information. For example, the Weibo messages did not highlight the conflicts between the villagers and the authorities. It was better to steer clear of such topics and instead focus on the villagers' desire for peaceful problem-solving and avoiding possible crackdowns. The less-organised Haimen protest, by contrast, lasted for only a few days. Local residents, of whom the most prominent elements were teenagers and youth in their 20s, used the Internet to disseminate information about their concerns over environmental problems, demands and the development of their protests. But they were still learning (and needed time to learn) how to organise their protest activities, and the short duration of the protest did not allow them to do so (MingPao 2011).

Online users' attention to an event is easily distracted by other subsequent events. In the Wukan case, where the original event was local but influences the national debate, intense attention from elites and news organisations is crucial for keeping online discussion alive and keeping the events in the national consciousness. The discourse in this case - of democracy and reform - moved it beyond local demands. Such a discourse in fact is the discourse of intellectuals, as it is highly associated with posts published by members of the elites and news organisations. They act as opinion leaders, shaping the development of the discourse and targeting political reform at the national level. On the other hand, in the Haimen case, elites

and news organisations seldom paid attention and the trend for online discussion was transient. The discourse centred mainly on local demands, requesting the local government to stop building the second coal-fired power plant and to release those who had been arrested in Haimen. Of course, we cannot be certain whether it is the participation of elites and news organisations that has led to the national importance placed on any local mass incident or whether the severity of the local mass incident and its potential national significance attracted the attention of elites and news organisations. This is a question of chickens and eggs.

In terms of the challenge to government legitimacy, the Haiman discourse is less challenging than the one of Wukan for two reasons. First, in the case of Haimen, there was only a critique of local government and an attempt to correct local wrongdoing. Local demands that do not correspond to the interests of other internet users have little revolutionary resonance. Central government appears to be separate from local governments and thus is more or less immune from demands focused on local issues. Second, the legitimacy of the Chinese government can even be reinforced by punishing local governments and focusing people's anger toward local officials. Local people are willing to appeal to central government in order to obtain justice. Through punishing local governments and correcting their wrongdoing, central government is able not only to establish a reforming image for itself but also

to reduce the social grievances of the people. Functioning as a safety valve that releases the pressure of the complaints of ordinary people, this in turn helps the solidity and maintenance of government legitimacy.

The discourse that is associated with political reform and democracy in the Wukan case, however, brings some big and fundamental challenges to the central government's legitimacy. Such demands touch on the very principles underlying the problems in China. The demand for democracy threatens the legitimacy of the Chinese government, while the requirement to increase villagers' literacy lays the basis for a greater degree of political transparency by allowing people to see through government attempts at obscurantism. This is a forbidden zone that the Chinese government has tried to prevent anyone from moving into since the 1989 political trauma over the Tiananmen protests. Discourses like this can resonate with the wider public rather than be limited to locales. Despite having the possibility of consolidating government legitimacy, when it comes to political reform, the village election also has some potential to undermine that legitimacy. When the discourse began associating mass incidents with democracy and protection of human rights rather than focusing on the battle with the police, doubts began to be raised about the legitimacy of the ruling government. The stronger the voices for democracy are, the greater the challenge to government legitimacy is.

Therefore, McLuhan's technologically determinist argument that the "medium is the message" is not really true in this case. What is meaningful is what communicated and who is participating in the communication. It is not meaningful to look into individual mass incidents in an isolated fashion. It is necessary to consider the social dynamics behind them and the different roles various actors have played in the process. Specifically, the appearance of Weibo has offered a means whereby communication is difficult for the authorities to control. Because of this, people have access to more information than before, which in turn might trigger further changes in society. This potentially spiralling influence of Weibo on society cannot be ignored.

Limitations and implications

The study which has been presented in this article suggests that the online world more or less functions as a "mirror" of the offline real world. The online virtual world now is part of our real life. We need to treat what happens in the virtual world as if it were in our real world. In other words, we cannot ignore what happens online and assume that it is separate from our real life. Our understanding however is still limited by the tools we have to analyse the massive amount of data generated online, which is also highly transient in some important ways. We definitely need to improve the tools and methods for a better understanding of this data.

The present study, led and driven by a huge amount of Weibo data, conducted

user activities' analysis and clustering analysis in order to find significant patterns in Weibo communication. In future studies, more sophisticated computational functions are needed for identifying networking patterns, the relationship between trends in discourse and users' activities and to examine the sentiments of messages. Despite its limitations, this present study suggests a research model that is able to test such discourses and communication patterns within a very large amount of Weibo data.

Bibliography

AFP (2011). China struggles to tame microblogging masses. The Independent. London.

Balasubramanian, R., F. Lin, et al. (2010). Node Clustering in Graphs: An Empirical Study. NIPS 2010 Workshop on Networks Across Disciplines in Theory and Applications. Vancouver, B.C., Canada.

Cai, Y. (2008). "Power Structure and Regime Resilience: Contentious Politics in China." British Journal of Political Science **38**(3): 411–432.

Chen, J., Y. Zhong, et al. (1997). "Assessing political support in China: Citizens' evaluations of governmental effectiveness and legitimacy,." Journal of Contemporary China **6**(16): 551-566.

Davidov, D., O. Tsur, et al. (2010). Enhanced sentiment learning using Twitter hashtags and smileys. COLING '10 Proceedings of the 23rd International Conference on Computational Linguistics: Posters

Diakopoulos, N. A. and D. A. Shamma (2010). Characterizing debate performance via

aggregated twitter sentiment. Proceedings of the 28th international conference on Human factors in computing systems New York.

Dickson, B. J. (2004). Dilemmas of Party adaptation: the CCP's strategies for survival. State and society in 21st-century China: crisis, contention, and legitimation. P. H. Gries and S. Rosen. New York and Oxon, RoutledgeCurzon.

Dong, F. (2012). "Controlling the internet in China: The real story." Convergence: The International Technologies Journal of Research into New Media **OnlineFirst Version**.

Downs, E. S. and P. C. Saunders (1998/99). "Legitimacy and the Limits of Nationalism: China and the Diaoyu Islands." International Security **23**(3): 114-146.

Esarey, A. and Q. Xiao (2011). "Digital communication and political change in China." International Journal of Communication **5**: 298-319.

Fewsmith, J. (2008). An "Anger-venting" mass incident catches the attention of China's leadership. The China Leadership Monitor.

Gilboy, G. and B. Read (2008). "Political and Social Reform in China: Alive and Walking." The Washington Quarterly **31**(3): 143-164.

Gilley, B. and H. Holbig (2009). "The Debate on Party Legitimacy in China: a mixed quantitative/qualitative analysis." Journal of Contemporary China **18**(59): 339-358.

Go, A., R. Bhayani, et al. (2009). Twitter Sentiment Classification using Distant Supervision. CS224N Project Report, Stanford, Technical Report.

Guo, B. (2003). "Political Legitimacy and China's Transition." Journal of Chinese Political Science **8**(1&2): 1-25.

Harwit, E. and D. Clark (2001). "Shaping the Internet in China: Evolution of Political Control over Network Infrastructure and Content." Asian Survey **41**(3): 377-408.

Hassid, J. (2012). "Safety Valve or Pressure Cooker? Blogs in Chinese Political Life." Journal of Communication **62**: 212-230.

He, Z. (2000). "Corruption and anti-corruption in reform China." Communist and

Post-Communist Studies **33**(2): 243–270.

Heberer, T. and G. Schubert (2006). "Political Reform and Regime Legitimacy in Contemporary China." ASIEN **99 April**.

Holbig, H. (2006). Ideological Reform and Political Legitimacy in China: Challenges in the Post-Jiang Era. GIGA Research Program: Legitimacy and Efficiency of Political Systems. G. G. I. o. G. a. A. S. L.-I. f. Globale and u. R. Studien. Hamburg.

Hughes, A. L. and L. Palen (2009). "Twitter adoption and use in mass convergence and emergency events." International Journal of Emergency Management **6**(3/4): 248 - 260.

Hunt, K. (2012). China tightens grip on social media with new rules. CNN.

Jansen, B. J., M. Zhang, et al. (2009). "Twitter power: Tweets as electronic word of mouth." Journal of the American Society for Information Science and Technology **60**(11): 2169-2188.

Jiang, M. (2006). Authoritarian deliberation on Chinese Internet. In Search for Deliberative Democracy in China. E. L. a. B. He. New York, Palgrave Macmillan.

Kirilenko, A. P. and S. O. Stepchenkova (2012). "Climate change discourse in mass media: application of computer-assisted content analysis." Journal of Environmental Studies and Sciences **2**(2): 178–191.

Lum, T. (2006, June 22nd, 2012). "Social unrest in China ". from <http://digitalcommons.ilr.cornell.edu/crs/19/>.

MacKinnon, M. (2012). Chinese 'netizens' intervene where judges fail. The Globe and Mail.

MacKinnon, R. (2008). "Flatter world and thicker walls? Blogs, censorship and civic discourse in China." Public Choice **134**: 31–46.

Mendoza, M., B. Poblete, et al. (2010). Twitter under crisis: can we trust what we RT? Proceedings of the First Workshop on Social Media Analytics New York, 71-79.

MingPao (2011). No good organisation in the Haimen protest that is different from Wukan. Ming Pao. Hong Kong.

Murphy, P. (2001). "Affiliation Bias and Expert Disagreement in Framing the Nicotine Addiction Debate." Science, Technology, & Human Values **26**(3): 278-299.

Orcutt, M. (2012). A Peek Behind China's 'Great Firewall'. Technology Review.

Pak, A. and P. Paroubek (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. Proceedings of LREC.

Popping, R. (2000). Computer-assisted text analysis . . Thousand Oaks, CA, Sage.

Potter, P. B. (1994). "Riding the Tiger: Legitimacy and Legal Culture in Post-Mao China." The China Quarterly **138**: 325-358

Schubert, G. (2008). "One-Party Rule and the Question of Legitimacy in Contemporary China: preliminary thoughts on setting up a new research agenda." Journal of Contemporary China **17**(54): 191-204.

Schubert, G. (2009). Village elections, citizenship and regime legitimacy in contemporary rural China. Regime Legitimacy in Contemporary China: Institutional Change and Stability. T. Heberer and G. Schubert. New York, Oxon, Routledge: 55-78.

Shangguan, L. (2011). Coat Electricity Plant Project is Haimen Shantou Questioned. Southern Weekend. Guangzhou.

Shue, V. (2004). Legitimacy crisis in China? State and Society in 21st-Century China: Crisis, Contention, and Legitimation. P. H. Gries and S. Rosen. New York, Routledge: 24-49.

Tian, Y. and C. M. Stewart (2005). "Framing the SARS Crisis: A Computer-Assisted Text Analysis of CNN and BBC Online News Reports of SARS." Asian Journal of Communication **15**(3): 289-301.

Wines, M. and S. LaFraniere (2011). In Baring Facts of Train Crash, Blogs Erode China Censorship. The New York Times. New York.

Wong, L. (2004). "Market reforms, globalization and social justice in China." Journal of Contemporary China **13**(38): 151-171

Wu, W. (2008). "Measuring Political Debate on the Chinese Internet Forum." Javnost-the Public **15**(2): 93-110.

Xu, W. (1997). "Mutual Empowerment of State and Peasantry: Grassroots Democracy in Rural China " World Development **25**(9): 1431-1442.

Yang, G. (2003). "The Co-evolution of the Internet and Civil Society in China." Asian Survey **43**(3): 405-422.

Yang, G. (2009). The Power of the Internet in China: Citizen Activism Online. New York, Columbia University Press.

Zhang, J. (2011). Wukan Village: a village paradigm of social movement. Sun Affairs. Hong Kong.

Zhang, J. and W. Lu (2011). Group of Enthusiastic and Patriotic Wukan Youth. Sun Affairs. Hong Kong.

Zhang, Q. (2011). Thousands Protesters Blocked Road to Protect Rights in East Guangdong. Asian Weekly (yazhou zhoukan).

Zheng, Y. (2008). Technological empowerment: The Internet, state, and society in China. Stanford, CA, Stanford University Press.

Zhou, X. (2009). "The political blogosphere in China: A content analysis of the blogs regarding the dismissal of Shanghai leader Chen Liangyu." New Media and Society **11**(6): 1003-1022

Appendix 1: Figures and tables

Figure 1: Sum number of tweets by top 100 most active weibo users in the Wukan case

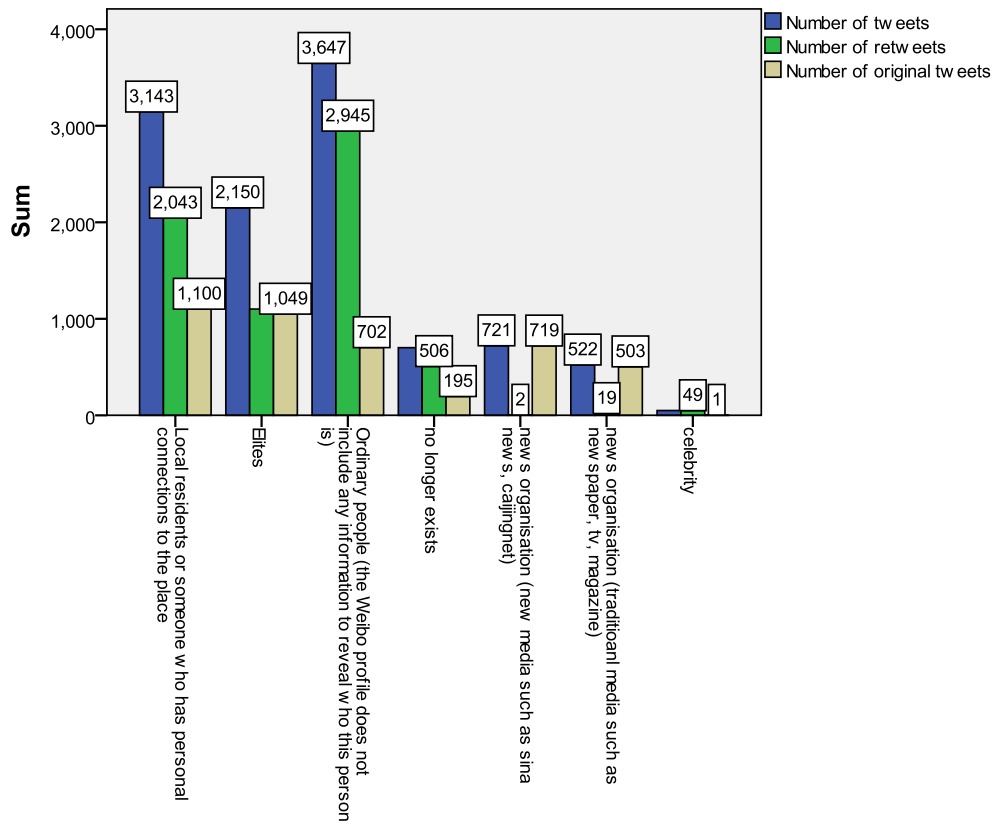


Figure 2: Sum number of tweets by top 100 most active weibo users in the Haimen case

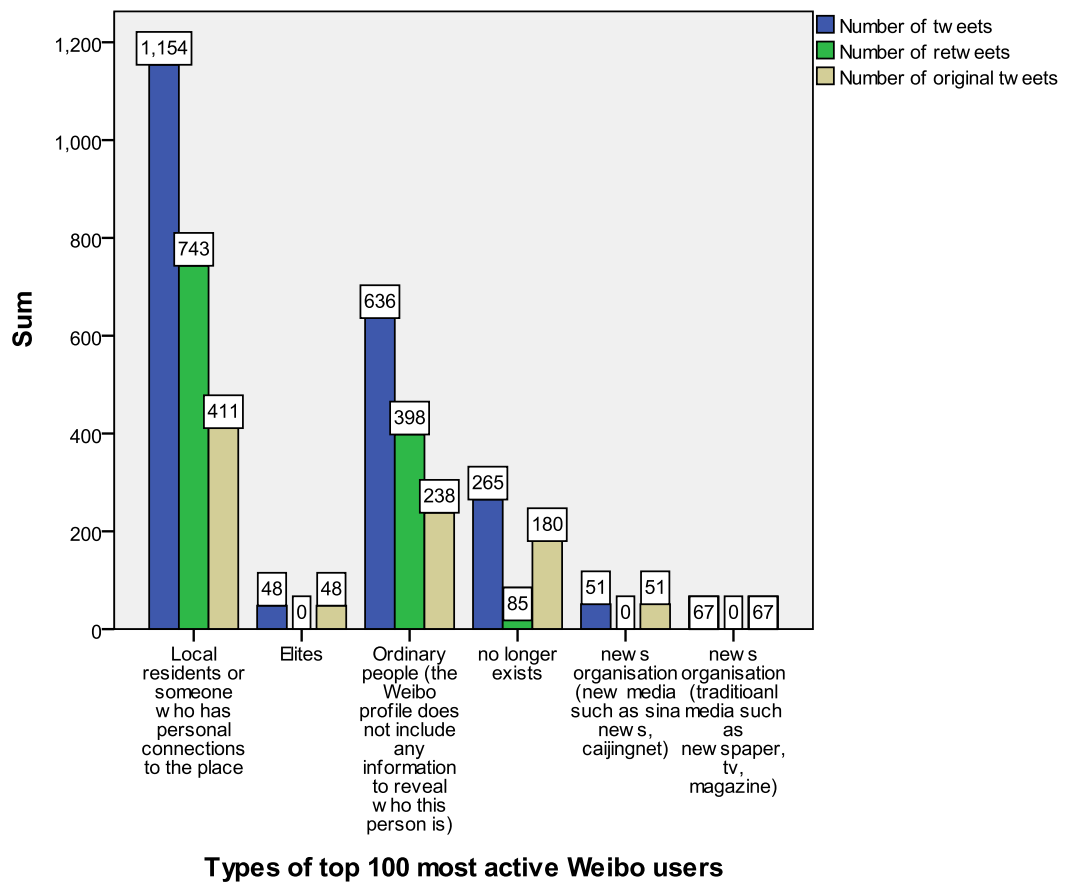


Figure 3 The overall distribution of tweets over time on a weekly basis in the Wukan and Haimen Cases

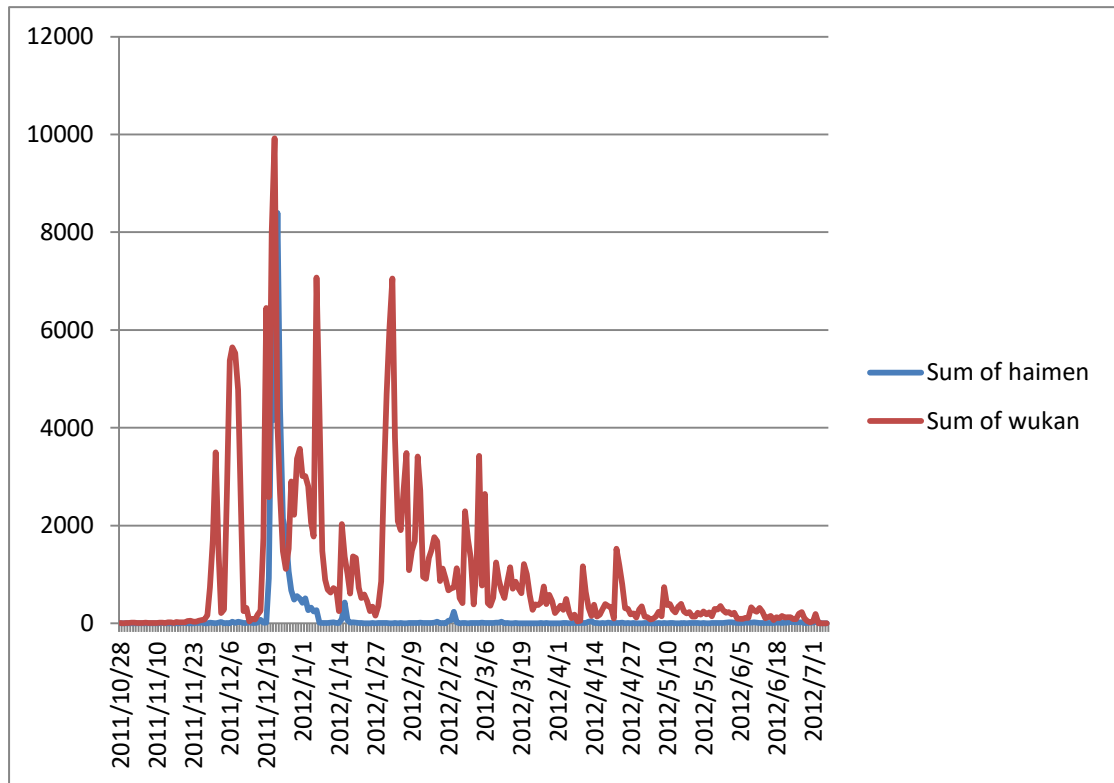


Figure 4 The comparison of types of users between “top 100 most active users” and “users who have produced top 100 most frequently retweeted posts” in the Wukan and Haimen cases

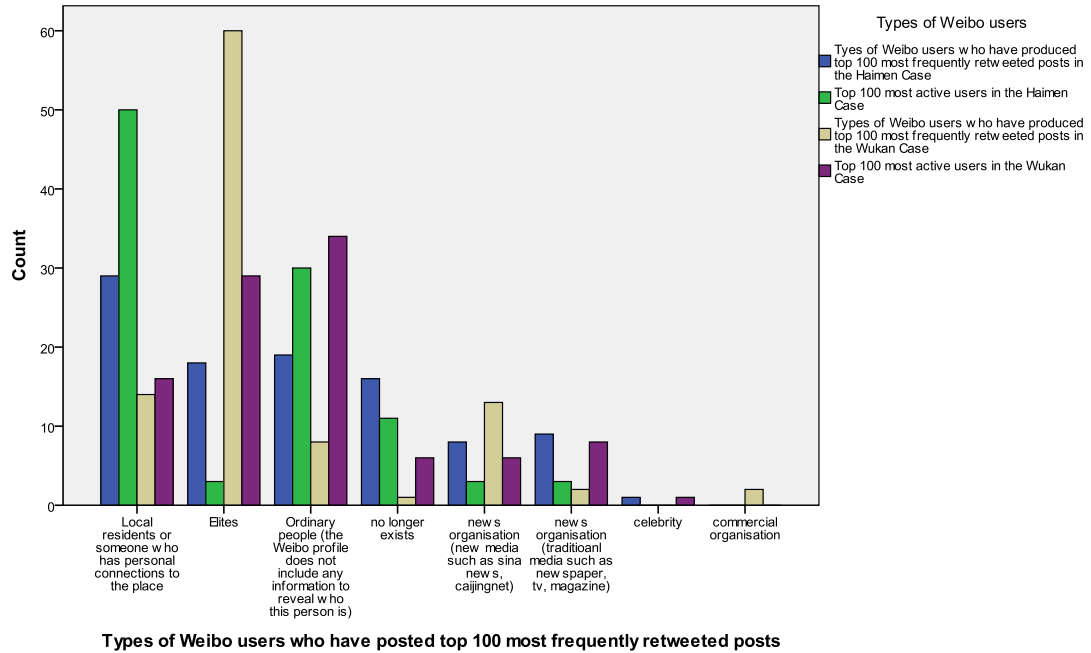


Table 1 The top 10 most active Weibo users¹⁰ in the Wukan Case

No.	Number of tweets	Types of Weibo users
1	1345	Local Wukan villager or someone who has personal connections to Wukan
2	525	Elites, including culture elites (such as journalists, university academics and writers) and economic elites (such as CEOs and other business people)
3	477	Ordinary people (the Weibo profile does not include any information to reveal who this person is)
4	368	Local Wukan villager or someone who has personal connections to Wukan
5	266	Ordinary people (the Weibo profile does not include any information to reveal who this person is)
6	258	Profile no longer exists
7	227	Local Wukan villager or someone who has personal connections to Wukan
8	197	News organisation (new media such as sina news, caijingnet)
9	186	Profile no longer exists
10	177	Ordinary people (the Weibo profile does not include any information to reveal who this person is)

Table 2 The top 10 most active Weibo users in the Haimen Case

Number of tweets	Types of Weibo users
68	Local Haimen resident or someone who has personal connections to Haimen
55	Local Haimen resident or someone who has personal connections to Haimen
47	Local Haimen resident or someone who has personal connections to Haimen
46	Local Haimen resident or someone who has personal connections to Haimen
43	Local Haimen resident or someone who has personal connections to Haimen
41	Ordinary people (the Weibo profile does not include any information to reveal who this person is)
40	Local Haimen resident or someone who has personal connections to Haimen
40	Ordinary people (the Weibo profile does not include any information to reveal who this person is)
37	Local Haimen resident or someone who has personal connections to Haimen
37	Ordinary people (the Weibo profile does not include any information to reveal who this person is)

Table 3 Haimen clusters

Clusters	Cluster Theme	Keywords
Cluster1	Event	Haimen, Shantou, Guangdong, http, cn, mass incident, happen
Cluster2	Cause	pollution, project, environment, power plant, second, town government, government, don't want
Cluster3	On-site scene 1	Block, traffic, share, tear bomb, they, dialogue, gather, peasants, again, Shen(zhen)Shan(tou), protest (shiwei), masses who do not get to know the truth (buming zhenxiang de qunzhong), media, report
Cluster4	Weibo	News, retweet, people (minzhong), we, people (renmin), mayor, Wukan, attention, one, Haimenneses, Weibo, (guanzhu)
Cluster5	On-site scene 2	East Sea, Hongkong, armed police, news, villagers, coal-fire, protest (kangyi), police
Cluster6	Demand	China, demand, solve (jiejue), problem, request, today (jinri), continue, handle (chuli), serious
Cluster7	On-site scene 3	on-site scene, develop, support, Li Yong, mothers, today (jintian), now, see, reply (huifu), haha

Table 4 Wukan clusters

Clusters	Cluster Themes	Keywords
Cluster1	Democracy and Weibo	Election, Xiong Wei, China, democracy, http, cn, Enlightenment, support, one, retweet
Cluster2	Weibo and Reform	Weibo, reform, attention, do not have, we, they, people (renmin), increasing pace, society, problem, solve, land, representatives, need, now, no, hope
Cluster3	Elites	Xue Manzi, change, Xiaoshu, one time (yici), everybody (gewei)
Cluster4	Meaning1	Quality, young people, Xiaogang village, important, basic levels, start, peasants, Han Zhiguo
Cluster5	Election	Village committee, vote, share, news, work, secretary, village, politics
Cluster6	Justice	Interests (liyi), lawyers, Xue Jinbo, worth (zhide), today (jintian), results, rights (quanli)
Cluster7	Solution	Development, politician, autonomy (zizhi), local, solve (chuli), economy, together, masses (baixing), Hongkong, journalists, haha, Wang Yang, people (minzhong), rights protection, history, media

Table 5 Haimen clusters in the three phases

Phase1 Rising:

Clusters	Cluster Theme	Keywords
Cluster1	Event	Mass incident happened in Haimen Town Shantou City
Cluster2	Event	Masses (qunzhong), Top news, share, traffic, Wang Keqing, mass incident, happen, town government, http
Cluster3	Cause and on-site scene	Do not understand truth (buming zhengxiang de), high way, power plant, pollution, people (renmin), coal-fire, protest, demonstration, build, oppose, second, people (minzhong), report, Shen(zhen)Shan(tou), environment, aggregate, project, worry, hundreds, part
Cluster4	Government reaction reported by mainstream media	Early stage, China energy, proof, delude, yesterday, work, morning, result in, city government, daily newspaper, value, appeal, decision city government committee
Cluster5	Weibo	Retweet, ChaoShan, electricity plant, Haimennese, we, Weibo, see, do not want, know, pay attention to, China, hope, event, Wukan
Cluster6	On-site scene and demand	Today, one, media, occupy, villagers, armed police, yesterday, solve, reply, thing, so, right protection, already, Hong Kong
Cluster7	Rumour	News, beaten to death, students, they, say, police, problem, today, this, now, angry

Phase2 Declining:

Clusters	Cluster Theme	Keywords
Cluster1	On-scene site and cause	Again, Shen(zhen)Shan(tou), throw, peasants, must, tear bomb, power plant, truth, second, recently, wisdom, dissatisfied, direct, corruption
Cluster2	Event and cause	Do not want, power plant, town, report, protest, masses (qunzhong)
Cluster3	On-scene site	Dialogue, high way, tear bomb, media, aggregate (juji), they
Cluster4	Weibo	Big, purpose, cn, retweet, Guangdong, people (renmin), people (minzhong), Wukan, attention, one, government, we, http, before, Shantou, Weibo, Haimennese, event
Cluster5		Same, continue (jixu), WK, Guangzhou, efforts, please, continue (chixu), handle (chuli), organize, opinion, my, on-scene site, serious, Yu Jianrong, come, agree, than, East
Cluster6		Science, Xu Xing, project, environment, news
Cluster7		Beat, fact, women, Hong Kong, Li Yong, East sea, see, police, Anti-riot, armed police, photo, mother, news
Cluster8		Request, pollution, villagers, victory, hope, now, China, Chaoyang, Beijing, development, support, oppose, Chaoshan, shanghai, freedom, Haichao association, today, hah, reply, know

Phase 3: Fading out

Clusters	Cluster Theme	Keywords
Cluster1	Mayer's opinion reported by mainstream media	Have to arrest (feizhua buke), Zheng Renhao, some people, masses, these people, intentions unpredictable, influenced by, media, influence, can not, mayer, long talk, few people, oversea
Cluster2	Mayer's opinion reported by mainstream media	Problem, expression, whether, project, China energy, response, already, continue to build, people (minzhong), journalists, recently
Cluster3	Mayer's opinion reported by mainstream media	Making troubles, lightly, release, things, appeal, firmly, stress, reasonable, complete, tone, violate, light case, (provincial) people's congress, the fifth (conference)
Cluster4	Event	Haimen event, big, government, coal-fire, build, today, riot, electricity plant, people (renmin), Wukan, one, Weibo, we
Cluster5	Neteast news	Neteast client end news, share, Neteast news, Haimen, happen, mass incident, cn, http
Cluster6	Question raised	Again, a small stamp, masses who do not understand the truth?, mass incident
Cluster7	Other mass incidents	Again, phoenix media, East, compared to the last time, this time, ten times, Xu Xing, outburst

Table 6 Wukan clusters in the four phases

Phase 1: Protest

Clusters	Cluster Theme	Keywords
Cluster1	Weibo	http, retweet, Weibo, Haha
Cluster2	Event	China, world, attention, event, Wukan, Guangdong, Zhengfu

Phase 2: Negotiation

Clusters	Cluster Theme	Keywords
Cluster1	Literacy and democracy	Xiao Shu, one time (yici), change, get rid of (baituo), Xue Manzi, read, the more the better, support, everybody, retweet, donation, book fund, one yuan (yiyuan), Xiong Wei, enlightenment
Cluster2	Election and Weibo	Election, they, attention, need, please, now, Weibo, one, thanks, weibo friends, together, masses (baixing), villagers, retweet, Wukan village, China, new
Cluster3	Democracy and solution	Economy, solve, politicians, Xiaogang village, local, autonomy (zizhi), http, government, democracy
Cluster4	Demand	Land, village head, we, people (renmin), hope, solve, work, secretary, problem, reform, society, working team, provincial, demand, news, interest, politics, Shanwei
Cluster5	Justice	Xue Jinbo, rights protection, lawyer, law, important, event, Hongkong, Wukan, rights, conflict, appraise, today
Cluster6	Elite	Yu Jiarong, develop, people (minzhong)

Phase 3: Election

Clusters	Cluster Theme	Keywords
Cluster1	Literacy	Wu Jiaxiang, democracy, quality theory, precise, history book (shice), Be written into (zairu), myself (benren), judgement
Cluster2	Historical meaning	Basic levels, peasants, country fellows, Han Zhiguo, young people, future, influence, today, development, take place, results
Cluster3	Democracy	Xiong Wei, enlightenment, villagers, quality, reform, China, election, democracy
Cluster4	Election	Village committee, http, representatives, event, committee members, vote
Cluster5	Historical meaning 2	History, people (renmin), we, attention, Xue Jinbo, retweet, country fellows (guoren), society, lawyers, land, problem, government, support, they, hope, Weibo

Phase 4: Post-Protection

Clusters	Cluster Theme	Keywords
Cluster1	Literacy	Enlightenment, Xiong Wei, Liu Jianfeng, library, books, boost, life
Cluster2	Democracy	cn, http, election, democracy, China, villagers, event, sport, retweet, representatives, journalists
Cluster3	Solution	Village committee, land, secretary, village, they, share, news, solution, village officials, Zhang Jianxing
Cluster4	Reform	Problem, reform, Wang Yang, meaning, politics
Cluster5	Weibo	Government, society, attention, people (renmin), Wukan, we, no, please, Weibo
Cluster6	Meaning	Experience, Wukan, hope, lawyer, Hongkong, interests (liyi), south (nanfang)

¹ CNNIC is short for China Internet Network Information Centre. Accessed on July 1

2012, at

<http://www.cnnic.cn/research/zx/qwfb/201206/W020120613472937186052.pdf>

² Ten of thousands of protesters clashed with police who had allegedly covered up a girl's death in the Guizhou Province.

³ We selected Sina's Weibo in the present study for two reasons: 1) it is the first social media site of this kind in China, and 2) it is one of the two most important Weibo sites.

⁴ Both are computer programming languages.

⁵ After several tests, 100 is considered as being appropriate for detailed analysis and interpretation.

⁶ The types of users are judged by information published on their Weibo profiles and messages.

⁷ Elites include cultural and economic elites such as journalists, university academics, lawyers and CEOs. No political elites such as politicians have been recognized.

⁸ Cluster tables are all adapted from dendrograms. Some types of keywords such as conjunctions and particles have been removed from the clusters.

⁹ With a title of "Zhen Renhao, the Shantou Mayor, talked about the Haimen incident: have to arrest some people"

¹⁰ All the real names of profiles have been removed with a concern over possible ethics issues. We decided to use types of Weibo users instead.