Is it all about politics? An analysis of the activities of the Swedish political Twitter elite

David Gunnarsson Lorentzen, University of Borås

The purpose of this paper is to map the topics discussed over three four week periods among a set of predefined prominent Twitter users, as well as their interaction with other users. The paper also makes a methodological contribution as it identifies prominent users from more complete conversations than a hashtag based dataset constitutes, and illustrates how methodological choices impact on the results. 985 Twitter users were selected from an eight week pilot study of conversations around a political hashtag. The messages sent by and to them were collected during twelve weeks over one year. Their activities were analysed focusing on hashtags in tweets posted by and to them, and to what extent they interacted with other users. Overall, political topics dominated the activities of and around the Twitter elite. Differences between usage of hashtags could be seen when considering only tweets posted by the tracked users compared to when considering all tweets in the dataset. The 258 most often used hashtags were closely related to most other prominent hashtags. The elite users replied to and retweeted many other users, however, most of these were addressed infrequently. This longitudinal study sheds light over how topics and hashtags evolve over time, as well as the elite users' interaction with other users.

Keywords: Twitter, longitudinal studies, hashtags, politics, topic mapping, opinion leadership

Gunnarsson Lorentzen, David, "Is it all about politics? An analysis of the activities of the Swedish political Twitter elite". *HUMAN **IT*** 13.3 (2017): 115–155

Twitter is an important alternative source for news as well as a platform for discussions within a wide variety of topics. One important topic is politics. With numerous studies finding evidence of political Twitter activity being dominated by a small set of elite users (e.g. Barberá & Rivero 2015; Bruns & Highfield 2013; Bruns & Stieglitz 2013; Lorentzen 2014; 2016; Tumasjan et al. 2011), it is relevant to study the activities of and around these users. Which topics are most often discussed? What content is most often filtered forward by Twitter users who follow the activities of these prominent users? To what extent do the elite users communicate with non-elite users?

The most prominent users can act as gatekeepers (Shoemaker & Vos 2009) or opinion leaders (Rogers 2003), or perhaps shift between these roles. As popular users are more visible (e.g. van Dijck 2013; van Dijck & Poell 2013), it follows that other participants in the conversations around the topic, and other stakeholders, need to consider, and perhaps adapt to, their behaviour. If one wants to be successful within a topic on this platform, it is reasonable to consider how to interact with the elite users, and how they themselves interact with other users. The present study takes a longitudinal approach and focuses on the activities of and around 985 prominent participants in the Swedish political Twittersphere. A Twittersphere is described by Ausserhofer and Maireder (2013) as a sphere of communication. The sphere is here defined by the hashtag #svpol, which is the dominant hashtag for Swedish political Twitter conversations (e.g. Larsson 2014).

Twitter is an open forum which has some specific characteristics. First, it is based on a uni-directed follower model, meaning that its users choose who to follow and by doing so they construct their own filters. Second, it facilitates different modes of communication: broadcasting by posting tweets, redistribution by retweeting, interaction by @mentions and @replies, and acknowledging by favourites/liking (e.g. Larsson 2015). Redistribution through retweets is a form of forward filtering (Weinberger 2011), which means that the retweeting user brings content to the front for its followers. The follower model could result in polarisation or echo chambers if Twitter users mostly follow like-minded, or users tweeting about a topic of mutual interest, and previous Twitter studies have found indications on such effects, especially regarding the redistribution aspect (e.g. Conover et al. 2011; Dyagilev & Yom-Tov 2014; Lorentzen 2014). Another important trait of Twitter is the notion of popularity. This has the effect of popular users being more emphasised by the platform (e.g. van Dijck 2013; van Dijck & Poell 2013), which introduces a self-reinforcing spiral. The more visible users are more likely to gain more followers, and have their content replied to and redistributed, which then results in more visibility.

In the context of news, mass media actors are traditional gatekeepers, filtering information to end users (Shoemaker & Vos 2009). On social platforms such as Twitter, information sources are not limited to mass media actors (e.g. Bastos, Raimundo & Travitzki 2013) and any ordinary user can take a gatekeeping role (Klinger & Svensson 2015). Following this, it is not only traditional elites that are part of the Twitter elite, but also any person or organisation that is well connected with other users, and happens to have a successful tweeting strategy. Opinions, among other own content, can be expressed as well, and given their connectedness through followership, the elite users are potential opinion leaders who, according to Rogers (2003), can be identified through network analysis. Compared to their followers, opinion leaders have higher socio-economic status, a higher degree of participation, greater contact with both change agents and mass media and they comply with the norms of their system to a higher degree (Rogers 2003). Dubois and Gaffney (2014) found that other Twitter users than political actors and mass media actors can be opinion leaders, and that these have the ability to influence other people in their personal networks, but offline authority matters too, as indicated by Xu et al. (2014).

With the artificial demarcation of a Twittersphere, it is not trivial to identify opinion leadership. Instead, I refer to the term elite users as a label for the users that are prominent within such a demarcation. This differs from the definition by Wu et al. (2011), which covers traditional elite such as celebrities, organisations and media, and early new media adopters such as prominent bloggers, although these user categories are certainly included among the elite users identified for this study. The elite users could have the potential to act as opinion leaders within and outside of the sphere, however, if they do influence other users is outside the scope of this paper, as other methods would be required for such conclusions to be drawn.

Although there are many examples of research on political Twitter usage and discussions, little is known about which topics are discussed besides politics or how these topics are connected. One related study collected tweets sent by and to a sample of 374 elite users of the Austrian political Twittersphere during four one week periods over four months (Ausserhofer & Maireder 2013). It was found that the tracked users tweeted about politics in every fifth post, and that large shares of their mentions were sent to users outside the sample. The present study takes a similar approach in the sense that it makes use of a sample of users chosen from Twitter activity within a political topic and that it analyses which topics are discussed, as well as interactions across user groups. However, this study also makes extensions in three ways. Firstly, the sample is based on more complete conversations, as replies to hashtagged tweets have been collected (see Method). Secondly, the sample here is larger and contains roughly the amount of users found to be dominant in the chosen setting (e.g. Lorentzen 2014), and the study spans over one year compared to the four months in Ausserhofer and Maireder (2013). Finally, this analysis also considers evolution of topics on a longer timeframe, as well as how the topics are connected.

Studies of the Swedish political Twittersphere have identified a core of elite Twitter users (e.g. Larsson & Moe 2012; Larsson 2014; Lorentzen 2014), but questions still remain regarding the broader context they are part of. Do they for example constitute a small group of people in an interest based forum largely disconnected to the surrounding network of Twitter users or are they a part of a larger structure of different discussions of various topics? The purpose of this study is twofold. Firstly, it aims to outline the most prominent topics the elite users are tweeting about given usage of hashtags, and which topics are most often amplified through retweets by their followers. This gives an indication of what the wider audience finds important among the issues the elite users are tweeting about. Through trend analysis of hashtags, the paper also outlines how the topics evolve over time which in turn indicates how this Twitter community reacts to various events. Are there for example topics that are more stable, and is it possible to identify sudden events that can be studied further? A co-occurrence analysis is then made for outlining how the prominent topics are connected, and finally, the extent to which the elite interact with other users is studied through networks of replies and retweets.

Secondly, the paper makes a methodological contribution as it provides an example of how one can study the effects of the activity among the elite users, by first identifying the users more accurately from a more complete conversation than what a hashtag based dataset constitutes, and then collecting the tweets sent by and to these users over a longer period. With the analysis made on one set including and another set excluding retweets, the paper illustrates how methodological choices impact on the results.

Literature review

Murthy argued that "one's 'banal' activity on Twitter [...] is as much a part of many people's identity as discussing current events" (2013, p. 149). Twitter has continuously moved from the banal even though banality still exists among its users (Rogers 2013a). One might wonder if banality is still visible even among the top topics discussed by the top users. Of Twitter's features, this paper focuses on hashtags, retweets and mentions. The hashtag is a mechanism for a Twitter user to add metadata to a message (tweet) by including the # symbol in front of any string of characters. By doing so, the user is able to signal a wish to take part in a wider conversation (e.g. Bruns & Moe 2013). Tweets can also be conversational in mainly two different ways. One way is to redistribute a tweet (retweet) and another way is to include a username in a tweet (mention or the sub-type reply if the tweet starts with the username). This paper makes use of a broad definition of conversation and considers both redistribution through retweets and interactions through mentions as conversational, similar to Bruns and Highfield (2013). The literature review starts with results from a number of studies regarding conversations on Twitter, focusing on three types of Twitter research. These are sample based studies, where a sample of all tweets have been taken without matching search criteria, hashtag based studies, where one or more hashtags have been tracked to collect tweets including the hashtag(s), and user account based studies, where a fixed or dynamic set of Twitter users have been tracked.

Twitter usage can be seen as event-driven (e.g. Murthy 2012). There are examples of studies where activity spikes have been found following or during a major event. Some of these are election related (e.g. Bruns & Burgess 2011; Larsson & Moe 2012) and others are protest related (e.g. Jungherr & Jürgens 2014). It seems as during important events the Twitter activity is more centred on distributing news than interacting through mentions or replies. This behaviour was found by Jungherr and Jürgens' study of Twitter activity related to protests (2014), where the share of retweets increased significantly during the protests while the share of mentions decreased. Similar findings were also presented by Lorentzen (2016).

In general, research on Twitter so far has found very different measures of interactions through mentions. Studies of Swedish political Twitter conversations indicating a lower level of interactions have been made by Larsson and Moe (2012), who found 7% mentions in a study of Twitter activity before and during the national election, and Larsson

(2014) and Lorentzen (2014), whose studies of non-election times had corresponding numbers of 9% and 18%, respectively. These figures seem to indicate that such interactions do not happen to a large extent on Twitter. However, these studies were hashtag based, and as such, they miss out on non-tagged replies ("follow-on tweets", see Bruns 2012) and thereby underestimating the level of conversation through mentions. While other hashtag based studies have reported lower interaction levels (e.g. Jungherr & Jürgens 2014; Small 2011), one exception is Bruns and Highfield (2013, p. 686), whose study found that roughly one third of the #qldvotes tweets were replies. A more true interaction level might be found in user account or sample based studies, as these do not have the requirement of a certain hashtag or keyword being included in the tweets. Around 30% of tweets have been identified as replies or mentions in the former type (e.g. Graham et al. 2013; Holmberg & Thelwall 2014; Kruikemeier 2014).

Twitter usage of hashtags has been described as belonging to a conversation (e.g. Bruns 2012; Huang, Thornton & Efthimiadis 2010; Larsson & Moe 2014; Lorentzen 2014). Lindgren and Lundström (2011) talked about the hashtag as a discourse, with the "linguistic space" around the studied hashtag #wikileaks found to be global and loosely knit. González-Bailón et al. (2014) argued that hashtags are labels that can be utilised both for the user's own classifications and while they also can be used to participate in a community or topic. The latter part was echoed by Bruns and Moe (2013) who added that Twitter users might simply tag their tweets to increase their visibility without following the conversations. Hashtags can be either topical or non-topical where the latter type is unlikely to be followed by many users (Bruns & Moe 2013).

Far from all tweets contain hashtags and research so far has found very different shares of tweets with hashtags. In two sample based studies, boyd, Golder and Lotan (2010) found hashtags in 5% of the collected tweets, while the corresponding number for Gerlitz and Rieder (2013) was 13%. In a study of Suh et al. (2010), 10% of all tweets and close to

21% of the retweets included one or more hashtags. In a user tracking study of 177 political candidates, Kruikemeier (2014) found hashtags in 25% of the tweets, and in Merry's (2013) study of environmental organisations' use of Twitter, 65% of all tweets included hashtags.

A common approach to study Twitter has been to use one or several hashtags or keywords to track (e.g. Bruns & Moe 2013). There are a few examples of studies made in a similar setting as in the present paper. Larsson and Moe (2012) focused their study on Twitter usage before and during the national election 2010, while Larsson (2014) and Lorentzen (2014) had their focus on non-election times. However, none of these studies attempted to analyse other topics or sub-topics within the conversations. A study that did so was made by Pearce et al. (2014), who investigated which hashtags were used in tweets about Intergovernmental Panel on Climate Change's Fifth Assessment Report.

Examples of longitudinal studies of political Twitter usage or conversations about political topics are few. Larsson and Moe (2014) compared the reply and retweet networks around hashtags during two elections in Norway and found elite domination, but also activity of other users including smaller political actors. Apart from some changes in retweet usage, there was little change between both periods. Kruikemeier (2014) tracked 177 political candidates during three months before and two months after the Dutch national elections 2010 and found that Twitter is mainly a means for the politicians to talk about their private persona. This usage was more evident during than after the campaign, as was interaction by mentions and retweets. A third example is Larsson's (2014) three month study of the hashtag #svpol. Its usage was fairly stable and was dominated by broadcasting and redistributing. Attention in the form of mentions was often given to politicians.

The retweeting practice in relation to political conversations seems to be polarised; Twitter users prefer to retweet tweets posted by like-minded (e.g. Conover et al. 2011; Lorentzen 2014). It has also been shown that tweets expressing more emotionality are more likely to be retweeted

(Dang-Xuan et al. 2013). Elite users such as mass media and political actors have been found to have a higher retweet per tweet ratio than citizens (Hawthorne, Houston & McKinney 2013), but network aspects seem to be important too. Users in central positions, followership-wise, are more likely to have tweets retweeted (González-Bailón et al. 2011), a finding supported by Xu et al. (2014), who also concluded that users geographically located close to an event were more likely to be successful in redistribution.

Method

Data Collection

Similar to some of the examples mentioned above, this study followed a fixed set of users identified through their previous activity. your Twapper Keeper (see Bruns & Liang 2012) was used to collect data. It was modified in two ways. First, the field for which tweet a reply is directed to was added, and second, the streaming functionality was used for tracking user accounts instead of keywords. Data were initially collected by tracking #svpol for eight weeks during spring 2013. It is argued here that this hashtag, which covers politics broadly, is stable enough for a relatively consistent group of prominent users to emerge over time. This claim is supported by Larsson's (2014) three months long study of usage of the hashtag, the only extensive investigation of #svpol so far. The aim of the data collection was to also capture follow-on conversation, that is, non-tagged replies to tagged tweets, and so getting a more true picture of the most active and visible participants in the conversations. This was made possible by simultaneously using the search API to collect tweets matching the hashtag, and the streaming API to collect tweets sent to the most frequent users of the hashtag in the dataset, storing tweets matching the hashtag and/or being replies to stored tweets. An improved version of this method has since this data collection been presented and tested by Lorentzen and Nolin (2017) and D'heer et al. (2017).

From this initial set, the most prominent users were identified based on their activity, visibility and spreadability, similar to Lorentzen's weighting scheme (2014). Underlying this is the assumption that it is possible to identify opinion leaders through network analysis, or more specifically, communication network analysis (Rogers 2003). Here, the broadcasting tweets were added to the conversational tweets and communication relationships. According to Zimmer and Proferes (2014), fewer than 1,000 users have been common as subject of study in the Twitter literature. 985 non private accounts were found to have an aggregated score above an arbitrary threshold value of 30. These users accounted for 71% of all non-spam tweets during the eight weeks, which was quite similar to the findings in Lorentzen's study (2014), in which 72% of all tweets were posted by 916 users. The main data collection was performed utilising the streaming API, through filtering the stream by user IDs. According to the Twitter documentation, using the streaming API for tracking users will capture tweets that are 1) posted by a given user A including retweets, 2) replies to tweets posted by A, 3) retweets of tweets posted by A, and 4) manual replies to A, which are tweets starting with @A but not created by using the reply button (Twitter, Inc. 2015). All the captured tweets are related to the tracked users.

A distinction between manual retweets and button retweets has been made (e.g. Bruns & Moe 2013), where the former is an edited version of the original tweet including "RT", "MT", or "via" before the Twitter handle of the original tweet author, and the latter is simply a copy of the tweet. It has been argued that the manual retweet is more conversational (Highfield, Harrington & Bruns 2013). Indeed, it requires more effort than simply clicking on a button, and by retweeting manually it is possible to edit the tweet somewhat. In the software used for collecting data there was no way of finding out whether a button retweet was a retweet and if so, of which tweet. your Twapper Keeper can be modified further to collect that information, but that option was not investigated here. The button retweet functionality means that it is difficult to identify retweets not starting with the pattern "RT @username", which in turn has an implication on the analysis. Retweets not matching this pattern are not identified as retweets in this study, an approach that has been utilised before (e.g. Bruns & Burgess 2012; Highfield, Harrington & Bruns 2013; Lorentzen 2014), and in some cases also taking other variants into account, such as "via @username" (e.g. Pearce et al. 2014).

Three four week long periods during one year were chosen as samples (2013-06-02 – 2013-06-30, 2013-11-24 – 2013-12-22, and 2014-05-11 – 2014-06-08). The third period coincided with the European election which was held at May 25. Four months later, the national election was held, but this was not covered here. The aim was to identify two periods without a major scheduled political event as well as one with such an event. The purpose of this was to make it possible to identify what other topics apart from politics, if any, that were discussed among these users. After the third period the number of users had decreased to 942 due to either the account being suspended or removed by the user.

During the 12th day of the first period problems with the streaming API were experienced. This explains the drop in amount of tweets for that day. The problems turned out to be on the client side and were repaired after a few hours through software update. This is a problem related to API based real-time data collection. The data collection tool needs to be constantly monitored to prevent missing out on data.

Data Description

In total, 2,303,403 tweets from 154,993 Twitter users were collected, 1,288,746 from the tracked users. The top ten language codes were en/en-gb (English, 1,246,081 tweets), sv (Swedish, 1,002,245), es (Spanish, 14,883), de (German, 9,286), tr (Turkish, 5,656), fr (French, 5,074), no (Norwegian, 3,220), it (Italian, 3,150), da (Danish, 2,745), and pt (Portuguese, 1,944). In Table 1 we see that not all tracked users posted tweets during the selected periods. In the first period, 953 of 985 active accounts provided tweets, and in the last period, the corresponding figures are 854 of 942. A constant decrease in the number of tracked users posting and the number of tweets posted by them can be seen, however, the tweets per user increased from 452 to 477 and 500. Around 25% of all tweets included hashtags with 10% including more than one. Hashtag usage increased slightly over time.

		P1	P2	P3
All users	Tweets	711,688	761,301	830,414
	Active users	62,788	65,777	71,772
	% Tweets with hashtags	25.0	25.0	26.8
	% Tweets with > 1 hashtag	9.8	9.1	12.1
Tracked users	Tweets	431,196	429,836	427,714
	Active users	953	901	854
	% Tweets with hashtags	24.6	25.5	27.0
	% Tweets with > 1 hashtag	9.1	9.3	12.0

Table 1. Description of the data.

Data Analysis

All hashtags were extracted from the tweets collected. Given the nature of tweets, including a hashtag is a conscious choice in order to either describe the tweet or tag it into a given topic. While focusing on hashtags restricts the dataset to a biased sample (slightly more than 25% included one or more hashtags, see Table 1), it is here postulated that the follow-on conversation belong to the topic the hashtag is reflecting (i.e. any reply to a hashtagged tweet is related to that hashtag). In the sets studied in this paper, 81,226 (93%) of all tweets with multiple hashtags (manual retweets excluded) are posted by the tracked users. The remaining 6,114 tweets were posted by snowballed users; i.e. users introduced into the conversations through replies or button retweets.

To identify stable, emerging, and disappearing hashtags over time, time-series analysis of the hashtags was conducted. The 100 most used hashtags overall for a set without retweets and a set including retweets were used as basis for the analyses. For each day, the number of tweets the hashtag occurred in, and the number of unique users of the hashtag were counted. The next step was to perform co-occurrence analysis of hashtags to map the topics discussed. According to Callon, Law and Rip (1986, p. 106), a text can be reduced to a "network of powerful words" by an indexer of the text. In the Twitter context, it is assumed that the hashtags can be considered as powerful words, as they are chosen to describe the topic or connecting it to an aggregate stream of related tweets. Hence, the tweeter takes the role as the indexer. Co-occurrence analysis as a method can provide an overview of the communication space around the tracked users and identify the topics within this space as well as their inter-relationships (Borra & Rieder 2014; Callon, Law & Rip 1986; He 1999). In this setting, it would emphasise the most prominent topics rather than the most common ones. Words can for example act as global hubs if they are connected to many words with low degree, or local hubs if they are connected to a global hub and are related to different contexts (Drieger 2013).

As in Pearce et al. (2014), all retweets were excluded in the co-occurrence analysis as they distort the results. Another reason for excluding the retweets is as when "RT" and a Twitter handle is added to the tweet and possibly other edits taking place, not just the content but also the hashtags of the original tweet might be affected, especially for those tweets that ends with hashtags. There were quite a few examples of hashtags being cut off at the end of a retweet, resulting in "new" hashtags, and some of them being omitted. The emphasis of this analysis is on the relationship between the hashtags and not the visibility of the hashtags. The latter part is analysed as the time series described above.

The co-occurrences for the three periods were aggregated into one hashtag network. The hashtags occurring in at least 100 tweets were kept and disconnected smaller clusters of hashtags were excluded as an initial filter (giant component filtering). By utilising the Gephi (Bastian, Heymann & Jacomy 2009) algorithm Modularity for community detection (Blondel et al. 2008; Lambiotte, Delvenne & Barahona 2009), larger clusters of hashtags were identified. While this algorithm does not result in absolute clusters (the borders are fuzzy) it does indicate which hashtags are more tightly connected. By setting the resolution to 2.0 two main clusters were identified, one larger including 74% of the nodes and one smaller including 15% of the nodes. A couple of small clusters surrounding these two were also identified by the algorithm. Finally, a force-directed algorithm was used for layout of the network. Such an algorithm positions hashtags with stronger links (more often co-occurring) closer to each other.

The last part of the analysis focuses on replies sent and retweets made by the elite users. A sample of 100,000 tweets from each period was taken and from this, all replies and retweets posted by an elite user were extracted. Through network analysis, the extent to which these elite users address other users and redistribute their messages was then analysed. Recalling Rogers' (2003) ideas of connectedness, the potential opinion leaders should have central positions in such networks.

Results

Descriptive statistics

The share of retweets was higher during the last two periods, most notably during the Election Day in the middle of the third period. As mentioned above, retweeting also seems to increase during sudden events. A clear example of this is when Nelson Mandela passed away, a heavy usage of #mandela in retweets was found but in contrast relatively few non-retweets included this hashtag. Another offline event possibly influencing the amount of retweets was the release of the PISA (Programme for International Student Assessment) report. Other days with dramatic increase of retweets were day 21 of period 1 (Snowden), day 28 of period 2 (demonstrations against racism in the suburb Kärrtorp), and days 15 and 16 of period 3 (the election). On days 19 and 20 of period 2, the case was the opposite with a dramatic decrease in retweets and increase in mentions. This was the first day the #kärrtorp hashtag was used within this dataset. These data indicate on a relationship between retweets and mentions. Typically, the share of mentions decreased as the retweets increased, while the share of original tweets was stable over all three periods.

		P1		P2		P3	
		N	%	N	%	N	%
All users	OT	111,194	15.6	95,844	12.6	91,705	11.0
	@	356,729	50.1	356,205	46.8	353,460	42.6
	RT	243,765	34.3	309,252	40.6	385,249	46.4
Tracked users	OT	111,194	25.8	95,844	22.3	91,705	21.4
	@	221,265	51.3	210,860	49.1	196,586	46.0
	RT	98,737	22.9	123,132	28.6	139,432	32.6
Non-tracked users	OT	0	0.0	0	0.0	0	0.0
	@	135,464	48.3	145,345	43.8	156,874	39.0
	RT	145,028	51.7	186,120	56.2	245,817	61.0

Table 2. Distribution of tweet types (OT: original tweets, @: @mentions, RT: retweets). Note: no original tweets from users outside the filter can be captured with this method.

Focusing on the tracked users, we see that slightly less than half of the tweets are mentions. It is a slightly higher share than overall. The figures in table 2 confirms that the large set of peripheral users retweet tweets of the tracked users to a larger extent than replying to them. The share of interactions through mentions can be compared to studies with similar approaches previously mentioned. Graham et al. (2013), Holmberg and Thelwall (2014) and Kruikemeier (2014) all found shares of mentions around 30%. The results in this paper show a far higher share of mentions at 40-50% of all tweets, depending on which period focused on. Both Bruns and Highfield (2013) and Lorentzen (2014) showed that the most active users are also the most conversational ones, so the deliberative tracking of the most prominent users of a given topic in this paper is perhaps quite likely to yield these kinds of results.

There are two distinct dips in the overall activity (Fig. 1 and 2), both occurring during the first period. The first of these was due to software failure as mentioned above while the other is probably due to lower activity (software or hardware failures at Twitter cannot be ruled out though). The activity spike in period 3 coincides with the election. What is interesting here is that the spike in activity of the tracked users (Fig. 2) is not as significant as the one where all users are considered (Fig. 1). The elite users do not have the same event-driven behaviour as the users around the elite. Similar findings were made by Lorentzen (2016) who concluded that the generally most active users had a stable activity level whereas the least active users were as most active following a major event.

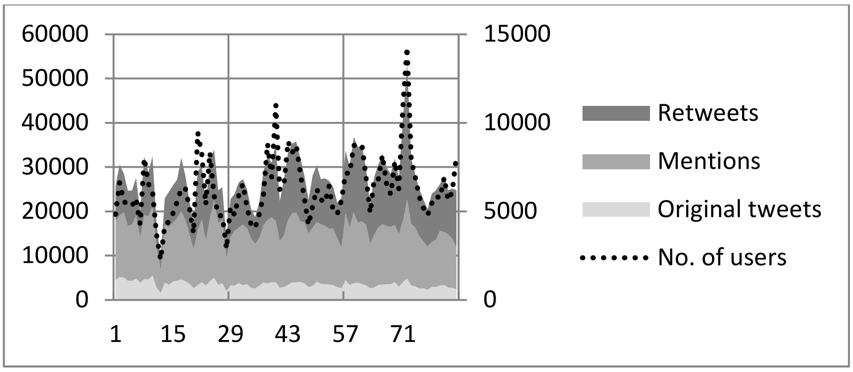

Figure 1. Overall activity partitioned by tweet types. Left: period 1, centre: period 2, right: period 3. The "Number of users" data series is attached to a secondary axis for the sake of clarity.

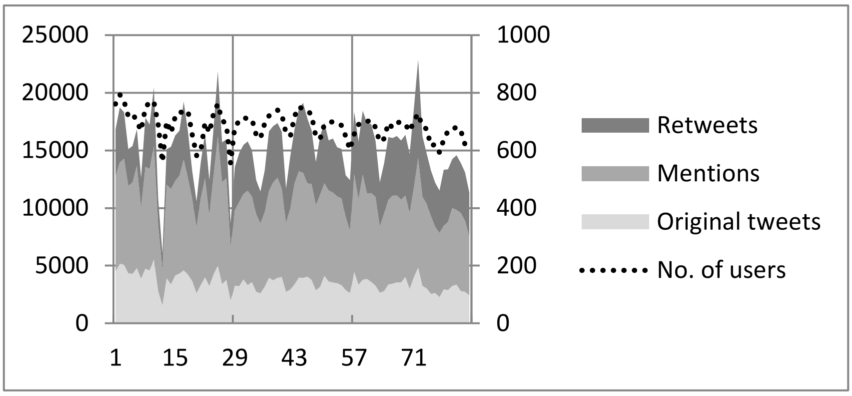

Figure 2. Overall activity partitioned by tweet types, tracked users. Left: period 1, centre: period 2, right: period 3. The "Number of users" data series is attached to a secondary axis for the sake of clarity.

The overall share of retweets is at its highest during these couple of days (57%) and then quickly drops to its more normal level (40-45%). For the tracked users, a similar pattern can be seen. The share of retweets reached 37% and then dropped to around 30%. The days before, during, and after Election Day a relative decrease in mentions can be seen, most notably when considering all users, but also among the tracked users. While these findings show some interesting patterns, it must be noted that the peripheral users here only contribute with replies to and retweets of tweets posted by the tracked users. Hence the overall activity is biased towards the activity of this set of 942 (last period) to 985 users (first period).

Hashtag usage

Tables 3 (retweets excluded) and 4 (retweets included) list the 30 most used hashtags. In total, 52,087 unique hashtags were identified, and 37,536 hashtags were identified in the set with retweets excluded. This difference indicates on the one hand that users do add their own hashtags

Hashtag	Description	# tweets
#svpol	Swedish politics	97,081
#aftonbladet	Newspaper	7,514
#sverige	Sweden	5,816
#migpol	Migration politics	4,177
#svt	Swedish Television	4,148
#val2014	Election 2014	3,285
#eupol	European (Union) politics	3,045
#pldebatt	Party leaders' debate (TV show)	2,915
#dinröst	Your vote	2,904
#assange	Julian Assange	2,854
#snowden	Edward Snowden	2,553
#euval2014	European Union Election 2014	2,393
#val14	Election 2014	2,335
#gbgftw	Gothenburg for the win	2,312
#nowplaying	Now playing	2,243
#göteborg	Gothenburg	2,191
#wikileaks	Wikileaks	2,103
#sd	The Sweden Democrats Party	1,925
#euval14	European Union Election 2014	1,697
#almedalen	Almedalen (annual political event)	1,689
#kärrtorp	Kärrtorp (Swedish suburb)	1,608
#euval	European Union Election	1,572
#svfm	Swedish Defence Forces	1,561
#sweden	Sweden	1,477
#nyheter	News	1,427
#nsa	National Security Agency	1,405
#agenda	Agenda (TV show)	1,282
#eu	European Union	1,264
#reinfeldt	Fredrik Reinfeldt	1,230
#alliansen	The Alliance	1,223

Table 3. Top 30 hashtags with description according to number of tweets (RTs excluded).

Hashtag	Description	# tweets
#svpol	Swedish politics	198,465
#snowden	Edward Snowden	27,982
#assange	Julian Assange	25,116
#wikileaks	Wikileaks	18,998
#svt	Swedish Television	11,177
#aftonbladet	Newspaper	10,626
#val2014	Election 2014	10,301
#sweden	Sweden	9,282
#dinröst	Your vote	8,413
#nsa	National Security Agency	8,359
#migpol	Migration politics	7,991
#sverige	Sweden	7,703
#sd	The Sweden Democrats Party	7,185
#eupol	European (Union) politics	7,103
#euval2014	European Union Election 2014	6,992
#manning	Chelsea (Bradley) Manning	6,571
#kärrtorp	Kärrtorp (Swedish suburb)	5,582
#mandela	Nelson Mandela	5,306
#euval14	European Union Election 2014	5,101
#euval	European Union Election 2014	5,047
#ukraine	Ukraine	5,023
#tpp	Trans Pacific Partnership	4,755
#freebrad	Free Chelsea (Bradley) Manning	4,407
#val14	Election 2014	4,338
#prism	PRISM (surveillance program)	4,253
#pldebatt	Party leaders' debate (TV show)	4,252
#svfm	Swedish Defence Forces	4,059
#gbgftw	Gothenburg for the win	3,836
#agenda	Agenda (TV show)	3,603
#piratpartiet	The Pirate Party	3,504

Table 4. Top 30 hashtags with descriptions (RTs included).

to retweets, but on the other that hashtags at the end of a tweet might be cut off as a consequence of adding "RT" and a Twitter handle to the tweet. 23 of the hashtags are included in both sets, and #svpol was by far the most popular. In general, politically related hashtags dominate both these lists. Some topics are more prominent when retweets are included, most notably Chelsea (formerly Bradley) Manning and Nelson Mandela. The surveillance programme PRISM and the Pirate Party are also included in this set, but not in the one without retweets, as were the Trans Pacific Partnership and Ukraine. We see from both these top lists that hashtags are used to talk about persons and organisations. Apart from the above mentioned we find Julian Assange, Edward Snowden and Fredrik Reinfeldt (prime minister at the time of the data collection), the Sweden Democrats Party and the (at the time of the data collection) leading coalition of parties named the Alliance. There are examples of recurring events reflected on the use of hashtags, for example the annual political event Almedalen, and the TV show Agenda. Some hashtags are not related to politics, for example #nowplaying. This hashtag could be seen as an example of banal topics appearing with the profound ones.

Hashtag Trends

Figures 3 (retweets excluded) and 4 (retweets included) shows the 20 most used hashtags over all three periods. Overall, #svpol dominated the conversations during all three periods and was the most used every day including the Election Day, even if we count #euval, #euval14, and #euval2014 as one hashtag. With the two elections occurring during the same year, the hashtag #dinröst ("your vote") was introduced by the Swedish Television. This hashtag was frequently used during the third period as were four other election related hashtags. Both EU (#euval*) and domestic (#val*) election hashtags were used, indicating that both the elections had some impact on the overall usage of hashtags. All of these had different usage patterns. #val14 (not prominent in the full set) and #val2014 were the stable ones over the whole period, with the latter more used. These were possibly used to discuss not only the European election, but also the national election four months later. All of the hashtags specifically used for the European election, #euval, #euval14 and #euval2014 quickly faded after Election Day. Of these, the 2014 variant was more popular, both regarding number of tweets and number of daily users, mirroring the use of the #val* variants although with smaller differences.

There are both similarities and differences between these two sets. The most striking difference between the non-retweet set and the full set is that there are more hashtags fluctuating in the latter compared to the formed, both in number of usages and number of users. This too can be explained by the event-driven behaviour of the less active users which is not as evident among the tracked elite. But some hashtags seem to be both stable and prominent regardless if retweets are included or not. Two of them are representing the mass media outlets the Swedish Television (#svt) and the newspaper Aftonbladet. Both hashtags were fairly stable over all three periods, with the latter more used during the first of them. Other stable hashtags represent migration politics (#migpol), Sweden (#sverige), the Sweden Democrats (#sd) and European politics (#eupol), with the latter increasing before the European election. Also interesting to note is that #val2014 was used during spring 2013 and increasingly so during November/December the same year.

One topic that was always ubiquitous was Wikileaks, here visible through the use of #assange and #wikileaks. Wikileaks and whistleblower Edward Snowden (#snowden) were more prominent in the full set, indicating that these topics are more often retweeted. Especially Snowden was mentioned by many tweeters, and the hashtag was amplified by a heavy usage of retweets, eclipsing #svpol in terms of number of tweets and users during a couple of days in period 1, following his release of reports to mass media outlets. During period 2, Wikileaks, Assange and Snowden were prominent in the full set but not in the non-retweet set.

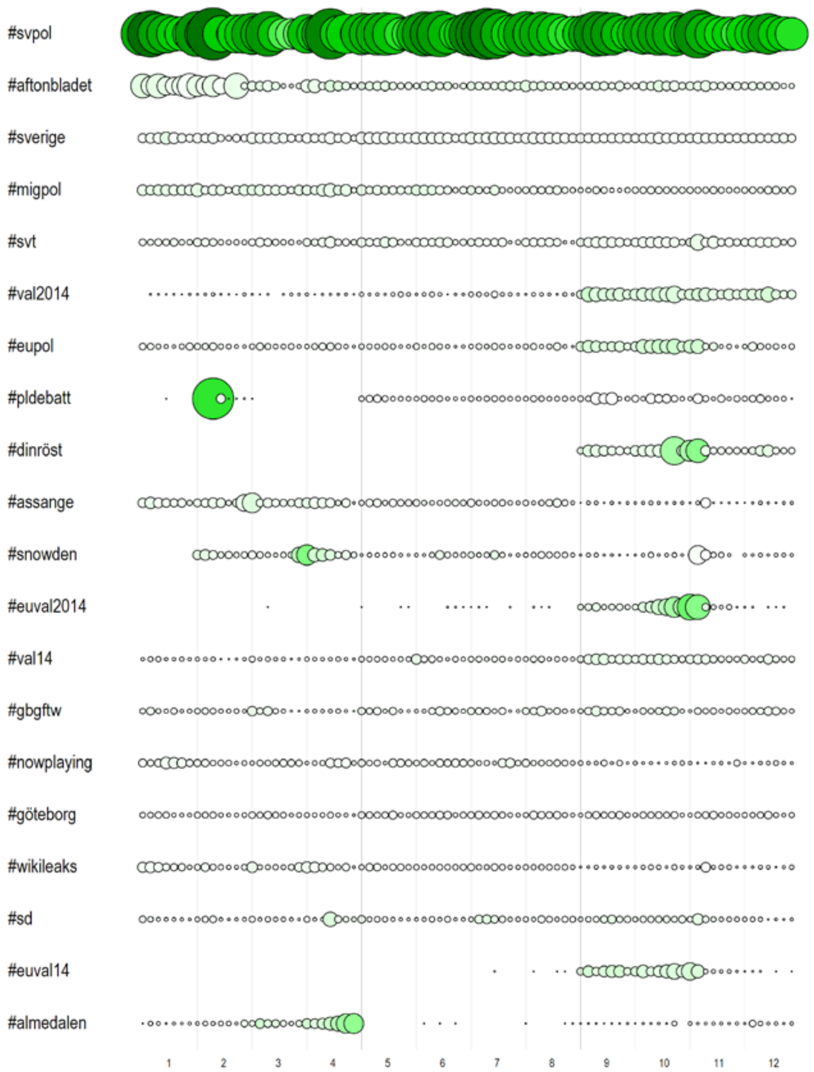

Figure 3. Top 20 hashtags for all three periods (RTs excluded). Size: number of tweets, colour: number of users (dark green: many users, white: few users).

Other related topics that were amplified through retweets were NSA and Manning, the latter also visible through the hashtag #freebrad, which fell outside of the top 20 tags overall. This amplifying could be the effect of people reacting to and spreading certain news, which is also seen in usage patterns of #kärrtorp and #mandela, both prominent in the full set but not in the other set. #kärrtorp was related to protests against racism, and was used during the last eight days of the second period, occurring 1,600 in the elite set and 5,561 times in the full set (1,608/5,582 when considering all periods). This hashtag is a good example of how a single topic can dominate the local Twitter usage for a couple of days and then disappear. It should be noted that #kärrtorp was among the most prominent hashtags in the non-retweet set during period 2, and as can be seen in table 7, is placed at 21st in the overall hashtag count. The full set again differed from the non-retweet set with the inclusion of Nelson Mandela. Following his death many users retweeted tweets including either the hashtag (visible here) or the name without the #, however, few of the tweets referring to Mandela were non-retweets.

Hashtags that were prominent in the non-retweet set, but not the other, were #pldebatt, #almedalen, hashtags related to Gothenburg and #nowplaying. A onetime high was noted for the party leaders' debate in period 1, and then it became more stable as the election came closer. Period 3 was dominated by election related discussions, and during this period there were small differences between both sets with regards to prominent hashtags, but, again, Wikileaks and Snowden were more popular topics in the full set.

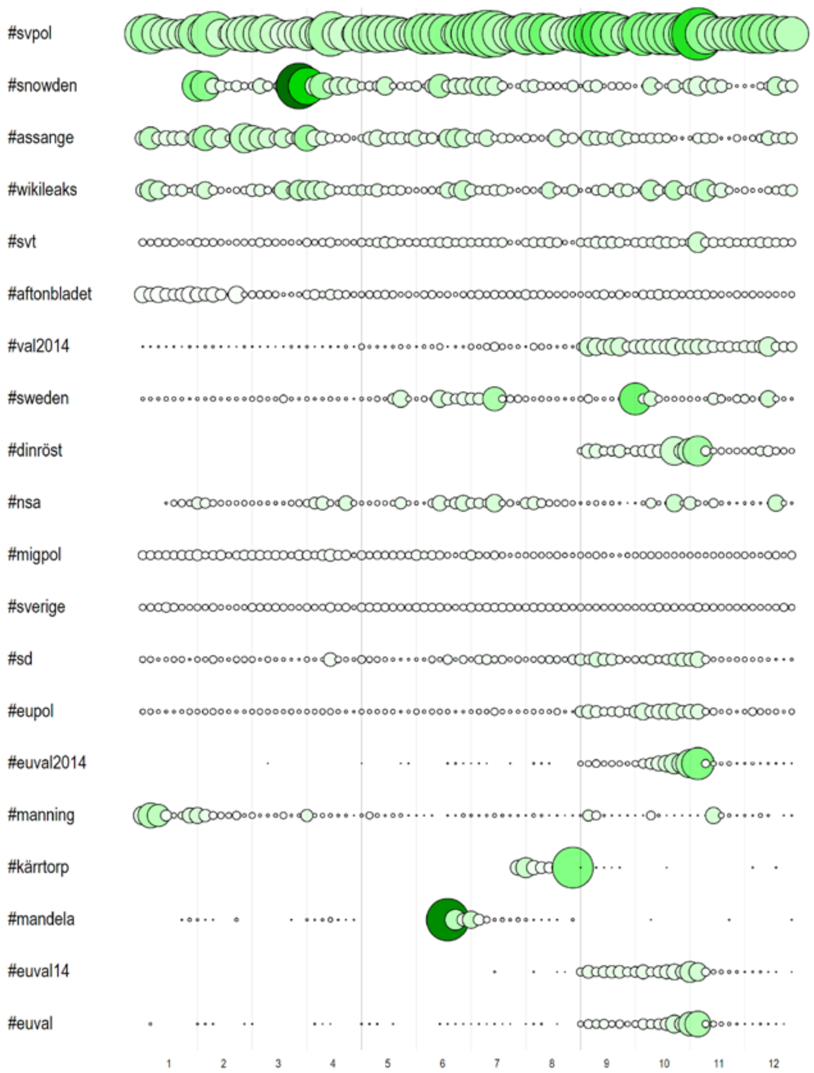

Figure 4. Top 20 hashtags for all three periods (RTs included). Size: number of tweets, colour: number of users (dark green: many users, white: few users).

Hashtag	Description	# co-occurrences
#svpol	Swedish politics	45,413
#sverige	Sweden	4,257
#migpol	Migration politics	3,796
#svt	Swedish Television	3,233
#val2014	Election 2014	2,756
#eupol	European (Union) politics	2,612
#dinröst	Your vote	2,477
#assange	Julian Assange	2,353
#pldebatt	Party leaders' debate (TV show)	2,327
#aftonbladet	Newspaper	2,227
#val14	Election 2014	2,178
#wikileaks	Wikileaks	1,933
#snowden	Edward Snowden	1,901
#euval2014	European Union Election 2014	1,874
#sd	The Sweden Democrats Party	1,798
#gbgftw	Gothenburg for the win	1,735
#nyheter	News	1,423
#euval14	European Union Election 2014	1,416
#sweden	Sweden	1,374
#euval	European Union Election 2014	1,354
#nsa	National Security Agency	1,306
#eu	European Union	1,233
#reinfeldt	Fredrik Reinfeldt	1,221
#alliansen	The Alliance	1,153
#svfm	Swedish Defence Forces	1,065
#säkpol	Security politics	1,059
#facket	The Union	1,029
#dn	Newspaper	1,027
#media	Mass media	994
#invandring	Migration	986

Table 5. Top 30 most co-occurring hashtags (RTs excluded).

Co-occurring hashtags		Description	Count
#assange	#wikileaks	Julian Assange and Wikileaks	702
#dinröst	#val2014	Your vote and Election 2014	586
#pldebatt	#sverige	Party leaders' debate and Sweden	551
#nyasverige	#reinfeldt	The new Sweden and Fredrik Reinfeldt	386
#assange	#snowden	Assange and Edward Snowden	349
#assange	#sun4assange	Assange and Sun for Assange	349
#snowden	#wikileaks	Snowden and Wikileaks	349
#dinröst	#euval2014	Your vote and European election 2014	309
#sif14	#snowden	Stockholm Internet Forum 2014 and Snowden	306
#invandring	#pldebatt	Migration and Party leaders' debate	288
#dn	#sverige	Dagens Nyheter (newspaper) and Sweden	261
#nsa	#snowden	National Security Agency and Snowden	257
#sverige	#sweden	Sweden and Sweden	244
#dinröst	#piratpartiet	Your vote and the Pirate Party	237
#föpol	#säkpol	Defence politics and Security politics	221
#nyheter	#sweden	News and Sweden	213
#manning	#wikileaks	Manning and Wikileaks	206
#invandring	#migpol	Migration and Migration politics	205
#sverige	#val14	Sweden and Election 2014	203
#svfm	#säkpol	Swedish Defence Forces and Security politics	201
#sverige	#svpbs	Sweden and Swedish political blogosphere	197
#dinröst	#slutdebatt	Your vote and Final debate	197
#euval	#euval2014	European election and European election 2014	196
#assange	#manning	Assange and Manning	194
#invandring	#sverige	Migration and Sweden	193
#journalister	#reinfeldt	Journalists and Reinfeldt	190
#dinröst	#euval	Your vote and European election	187
#invandring	#islam	Migration and Islam	185
#föpol	#svfm	Defence politics and Swedish Defence Forces	184
#dinröst	#svt	Your vote and Swedish Television	184

Table 6. The most connected hashtags with description. Count denotes number of co-occurrences (#svpol and RTs excluded).

Co-occurring hashtags

As #svpol was used to identify prominent users it is not surprising that this hashtag turned out to be far most used and connected. It co-occurred with other hashtags in 45,413 of all tweets with more than one hashtag, which is more than half of them (Table 5). This indicates that Swedish politics is a dominant topic for this group of Twitter users. Table 6 shows the most connected pairs of hashtags with #svpol excluded. It gives an indication of which topics are the most heavily discussed in this setting. Political issues dominate the agenda. All these pairs of hashtags are closely related to politics. The election, Wikileaks and whistleblowers, migration politics and security politics dominate this list. The Pirate Party is the only party in this list, appearing in a strong connection with "your vote". Interestingly enough, the Pirate Party had a poor European election.

The network of hashtags occurring in at least 100 tweets resulted in a network heavily dominated by domestic politics (Fig. 5). There is a smaller cluster, which is dominated by international matters (Fig. 6). There are few examples of topics that can be considered being banal. In both clusters, Eurovision Song Contest is represented through a number of hashtags, and in the Swedish cluster, the #nowplaying and #spotify hashtags have peripheral positions. #svpol acts as a global hub being directly connected to 253 of the top 258 hashtags, however, local hubs are difficult to discern. Overall, the network of these hashtags is fairly dense. The 258 hashtags shares 6,589 connections, giving the network a density of 0.199. The clustering coefficient, i.e. the probability that any two hashtags co-occur with a given hashtag, are also co-occurring with each other (Easley & Kleinberg 2010), was 0.562. A random network with the same number of nodes and a similar number of edges has a clustering coefficient of 0.1.

This indicates that the hashtag network is far more likely to cluster together than the random network. Neighbourhood size measure how many nodes (hashtags) are reachable within a given number of steps (i.e.

depth) from each node. The average neighbourhood size at depth 2 is 252 and the median is 254. This means that the average hashtag is connected to 252 of the other hashtags, either directly or through its direct connections.

The neighbourhood overlap measures the number of hashtags co-occurring with both of a pair of co-occurring hashtags divided by the number of hashtags co-occurring with one of these hashtags (e.g. Easley & Kleinberg 2010). In this network, the neighbourhood overlap is 0.275 which can be compared to 0.109 in the random network. This suggests that these dominating topics are closely inter-related and that many hashtagged exchanges are intertwined with each other. As the figures reveal there is a fuzzy border between the clusters (top part of fig. 5 and bottom part of fig. 6). Next to all hashtags below this border are written in Swedish while most of the hashtags above the border are in English.

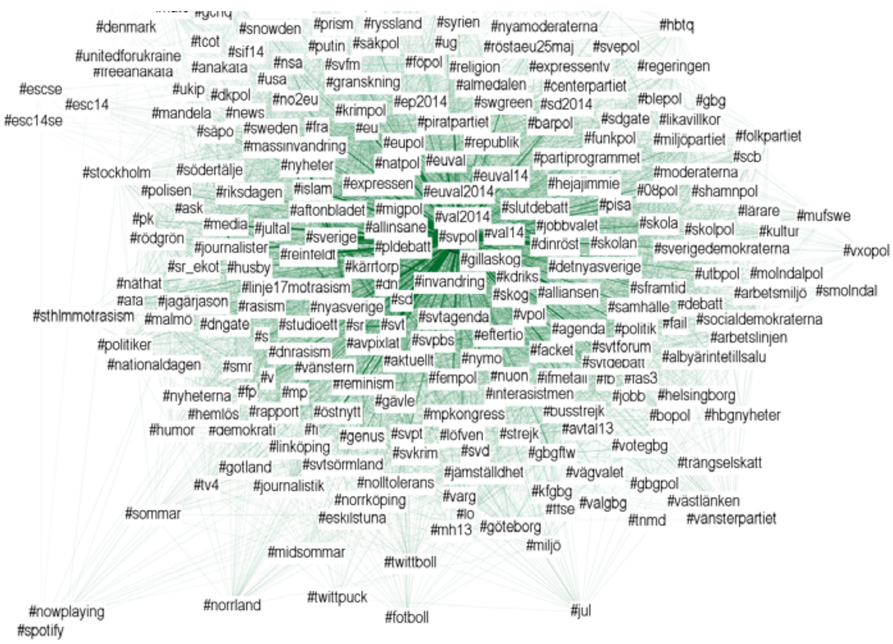

Figure 5. Domestic cluster, hashtags co-occurring in at least 100 tweets.

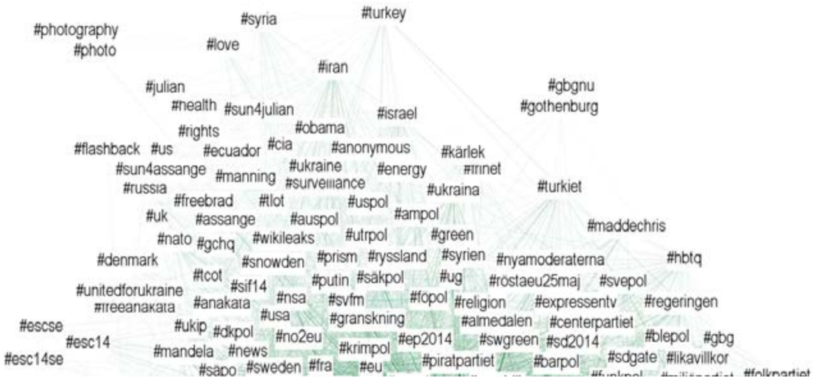

Figure 6. International cluster, hashtags co-occurring in at least 100 tweets.

Interaction with other users

In the sample of 300,000 tweets, 77,380 were replies and 46,692 were retweets posted by 899 and 908 elite users, respectively. Similar to Ausserhofer and Maireder's (2013) findings, there was much interaction between the sample group and other users. A large share of all replies (65.44%) was sent to a total of 10,142 users outside the elite group and, similarly, a large share of retweets (69.24%) from the elite users redistributed tweets from 10,289 non-elites. However, this non-elite group is so large that the replies and retweets per user are still fairly small (4.99 and 3.14). In comparison, the average elite user had 86.07 replies and 51.42 retweets received.

Figures 7 and 8 depict replies and retweets posted by elite users. In these networks, the number of received replies or retweets was used for node sizes (weighted in-degree). These networks represent the entire sample of replies and retweets, and filtered versions with the top 1,000 and 100 users according to messages received. The reply networks were comprised of one main cluster in which several elite users are prominent, while the retweet networks were clustered around a few dominating hubs. Although elite users certainly were dominant in both networks, the non-elite users were not just confined to the periphery as there were quite a few examples of non-elite users with central positions. This indicates that users outside the sample of elite users can act as gatekeepers or opinion leaders. Overall, members of the elite were more often replying than retweeting but still replied to a similar amount of users as they retweeted (30,704 and 30,966, respectively). A similar behaviour among the most active percentile was seen in Lorentzen (2014).

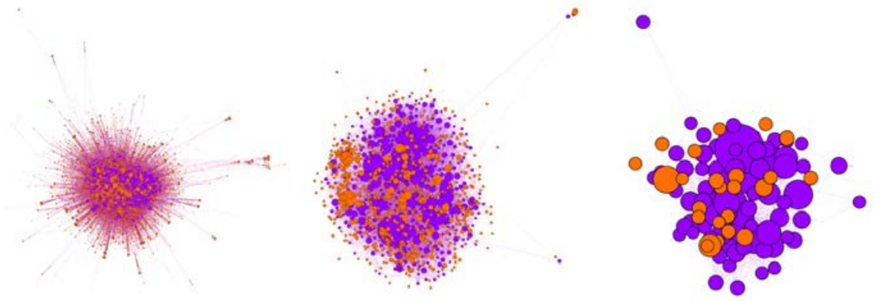

Figure 7. Replies sent by elite users (purple). Left: full network, centre: top 1,000 users, right: top 100 users.

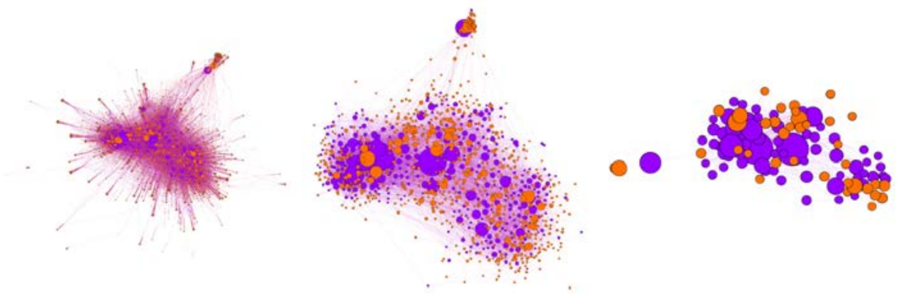

Figure 8. Retweets posted by elite users (purple). Left: full network, centre: top 1,000 users, right: top 100 users.

Obviously, there is only so much a small set of elite users can do with regards to replying and retweeting the larger set of non-elite users. 37 of them were amongst the 100 most replied to, and the corresponding figure for the retweet network was 22. However, the vast majority of non-elites were given little attention. Slightly more than one tenth of the non-elites (1,213 and 1,263 respectively) received at least ten replies or retweets. 115 non-elite users received at least 50 replies and 20 received 100 replies or more. Similar figures were found in the retweet network as 130 non-elite users received 50 retweets and 23 of them were retweeted at least 100 times. Even though this analysis has indicated that the elite users do send messages to and redistribute messages sent by other users than themselves, most of the attention was given to fellow elite users and a few non-elites.

Discussion and conclusions

This paper aimed to map the topics discussed of a given set of elite Twitter users, and to outline the extent to which they communicated with other users. By tracking a large set of political oriented users over a longer time period, evolution of topics and hashtags could be identified. In total, 154,993 Twitter users participated in the conversations. The vast majority of these did so by either replying to, or retweeting, a tweet posted by a member of these elites. Given the research design and the purpose of the study, a bias towards political conversations was introduced. It is no surprise that these users are more interested in discussing politics, but other topics were also found. Overall, hashtags related to whistleblowers and Wikileaks were frequently used. There were a few examples of stable hashtags representing migration politics and mass media while some hashtags reflected sudden or scheduled events. These are characterised by a large volume of tweets from many users, and a dramatic decrease in volume after the event. #snowden was such a hashtag, although it reappeared later on in the conversations. Other hashtags were related to TV shows (party leaders' debate), protests and elections, the latter type increasing in usage in conjunction with the EU election. During the first period, the newspaper Aftonbladet, the party leader debate and whistleblowers were among the top topics. The second period did not have any peaking hashtags. The third period was dominated by election related hashtags.

There were a few differences between the usage of hashtags by the tracked users compared to what was retweeted by their followers. Some hashtags that did not make it into the top ten list of the elite usage were highly retweeted. The passing away of Nelson Mandela and the demonstrations in Kärrtorp were such topics. The hashtags reflecting whistleblowers and Wikileaks and topics such as national security were also more visible in the full set. There were some similarities as well; including the usage of hashtags related to whistleblowers and Wikileaks, and, the stable levels of discussions around migration politics and the election hashtags.

Looking at how hashtags were connected to each other, there was a quite dense cluster around the most prominent hashtag #svpol. The co-occurrence hashtag network revealed an international cluster and a domestic cluster, but the border between these was very fuzzy. The most common topics in pairs were #svpol, and hashtags representing migration politics and the elections. #assange, #snowden and #wikileaks were all present in different constellations. "Your vote" co-occurring with election tags, and the "new Sweden" combined with the current Prime minister were prominent pairs. Other popular topics were migration, security politics and mass media. It was however somewhat surprising to see migration politics and related topics decrease during the third period, considering the topics being major during the election campaigns. The results also indicate that banal activity can still exist alongside more dominant topics, though seemingly being quite peripheral.

The Twitter community did respond to major events, both when considering the inner group and the snowballed group. Offline events tended to spark actions from more users, even the peripheral ones, who seemed to be more active following, or during, an event. This study found a few examples of this. The death of Nelson Mandela, the release of reports by Edward Snowden, the Kärrtorp protest, the PISA report, and the election were all followed by an increase in retweets and a decrease in mentions. This study has also shown that different topics emerge as prominent if retweets are included compared to when they are not. These differences highlight an important methodological aspect; when trend analyses are made researchers should investigate what is broadcasted and what is redistributed. Something not utilised here, but would be of interest, is to ground the findings in trending topics or with other sources on the web (e.g. Rogers 2013b). The findings also open up for more qualitative questions such as tweeter motivation, perceived experience of reader, both lurker and producer.

A very plausible argument against these findings is that the study is based on what seems like the tip of the iceberg; the elite of political tweeters. However, as other studies have found, these users account for a very large share of the tweets and this study found that with their activity, they engage many more users in the conversational exchanges. As so many tweets are produced by these potential opinion leaders it is very difficult for a human reader or follower of the conversations to assess whether the conversations are dominated by a few or many users, or how representative the opinions presented on Twitter are when considering the whole population.

This paper made use of the concept opinion leadership, however, it did not investigate to what extent or how these potential opinion leaders affected the opinions of their followers. Given previous findings indicating that Twitter users prefer to retweet like-minded (e.g. Conover et al. 2011; Lorentzen 2014) and that tweets expressing emotionality are more likely to be retweeted (Dang-Xuan et al. 2013), this approach would benefit from being complemented with analysis of who the users are and sentiment analysis of tweets. Another relevant aspect is the connections between local, national and global clusters. Is there for example a specific type of users bridging the national and global clusters?

While the identified elite of around 1,000 users did account for a very large share of the tweets within a topic such as domestic politics, the interaction analysis suggests that perhaps the elite is larger than that. With quite large shares of non-elite users among the top 100 in the reply and retweet networks it might be so that there are more users that could be categorised as elite users. Another explanation could be that elite users outside of this sphere were introduced into the dataset through replies or retweets when elite users tweeted about other topics than politics. The most likely explanation, though, is that during the year in which data were collected, new elite users emerged alongside the tracked ones.

Twitter seems to be elite centred, at least in this context, but Twitter research is also elite centred with a common focus on top 1,000 users within a given domain. While a large share of all content is produced by around 1,000 users, it would be relevant to also study the most common users. This study took one step out of the "elite bubble" as it considered what the cluster around the elite amplified and if the elite interacted with other users. This type of research also opens up other interesting research ideas. By adjusting the data collection method to follow the elite and the most active users around these, conversations initiated by the elite, and conversations the elite participate in, can be collected. We can then see how the elite and non-elite interact to a larger extent and so widen our perspective on Twitter activity.

Finally, if we want to fully understand how people use a platform such as Twitter, the event-driven behaviour needs to be studied as well. However, with data being collected in real time, it is almost impossible to completely track a sudden event. One idea is to utilise the Twitter Trends APIs to detect sudden events, but the issue here is how quickly something can become a trend. A possible option is to shift the focus from the elite users to the less active users. As they have been found to have an event-driven behaviour, how the Twitter community reacts to sudden events is perhaps best studied by tracking the activities of these users.

David Gunnarsson Lorentzen (PhD in library and information science) is a senior lecturer and a member of the Social Media Studies research group at the Swedish School of Library and Information Science. His main research interests concerns social media studies with a focus on method development.

Contact: david.gunnarsson_lorentzen@hb.se

References

AUSSERHOFER, JULIAN & AXEL MAIREDER (2013). "National Politics on Twitter." Information, Communication & Society 16.3: 291-314.

BARBERÁ, PABLO & GONZALO RIVERO (2015). "Understanding the Political Representativeness of Twitter Users." Social Science Computer Review 33.6: 712-729.

BASTIAN, MATHIEU, SEBASTIEN HEYMANN & MATHIEU JACOMY (2009). "Gephi: An Open Source Software for Exploring and Manipulating Networks." IC-WSM 8, May 17-20, 2009. San José, California: AAAI. 361-362.

BASTOS, MARCO TOLEDO, RAFAEL L. G. RAIMUNDO & RODRIGO TRAVITZKI (2013). "Gatekeeping Twitter: Message Diffusion in Political Hashtags." Media Culture Society 35.2: 260-270.

BLONDEL, VINCENT D. ET AL. (2008). "Fast Unfolding of Communities in Large Networks." Journal of Statistical Mechanics: Theory and Experiment 10: P10008.

BORRA, ERIK & BERNHARD RIEDER (2014). "Programmed Method: Developing a Toolset for Capturing and Analyzing Tweets." Aslib Journal of Information Management 66.3: 262-278.

BOYD, DANAH, SCOTT GOLDER & GILAD LOTAN (2010). "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter." In Proceedings of the 43rd Hawaii International Conference on System Sciences, January 5-8, 2010. Kauai, Hawaii: IEEE. 1-10. <http://www.danah.org/papers/TweetTweetRetweet.pdf> [2017-09-10]

BRUNS, AXEL (2012). "How Long is a Tweet? Mapping Dynamic Conversation Networks on Twitter Using Gawk and Gephi." Information, Communication & Society 15.9: 1323-1351.

BRUNS, AXEL & JEAN BURGESS (2012). “Researching News Discussion on Twitter.” Journalism Studies 13.5: 801-814.

BRUNS, AXEL & JEAN BURGESS (2011). “#ausvotes: How Twitter Covered the 2010 Australian Federal Election.” Communication, Politics & Culture 44.2: 37-56.

BRUNS, AXEL & TIM HIGHFIELD (2013). “Political Networks on Twitter.” Information, Communication & Society 16.5: 667-691.

BRUNS, AXEL & YUXIAN E. LIANG (2012). ”Tools and Methods for Capturing Twitter Data During Natural Disasters.” First Monday 18.4.
<http://firstmonday.org/ojs/index.php/fm/article/view/3937/3193> [2017-09-10]

BRUNS, AXEL & HALLVARD MOE (2013). “Structural Layers of Communication on Twitter.” Twitter and Society. Eds. K. Weller et al. New York: Peter Lang. 15-28.

BRUNS, AXEL & STEFAN STIEGLITZ (2013). “Metrics for Understanding Communication on Twitter.” Twitter and Society. Eds. K. Weller et al. New York: Peter Lang. 69–82.

CALLON, MICHAEL, JOHN LAW & ARIE RIP (1986). “Qualitative Scientometrics.” Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World. Eds. M. Callon, J. Law & A. Rip, A. Basingstoke: Macmillan: 104-123.

CONOVER, MICHAEL ET AL. (2011). “Political Polarization on Twitter.” In Proceedings of the 5th International Conference on Weblogs and Social Media, Barcelona, Spain, July 17-21, 2011. Menlo Park, California: AAAI. 89-96.

D’HEER, EVELIEN ET AL. (2017). “What Are We Missing? An Empirical Exploration in the Structural Biases of Hashtag-Based Sampling on Twitter.” First Monday 22.2. <http://firstmonday.org/ojs/index.php/fm/article/view/6353/5758> [2017-09-10]

DANG-XUAN, LINH ET AL. (2013). “An Investigation of Influentials and the Role of Sentiment in Political Communication on Twitter During Election Periods.” Information, Communication & Society 16.5: 795-825.

DRIEGER, PHILIPP (2013). “Semantic Network Analysis as a Method for Visual Text Analytics.” Procedia - Social and Behavioral Sciences 79. 4-17.

DUBOIS, ELIZABETH & DEVIN GAFFNEY (2014). "The Multiple Facets of Influence: Identifying Political Influentials and Opinion Leaders on Twitter." American Behavioral Scientist 58.10: 1260-1277.

DYAGILEV, KIRILL & ELAD YOM-TOV (2014). "Linguistic Factors Associated With Propagation of Political Opinions in Twitter." Social Science Computer Review 32.2: 195-204.

EASLEY, DAVID & JON KLEINBERG (2010). Networks, Crowds, and Markets: Reasoning about a Highly Connected World. Cambridge: Cambridge University Press.

GERLITZ, CAROLIN & BERNHARD RIEDER (2013). "Mining One Percent of Twitter: Collections, Baselines, Sampling." M/C Journal 16.2. <http://journal.media-culture.org.au/index.php/mcjournal/article/view/620> [2017-09-10]

GONZÁLEZ-BAILÓN, SANDRA ET AL. (2011). "The Dynamics of Protest Recruitment through an Online Network." Scientific Reports 1. Article number: 197. doi:10.1038/srep00197.

GONZÁLEZ-BAILÓN, SANDRA ET AL. (2014). "Assessing the Bias in Samples of Large Online Networks." Social Networks 38: 16–27.

GRAHAM, TODD ET AL. (2013). "Between Broadcasting Political Messages and Interacting with Voters." Information, Communication & Society 16.5: 692–716.

HAWTHORNE, JOSHUA, J. BRIAN HOUSTON & MITCHELL S. MCKINNEY (2013). "Live-Tweeting a Presidential Primary Debate: Exploring New Political Conversations." Social Science Computer Review 31.5: 552-562.

HE, QIN (1999). "Knowledge Discovery through Co-Word Analysis." Library Trends 48.1: 133-159.

HIGHFIELD, TIM, STEPHEN HARRINGTON & AXEL BRUNS (2013). "Twitter as a Technology for Audiencing and Fandom: The #Eurovision Phenomenon." Information, Communication and Society 16.3: 315-339.

HOLMBERG, KIM & MIKE THELWALL (2014). "Disciplinary Differences in Twitter Scholarly Communication." Scientometrics 101: 1027-1042.

HUANG, JEFF, KATHERINE M. THORNTON & EFTHIMIS N. EFTHIMIADIS (2010). "Conversational Tagging in Twitter." In Proceedings of the 21st ACM Conference on Hypertext and Hypermedia, Toronto, Canada, June 13-16, 2010. New York: ACM. 173-178.

JUNGHERR, ANDREAS & PASCAL JÜRGENS (2014). "Through a Glass, Darkly: Tactical Support and Symbolic Association in Twitter Messages Commenting on Stuttgart 21." Social Science Computer Review 32.1:74-89.

KLINGER, ULRIKE & JAKOB SVENSSON (2015). "The Emergence of Network Media Logic in Political Communication: A Theoretical Approach." New Media & Society 17.8: 1241-1257

KRUIKEMEIER, SANNE (2014). "How Political Candidates Use Twitter and the Impact on Votes." Computers in Human Behavior 34: 131-139.

LAMBIOTTE, RENAUD, J-C. DELVENNE & MAURICIO BARAHONA (2009). "Laplacian Dynamics and Multiscale Modular Structure in Networks." arXiv. <http://arxiv.org/abs/0812.1770> [2017-09-10]

LARSSON, ANDERS O. (2015). "Comparing to Prepare: Suggesting Ways to Study Social Media Today – and Tomorrow." Social Media + Society 1.1: 1-2.

LARSSON, ANDERS O. (2014). "Everyday Elites, Citizens, or Extremists? Assessing the Use and Users of Non-Election Political Hashtags." Mediekultur, Journal of Media and Communication Research 30.56: 61-78.

LARSSON, ANDERS O., & HALLVARD MOE (2014). "Triumph of the Underdogs? Comparing Twitter Use by Political Actors During Two Norwegian Election Campaigns." SAGE Open 4.4: 1-13.

LARSSON, ANDERS O. & HALLVARD MOE (2012). "Studying Political Microblogging: Twitter Users in the 2010 Swedish Election Campaign." New Media & Society 14.5: 729-747.

LINDGREN, SIMON & RAGNAR LUNDSTRÖM (2011). "Pirate Culture and Hacktivist Mobilization: The Cultural and Social Protocols of #WikiLeaks on Twitter." New Media & Society 13.6: 999-1018.

LORENTZEN, DAVID G. (2016). “Twitter Conversation Dynamics of Political Controversies: The Case of Sweden’s December Agreement.” Information Research 21.2: paper SM3. <http://InformationR.net/ir/21-2/SM3.html> [2017-09-10]

LORENTZEN, DAVID G. (2014). “Polarisation in Political Twitter Conversations.” Aslib Journal of Information Management 66.3: 329-341.

LORENTZEN, DAVID G., & JAN NOLIN (2017). “Approaching Completeness: Capturing a Hashtagged Twitter Conversation and Its Follow-on Conversation.” Social Science Computer Review 35.2: 277-286.

MERRY, MELISSA K. (2013). “Tweeting for a Cause: Microblogging and Environmental Advocacy.” Policy & Internet 5.3: 304-327.

MURTHY, DHIRAJ (2012). “Towards a Sociological Understanding of Social Media: Theorizing Twitter.” Sociology 46.6: 1059-1073.

MURTHY, DHIRAJ (2013). Twitter: Social Communication in the Twitter age. Cambridge: Polity.

PEARCE, WARREN ET AL. (2014). “Climate Change on Twitter: Topics, Communities and Conversations about the 2013 IPCC Working Group 1 Report.” PLoS ONE 9.4. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0094785> [2017-09-10]

ROGERS, EVERETT M. (2003). Diffusion of Innovations. (5. ed.) New York: Free press.

ROGERS, RICHARD (2013a). “Debanalizing Twitter: The Transformation of an Object of Study.” In Proceedings of the 5th Annual ACM Web Science Conference, May 2-4, 2013, Paris, France. New York: ACM. 356-365.

ROGERS, RICHARD (2013b). Digital Methods. Cambridge, Massachusetts: The MIT Press.

SHOEMAKER, PAMELA J. & TIMOTHY VOS (2009). Gatekeeping Theory. New York: Routledge.

SMALL, TAMARA A. (2011). "What the Hashtag? A Content Analysis of Canadian Politics on Twitter." Information, Communication & Society 14.6: 872-895.

SUH, BONGWON ET AL. (2010). "Want to Be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network." In Proceedings of IEEE International Conference on Social Computing, Minneapolis, Minnesota, August 20-22. Los Alamitos, California: IEEE. 177-184.

TUMASJAN, ANDRANIK ET AL. (2011). "Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape." Social Science Computer Review 29.4: 402-418.

TWITTER, INC (2015). Twitter Developer Documentation. Streaming API Request Parameters. <https://dev.twitter.com/streaming/overview/request-parameters> [2017-09-10]

VAN DIJCK, JOSÉ (2013). The Culture of Connectivity: A Critical History of Social Media. New York: Oxford University Press.

VAN DIJCK, JOSÉ & THOMAS POELL (2013). "Understanding Social Media Logic." Media and Communication 1.1: 2-14.

WEINBERGER, DAVID (2011). Too Big to Know: Rethinking Knowledge now that the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room. New York: Basic Books.

WU, SHAOMEI ET AL. (2011). "Who Says What to Whom on Twitter." In Proceedings of the 20th International Conference on World Wide Web, Hyderabad, India, March 28-April 1, 2011. New York: ACM. 705-714.

XU, WEIAI WAYNE ET AL. (2014). "Predicting Opinion Leaders in Twitter Activism Networks: The Case of the Wisconsin Recall Election." American Behavioral Scientist 58.10: 1278-1293.

ZIMMER, MICHAEL & NICHOLAS JOHN PROFERES (2014). "A Topology of Twitter Research: Disciplines, Methods, and Ethics." Aslib Journal of Information Management 66.3: 250-261.